UCLouvain

LOUVAIN
School of Management

Louvain School of Management

Portfolio Optimization with Functional Return Target

Author : Alexander ACKAERT
Supervisor : Frédéric VRINS
Reader : Nathan LASSANCE
Academic year : 2018-2019

As far as the laws of mathematics refer to reality, they are not certain, and as far as they are certain, they do not refer to reality

- Albert Einstein

Acknowledgements

Firstly, I would like to express my sincere gratitude to my advisor Pr. Frédéric Vrins for his continuous support, patience, and immense knowledge. He allowed me to challenge myself on a topic that was tremendously interesting. His guidance helped me during the time of research and writing of this thesis. I also would like to thank him for the classes he taught me during my Master years. His passion and clear explanations provided me with even more passion about what I studied.

I also wish to express my sincere thanks to Nathan Lassance. Whenever I ran into a trouble spot or had questions about my research, writing or coding, his door was always open. His expertise and research questions have been a source of motivation and inspiration for me. I really wish him the best for his future.

Moreover, I would like to thank all my family. My mother and stepfather for their constant encouragements, support, help and the opportunities, they enabled me to live during my studies. To my father, who has always push me to find interest in everything and to think big in life. I also would like to thank him for all the prints he made for me during those years. Finally, a special thanks to my brothers and sister, who have always been ready to listen to me and to be there for a nice break.

Last but not least, I would like to thank Camille for what you have brought to my life since we met. You never stopped supporting me and always pushed me to reach my goals. Far from being easy, you also supported me during the redaction of this thesis and you even tried to find some interest in it even though the subject is far from your liking. You will never stop amazing me with your kindness and your wit as well as your ability to always bring out the best of me. I can't wait to start our new life together in September.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The choice of asset allocation among securities is an important issue for asset managers or institutions. Standard approaches in portfolio selection determine the optimal asset allocation from a set of investor's constraints. Markowitz [1952] laid the foundations of portfolio selection using mathematical methods in order to find the best trade-off between the portfolio mean and variance.

Mean-variance analysis received many critics for being sensitive to estimation of the inputs and especially estimation errors in the mean. Consequently, many studies focus on the minimum variance portfolio which only includes the estimation of the variance-covariance matrix. In order to improve these estimations, well-known approaches such as the shrinkage estimator Ledoit and Wolf [2004] or robust estimation of DeMiguel and Nogales [2009] have been developed.

From an investor's perspective, the mean-variance portfolio selection also presents another drawback. It does not take into account specific preferences of an individual investor even though every investor has different preferences towards his investments. Scott and Horvath [1980] showed that investors exhibit positive

preferences for odd moments and negative preferences for even moments. For this reason, practitioners looked for higher-moments portfolio strategies. Several approaches exist but none of them is a clear leader. Martellini and Ziemann [2010] based their approach on Taylor's expansion to approximate the investor's utility function via the third and fourth moments. This approach, often based on a constant relative risk aversion parameter, is only impacted by higher moments when excessive value for the risk parameter is taken. Another way for managing higher moments has been developed by Favre and Galeano [2002] where an investor attempts to minimize the modified value-at-risk. Nevertheless, this methodology is also subject to estimation error as shown by Lim et al. [2011].

In this context, Lassance and Vrins [2019] rethink the current paradigm and introduced a new methodology that better accounts for investor's preferences and higher moments in portfolio optimization.

They investigate a new framework based on a functional approach where an investor targets a full distribution instead of a finite number of statistics. These finite statistics being characteristics captured in the probability density function. Concretely, the framework consists of minimizing the distance between a target-return density and the portfolio-return density. This allows for a full flexibility in terms of target and reduces the risk of estimation error in the inputs. In order to choose a reasonable target, they propose to fix the first two moments of the target according to a reference portfolio respectively known as the minimum variance portfolio or the maximum Sharpe ratio portfolio. To quantify the distance between these densities, they rely on the Kullback and Leibler [1951] divergence measure. This measure quantifies the amount of information that is lost when approximating a distribution by another.

Lassance and Vrins [2019] demonstrated that the minimum divergence portfolio

yields to similar mean-variance trade-off than the reference portfolio combined with an improvement in higher orders.

Having outlined the context and the practical relevance of the subject, we propose to investigate the following questions :

Does the minimum divergence portfolio outperforms out-of-sample portfolio selection strategy based on higher moments?

This research question also raises several interrogations. The cornerstone of this framework is on how to specify a target distribution that make sense according to the financial world.

Does the minimum divergence portfolio enhance higher orders when the reference portfolio is switched to a constant relative risk aversion utility function, modified value-at-risk or equally weighted portfolio?

How does the choice of the different moments of the target impact the Kullback-Leibler function?

What is the impact of the choice of the divergence measure? Comparison with the Rényi divergence and analysis of the α parameter.

After having introduced the context of our work and the research questions, let us look at the different chapters that structure the remainder of this thesis.

Chapter 2 First, we laid the foundation of portfolio theory by introducing the Modern Portfolio Theory of Markowitz [1952]. We describe the main concepts used in this framework and then we dedicate a section to the empirical results and drawbacks of this theory. We also present portfolio optimization strategies that account for higher orders.

Chapter 3 In this chapter, we present the minimum divergence portfolio theory.

We explain more in detail how this strategy can be implemented and its key features.

Chapter 4 The goal of this chapter is the description of the methodology we followed to study the minimum divergence portfolio. Firstly, we describe the data we used and provide some statistical insight. Secondly, we develop the structure of our analysis.

Chapter 5 In this chapter, we detail our results and answer the questions that have been raised.

Chapter 6 This final chapter aims to conclude this thesis. We point out implications and underline some limitations of our study to conclude with some suggestions for future research.

Chapter 2

Modern Portfolio Theory

Markowitz [1952] proposed a revolutionary approach for portfolio selection in his paper "Portfolio Selection" published by the Journal of Finance. Even though many criticisms have been made on his theory, Markowitz is considered as the father of portfolio theory and is still studied by every student in finance. In this section, we will first explain the approach of the Modern Portfolio Theory, which represents the sole basis of portfolio analysis then highlight the limitations and problems encountered by this framework in the real world.

2.1 Mean-Variance and Global Minimum Variance Portfolio

The portfolio selection process can mainly be separated into two steps. The first one requires an investor to select the potential assets in which he could invest and secondly, to decide, following a strategy or not, how much to invest in these assets. Every portfolio is subject to two kind of risks respectively known as the

systematic risk and the non-systematic risk. Systematic risk can be eliminated or highly reduced by the concept of diversification while non-systematic risk can not be reduced and is often called the non-diversifiable risk Zvi et al. [2014].

According to Markowitz [1952], an investor should not evaluate the asset's risk and return separately but on how these assets affect the overall portfolio's risk and return. For this reason, his approach relies on the expected rate of return, the risk of the individual assets and their interrelationship measured by the covariance or correlation. Basically, Markowitz's theory is built on the assumption that investors wealth can be illustrated by a quadratic utility function which will be presented later on in this thesis.

Suppose the financial world could be defined by N assets, then a portfolio could be composed of any asset N. Each asset in the portfolio has a weight noted w_i where each weight i represents the proportion allocated to the i-th asset in the portfolio. Analytically, the weights can be represented by an N-vector $w = (w_1, w_2, ..., w_N)$ and $\sum_{i=1}^{N} w_i = 1$.

The portfolio being composed of several assets, the actual return that is delivered depends on the weights allocation and the return of each asset. Assume that, asset returns $R = (R_1, R_2, ..., R_N)$ have expected rates of return $\mu = (\mu_1, \mu_2, ..., \mu_N)$. Since the portfolio is composed of several assets, the expected rate of return of the portfolio is the weighted average of the expected returns of the assets in the portfolio. The expected rate of returns being stochastic, the global return of a portfolio is given by a random variable R_P.

We have then:

$$R_P = \sum_{i=1}^{N} w_i R_i \tag{2.1}$$

where R_i are arithmetic returns. The expected return of the portfolio is, by linearity of the expectation, a weighted average of the expected returns on the component securities with portfolio proportions as weights:

$$\mu_p = \sum_{i=1}^{N} w_i \mu_i \tag{2.2}$$

For more simplicity, we can write these equations in a matrix form:

$$R_P = w^T R \tag{2.3}$$

$$\mu_p = w^T \mu \tag{2.4}$$

If the expected value of a RV presents a useful summary of the probabilistic nature of a variable, an investor would also like to have an intuition on the possible deviation from this expected value which could be considered as a risk measure. This latter is defined by the variance σ^2 and Markowitz used this measure to define the portfolio risk. When considering two or more assets, we also need to compute their mutual dependence and this can be achieved by the "covariance". The covariance between R_i and R_j is defined as:

$$\sigma_{ij} = \mathbb{E}[R_i R_j] - \mathbb{E}(R_i)\mathbb{E}(R_j) \tag{2.5}$$

Analogous to (2.5), there is an alternative shorter expression where we use the concept of correlation, denoted as ρ_{ij}, and we have:

$$\sigma_{i,j} = \rho_{i,j}\sigma_i\sigma_j \tag{2.6}$$

Two variables are said uncorrelated if $\rho_{ij} = 0$, meaning that knowing the infor-

mation of one value provides no information about the other. Two variables can also be positively correlated or negatively correlated which means respectively that if one variable is above the mean the other is likely to be above the mean and inversely.

From the description above, we can now define the portfolio variance as follows:

$$\sigma_P^2 = \sum_{i=1}^{N} w_i^2 \sigma_i^2 + 2 \sum_{i=1}^{N} \sum_{j \neq i}^{N} w_i w_j \sigma_{i,j} \tag{2.7}$$

We also can define the variance of r_i as σ_{ii} and thus,

$$\sigma_P^2 = \sum_{i=1}^{N} \sum_{j=1}^{N} w_i w_j \sigma_{i,j} \tag{2.8}$$

When dealing with many assets, notations and calculations can become increasingly complex and this is why we rely on matrix notation and calculation. In the matrix form, the $N \times N$ variance-covariance matrix between returns is defined as follows:

$$\Sigma = \begin{bmatrix} \sigma_{11} & \dots & \sigma_{1N} \\ \dots & \dots & \dots \\ \sigma_{N1} & \dots & \sigma_{NN} \end{bmatrix} \tag{2.9}$$

The portfolio variance is thus given by:

$$\sigma_P = w^{'} \Sigma w \tag{2.10}$$

It is obvious that for different combinations of w, the investor will attain different levels of μ_P and σ_P^2. In this framework, portfolios are considered being efficient if they are either risk minimal for a given return level or have the maximum return for a given level of risk. Portfolios are made by letting the weights w_i vary

across all possible combinations such that $\sum_{i=1}^{n} w_i = 1$. The set of points that correspond to portfolios is called the feasible set, feasible region or attainable set as in figure 2.1.

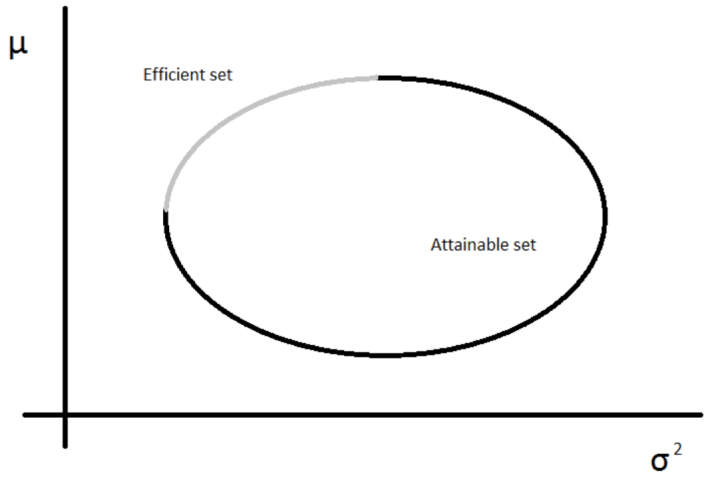

Figure 2.1: Attainable and Efficient portfolios in Markowitz

We are now in position to formulate a mathematical problem where the investor would like to maximize his wealth. It can be shown that this optimization takes the following form:

$$U(w) = \max_{w \in W} \mathbb{E}(R_P) - \frac{\gamma}{2}\mathbb{V}(R_P) = w'\mu - \frac{\gamma}{2}w'\Sigma w \tag{2.11}$$

Where R_P is the uncertain expected portfolio return, Σ is the variance-covariance matrix , μ is the vector of assets expected returns and γ the relative risk aversion parameter. However, this optimization is generally expressed in terms of risk

return trade-off and is given by the following equivalent formulation :

$$\begin{aligned} \min_{w} \quad & \sigma_P^2 = w^{'}\Sigma w \\ \text{subject to} \quad & w^{'}\mu = r \\ & w^{'}\mathbf{1} = 1 \end{aligned} \tag{2.12}$$

where $\mathbf{1}$ denotes the $(N \times 1)$ vector of ones and r a given level of portfolio return.

We only present the formulation where the objective is to minimize the portfolio's variance but we can also present it as the maximization of the expected return of the portfolio for a given level of risk. Both views are equivalent, even though the former is a quadratic optimization with linear constraints while in the latter, the objective function is linear and the constraints are quadratic.

The solution to this optimization can be derived using Lagrange multipliers and differentiate with respect to each variable and set this derivative to zero. Merton [1972] also used Lagrange multipliers and his framework is more frequently cited in the literature. We will partly present the solution developed by Merton :

$$w^* = \mu_p w_0^* + w_1^* \tag{2.13}$$

with

$$\begin{aligned} w_0^* &= \frac{1}{d}(c\Sigma^{-1}\mu - b\Sigma^{-1}\mathbf{1}) \\ w_1^* &= \frac{-1}{d}(b\Sigma^{-1}\mu - a\Sigma^{-1}\mathbf{1}) \end{aligned} \tag{2.14}$$

where $a = \mu^{'}\Sigma^{-1}\mu$, $b = \mu^{'}\Sigma^{-1}\mathbf{1}$ $c = \mathbf{1}^{'}\Sigma^{-1}\mathbf{1}$ and $d = ac - b^2$. The portfolio standard deviation is given by :

$$\sigma_P = \sqrt{\frac{1}{d}(c\mu^2 - 2b\mu + a)} \tag{2.15}$$

The combination that minimizes the variance for a specified expected return is called a "frontier portfolio". In fact, it is impossible for an investor to reach a higher expected return without bearing more risk. From equation(2.13), we notice that the portfolio weights are a linear function of the expected returns.

A particular point that attracts a lot of attention is the one at the bottom of the efficient frontier. This latter represents the portfolio with the smallest variance and is called the global minimum variance portfolio and can be easily computed. The latter is a specific portfolio often described as the apex of the hyperbola formed by equation (2.13). In fact, this equation forms an hyperbola enclosed by the asymptotes $\mu = b/c \pm \sqrt{\frac{d}{c}}\sigma$. The optimization function is given by :

$$\begin{aligned} \underset{w}{\text{minimize}} \quad & w^{'}\Sigma w \\ \text{subject to} \quad & w^{'}\mathbf{1} = 1 \end{aligned} \tag{2.16}$$

As mentioned above, the global minimum variance is the apex of the hyperbola and the weights corresponding to the apex are given by $w^*_{GMV} = \Sigma^{-1}\mathbf{1}/\mathbf{1}^{'}\Sigma^{-1}\mathbf{1}$. The reader will notice that in contrast with the mean-variance framework, the weights of this portfolio do not depend on the expected return of the assets.

2.1.1 Sharpe ratio

Another particular point located on the efficient frontier is also extensively studied. The point that maximizes the Sharpe [1994] ratio corresponds to the point where the ratio between the excess return (i.e. expected return minus risk-free rate) and the volatility is maximized. The Sharpe ratio is defined as follows :

$$SR = \frac{\mu_P - R_f}{\sigma_P} \tag{2.17}$$

where R_f represents the risk-free rate. The future returns for such assets are known with certainty today and are thus called risk-free. For the optimization process, we only have to define in the constraints the mean being equal to the Sharpe ratio to reach this point. An increasing Sharpe ratio outlines the fact that diversification of the portfolio is better when the new asset is hold in the portfolio or that the combination of weights is better in terms of diversification (and vice versa).

2.2 Assumptions of the Modern Portfolio Theory

Most theories are constructed under some assumptions and the mean-variance framework is no exception in that regard. The theory of the mean-variance is based on the following assumptions :

- Investors are risk-averse. To put it differently, for the same rate of expected return, they prefer the one with lower risk.
- Expected returns for all assets are known.
- The variances and covariances of all asset returns are known.
- Investors can ignore moments of order $k > 2$ if investors display quadratic utility function. In this framework, investors only need to know about the expected returns,variances, and covariances of returns to arbitrate the optimal portfolios.
- No transaction costs nor taxes.

2.3 Limits of Markowitz Mean-Variance Model

The groundbreaking theory of Markowitz has been the standard of asset allocation for decades. However, in practice, the mean-variance optimization has many shortcomings. Theses shortcomings focus mainly on two different areas. The first one being about the input of the model and the second one being about the output of the model. The main shortcomings could be resumed as follows :

- Extreme allocation and sensitivity : Indeed, the model relies on estimates for the value of μ_i and σ_i and as shown by Chopra and Ziemba [1993], small changes in the input would lead to substantial variations in the weights of the portfolio. Secondly, the output of the mean-variance portfolio often leads to extreme values like several large long and short positions that are not realistic in practice.
- Estimation of inputs : Kallberg [1984] showed that the estimation errors in the means are about ten times as important as estimation errors in variances and covariances. Nevertheless, estimation errors in variances and covariances can also have a considerable impact on the portfolio weights and results.

In the previous chapter, we have laid the foundation of the portfolio selection process. However, as previously mentioned, the use of Markowitz approach to define the asset allocation can be arguable. Indeed, the method is based on several assumptions that are not met in reality as well as an extreme sensitivity to small changes in the data. Therefore, a lot of researches have been made to enhance the process of portfolio selection.

In the following chapter, we will present a few evidences on stock returns and explain how they are in contradiction with Markowitz. Then, we will present some

risk measurements commonly used in the financial industry. Finally, portfolio optimization that accounts for higher moments will be addressed.

2.4 Expected Utility and Higher Moments

In the MV framework, an investor starts from a concave utility function U : $\mathbb{R} \to \mathbb{R}$ and seeks to maximize the expected utility of the portfolio defined as $R_P = \sum_{i=1}^{n} w_i R_i$ where w_i represents the portfolio weights and R_i the asset's returns. We have then :

$$w^* = \arg\max_{w \in W} \mathbb{E}[U(R_P)] \tag{2.18}$$

where W represents the set of all allowable portfolios. For infinitely differentiable utility functions, we can approximate it by a Taylor's expansion and this leads to :

$$\begin{aligned} \mathbb{E}[U(R_P)] = U(\mathbb{E}[R_P]) + \frac{U''\mathbb{E}[R_P]}{2}(\mathbb{E}[R_P - \mathbb{E}[R_P])^2 + \\ \sum_{k=3}^{\infty}(\mathbb{E}[R_P]\frac{U^K(\mathbb{E}[R_P])}{k!}(\mathbb{E}[R_P - \mathbb{E}[R_P]])^k \end{aligned} \tag{2.19}$$

where $(\mathbb{E}[R_P - \mathbb{E}[R_P]])^k$ is the kth-order centered moment. In the previous section, we only relied on the first two moments. In fact, as stated, when investors have a quadratic utility function we can ignore the moments of higher orders.

Theoretically, following Taylor's expansion, any k moments could have an impact on the utility function. We do not have specific intuitions for moments higher than fourth and can thus assume that the utility is well approximated by a fourth Taylor's expansion.

2.5 Higher orders and Stylized facts

Before going further with the optimization of a portfolio under higher moments, it is worthwhile to consider and review some typical characteristics of financial market data. Based on the paper of Cont [2001], a stylized fact is a typical characteristic of the stock returns that make their distribution different from a Gaussian distribution. Many stylized facts are observed such as :

- Time series are in general not independent and identically distributed
- Distributions of financial market returns are leptokurtic.
- Extreme events are observed closely in time, this is defined as volatility clustering.
- Distributions of financial market returns are mainly skewed to the left, meaning that negative returns are more likely to happen than postitive returns.

Indeed, Mandelbrot [1963] showed that asset returns display skewness and fat tails. Numerous stylized facts can be observed on the market but we will focus more specifically on the degree of asymmetry and the tailedness of the distribution. These concepts are known respectively as the skewness and the kurtosis Cont [2001].

The skewness and the kurtosis represent the third and fourth standardized moments of a probability distribution X and can be defined as follows :

$$\begin{aligned} S(X) &= \mathbb{E}\left[\left(\frac{X-\mu}{\sigma}\right)^3\right] \\ K(X) &= \mathbb{E}\left[\left(\frac{X-\mu}{\sigma}\right)^4\right] \end{aligned} \tag{2.20}$$

In other words, the first four moments represent the expected value of a specific integer power of the deviation of the random variable from the mean.

In a Gaussian world, the skewness and kurtosis of the distribution are respectively 0 and 3. On the contrary, stock returns commonly exhibit a negative skewness and a kurtosis that is higher than 3 meaning that extreme negative events are more likely to happen. Xiong and Idzorek [2011] demonstrated that asymmetry (i.e. third moment) and fat tails (i.e. fourth moment) are important in asset allocation and especially during crisis. Scott and Horvath [1980] also showed that risk-averse rational investors have preferences for odd moments and negative preferences for even moments.

2.6 Portfolio under Higher Moments

Based on the development in section 3.1, the portfolio choice is no longer a trade-off between the expected return and the variance. In this section, we present a portfolio optimization that accounts for higher moments. Martellini and Ziemann [2010], introduced a robust portfolio choice built on the higher order moment tensors of Jondeau and Rockfinger [2011]. In other words, these higher orders represent the estimation of all the co-moments i.e. covariance, coskewness and cokurtosis. As for the covariance, these co-moments provide an estimation on how the skewness and kurtosis move one from each other.

2.6.1 Portfolio moments

We now consider a random asset-return vector $X = (X_1, ..., X_n) \in \mathbb{R}^N$ of dimension $N \times 1$ where N is the number of assets available for investment. Let

$P := w'X$ be the portfolio return where $w = (w_1, ..., w_N)'$ is the $N \times 1$ vector portfolio weights. In order to stay in a matrix form, we use the Kronecker product ($\otimes$). The first fourth moments of the portfolio are determined by:

$$\begin{aligned} m_1 &= w^T \mu_X \\ m_2 &= w^T \Sigma_X w \\ m_3 &= w^T M_3 (w \otimes w) \\ m_4 &= w^T M_4 (w \otimes w \otimes w) \end{aligned} \tag{2.21}$$

where M_3 and M_4 are the co-skewness and the co-kurtosis of the asset-return vector X.

2.7 Portfolio Optimization

In this regard, the investor can set his allocation using the estimation of the higher orders. The following optimization is one possibility :

$$U(w) = \underset{w \in W}{\arg\max}\ m_1 - \gamma m_2 + \gamma m_3 \tag{2.22}$$

Alternatively,

$$U(w) = \underset{w \in W}{\arg\max}\ w^T \mu - \gamma w^T \Sigma w + \gamma w^T M_3 (w \otimes w) \tag{2.23}$$

where γ represents the risk aversion parameter. In general, the coefficient before the moments are different but we assess the optimization based on equal importance of each term. While this optimization takes the skewness into account, we also rely on Taylor's expansion developed in section 3.1. This methodology is based on the work of Martellini and Ziemann [2010] where we use a constant

relative risk aversion utility function. As proposed by Martellini and Ziemann [2010], we offset the large estimation risk of μ by setting it to a constant.

$$U(w) = \arg\min_{w \in W} \frac{\gamma}{2} m_2 - \frac{\gamma(\gamma+1)}{6} m_3 + \frac{\gamma(\gamma+1)(\gamma+2)}{24} m_4 \tag{2.24}$$

2.8 Risk Measurement for portfolios

2.8.1 Value-at-Risk

Within asset allocation strategies also arises the need to adequately measure risks that investors or institutions are facing. In practice, one of the most commonly used risk measure is the value-at-risk. The method was introduced by JP Morgan around 1990 but as stressed in Alexander and Baptista [2002], this sort of measure was already promoted by Baumol [1963]. The VaR is an estimate of how much a certain portfolio can lose within a given time period, for a given confidence level. Mathematically speaking, the VaR can be defined for a coverage ratio of α % where α corresponds to the quantile of the profit and loss distribution function over a given period of time :

$$VaR_t(\alpha) = -F_{R_t}^{-1}(\alpha) \tag{2.25}$$

Where $F_{R_t}^{-1}(.)$ is the quantile function defined as the inverse of the cumulative distribution function. Under a multivariate normal assumption for the asset returns, the portfolio's distribution is also represented by a normal. Consequently, the VaR of the portfolio is given by :

$$VaR(\alpha) = -w^T \mu - \sqrt{m_2} * G^{-1}(\alpha) \tag{2.26}$$

where $G^{-1}(\alpha)$ is the Gaussian quantile at the confidence level α

Nevertheless, recall from section 2.5 that return distributions are ordinarily not normally distributed and exhibit skewness and kurtosis which contradicts the former hypothesis.

For this reason, the need to use a risk measure that takes into account higher order is necessary. We rely on the modified value-at-risk developed as in Boudt et al. [2008]. The approximation of the distribution can be enhanced by readjusting it for higher moments and this can be done using a Cornish Fisher Expansion that is a method to transform a standard Gaussian random variable Z into a non Gaussian Z random variable by taking into account superior moments in its computation. We skip some mathematical developments which are developed in Boudt et al. [2008]. Using the latter, the modified value-at-risk of a portfolio return can be defined as :

$$\begin{aligned} MVaR_\alpha &= -w^{'}\mu - \sqrt{m_2}\ G_2^{-1}(\alpha) \\ G_2(\alpha) &:= q_\alpha + \frac{1}{6}(q_\alpha^2 - 1)S_P + \frac{1}{24}(q_\alpha^3 - 3q_\alpha)K_P - \frac{1}{36}(2q_\alpha^3 - 5q_\alpha)S_P^2 \end{aligned} \tag{2.27}$$

where $q_\epsilon = \Phi^{-1}(\epsilon)$ represents the quantile of the normal distribution. To put it simple, we modify the quantile of the normal distribution that takes into account the value of the skewness and the kurtosis. We also note that if the distribution is normal (i.e. skewness $= 0$, excess kurtosis $= 0$) we retrieve the Gaussian value-at-risk.

2.9 Modified Value-at-Risk portfolio

Accounting for higher moments also means that the investors or asset managers desire to "control" the tail risk they are facing in their portfolio. In this section,

we present a portfolio framework based on the work of Favre and Galeano [2002]. The goal of the authors is to minimize the value-at-risk of the portfolio but with a value-at-risk that accounts for skewness and kurtosis. With this aim, they rely on portfolio optimization under the modified value-at-risk. This latter depends on the mean of the portfolio but in this optimization, as in the CRRA utility function, we assume a constant mean of zero to overcome its large estimation risk. The optimization is presented as follows :

$$w^* = \arg\min_{w \in W} MVaR(\alpha) \tag{2.28}$$

with the $MVaR_\alpha$ being defined as follows :

$$MVaR_\alpha = -\sigma_P[q_\alpha + \frac{1}{6}(q_\alpha^2 - 1)S_P + \frac{1}{24}(q_\alpha^3 - 3q_\alpha)K_P - \frac{1}{36}(2q_\alpha^3 - 5q_\alpha)S_P^2] \tag{2.29}$$

where $q_\alpha = \Phi^{-1}(\alpha)$ represents the quantile of the normal distribution. Most investors are risk averse against extreme losses and thus using an $\alpha = 1\%$ will provide a fair portfolio with a high confidence level that losses will not exceed the modified value-at-risk. We also note that common values for α are 1% or 5%, however Cavenaile and Lejeune [2012] showed that for $\alpha > 4.16\%$ the modified VaR incorrectly reflects investors negative preferences for kurtosis.

2.9.1 Adjusted Sharpe Ratio

Given that higher moments are now taken into the portfolio selection techniques, the Sharpe ratio alone is not sufficient to assess the performance of the portfolio. Hence, we rely on the adjusted Sharpe ratio developed by Pézier [2004] that

accounts for higher orders and that is in line with investors preferences :

$$ASR := SR\left(1 + \frac{S_P}{3!}SR - \frac{K_P}{4!}SR^2\right) \tag{2.30}$$

2.10 Limits to Higher Moments

While these optimizations remain practicable and mathematically correct, several aspects of higher-moment restrain this way of doing.

- First, including higher-orders in the optimization process leads to complex non-convex optimization. This latter can bring out unreliable solutions because of the different sub-optimal solutions.
- Furthermore, estimations of all these higher moments are on the one hand complex and time-consuming when a lot of assets are included and on the other hand, the robustness of the solution is challenging. In fact, the MV framework is already sensitive to small changes in the input data, including higher orders exacerbates the sensitivity of the optimizations.

Chapter 3

Minimum Divergence Portfolio

This chapter introduces a new way for portfolio selection based on divergence measures. The framework introduces a portfolio selection process where the investor seeks the portfolio that is as close as possible to his ideal portfolio. This method allows an investor to define various constraints in the portfolio selection and to optimize a single objective function.

3.1 General Principal

- The first step consists in specifying the investor's target return probability density function that we define as being f_T. We note, T a random variable that is distributed according to f_T : $T \sim f_T$. In order for the method to be implemented, the density function that is defined embeds a set of J investor's constraints. Practically, one chooses $f_T \in \Omega_f$, with :

$$\Omega_f = \left[f \in L : \int \phi_j(x) f(x) dx = c_j \right] \neq \emptyset \tag{3.1}$$

where L represents all the probability densities defined on $\mathbb{R}$.

- In the second step, the optimal portfolio weights are determined such that the portfolio density, defined as f_P is as close as possible to the target density defined in the first step. The notion of distance - divergence - measure between the densities requires the use of a divergence measure. Denoting, $\langle f_P | f_T \rangle$ as divergence measure between P and T, the minimum divergence portfolio solves the following problem :

$$w^* = \arg\min_{w \in W} \langle f_P | f_T \rangle \tag{3.2}$$

The minimum divergence portfolio is thus a functional extension of the standard moments-based approach Lassance and Vrins [2019]. Indeed, instead of optimizing an objective subject to different constraints, this is all defined in a density function. Generally, the minimum divergence portfolio will not exactly comply with investor's constraints. In fact, investor's constraints must be in line with the reality of the market. Now, two main features need to be developed namely the divergence measure and how to specify the investor's constraints.

3.2 Divergence Measure

A central piece in this method is the use of a divergence measure between the investor's target and the portfolio return. The notion of distance between probability measures comes from Mahalanobis [1936]. Later, new methods arose from different authors. Naturally, we expect different results according to the measure that will be used. We first present the Kullback and Leibler [1951] divergence and then theRényi [1951] divergence measure.

3.2.1 Properties of Divergence Measure

As already stated, several divergence measures have been suggested in the statistical literature to express the fact that probability distributions are "close" to each other. Rao called these measures, measures of separation while Kolmogorov called it measures of variation-distance. Even though these measures have not been constructed with exactly the same objective, they have one common property: They all increase when the two distributions are "further away from each other" Pardo [2006].

Ullah [1996] presents some useful concepts of distance measure. However, the divergence measure only satisfies one criteria of a distance measure which is the positive definiteness. For, $D(f_P, f_T)$ the positive definiteness is defined as follows:

$$\begin{aligned} &(i)\ \textit{Positive definiteness} : D(f_P, f_T) \geq 0 \\ &(ii)\ D(f_P, f_T) = 0 \iff f_P = f_T \text{ everywhere} \end{aligned} \tag{3.3}$$

3.2.2 Kullback-Leibler Divergence

Chalabi and Wurtz [2014] investigated the impact of divergence measures in portfolio optimization where they present the ϕ divergence measure and explore the main advantages of using this divergence. In this section, we present the most common divergence measure developed by Kullback and Leibler [1951]. Suppose the two probability density function defined in section 3.1, f_P and f_T continuous with respect to x then the Kullback-Leibler divergence (KL) denoted as $D_{KL}(f_P, f_T)$ is defined as follows :

$$\langle f_P | f_T \rangle = D_{KL}(f_P, f_T) = \int f_P(x) \ln \left(\frac{f_P(x)}{f_T(x)} \right) \mathrm{d}x \tag{3.4}$$

where f_P is the portfolio density and f_T the target distribution. In simple words, the KL is a measure of how similar two probability distributions are. The KL was first used in information theory and measures the increase of entropy meaning the amount of information that is lost when approximating a function by another. In our framework, it means the amount of information that is lost when we approximate the target distribution by the portfolio's distribution. The measure is a pseudo-distance in the sense that it only satisfies the positive definiteness.

3.2.3 Rényi Divergence

The Rényi divergence was introduced by Rényi [1951] as a measure of information and depends on a parameter α called the order of the function. Therefore, the Rényi divergence of order $\alpha(\alpha > 0, \alpha \neq 1)$ is defined as follows :

$$\langle f_P | f_T \rangle = D_\alpha(f_P, f_T) = \frac{1}{\alpha - 1} \ln \int f_P(x) \left(\frac{f_P(x)}{f_T(x)} \right)^{\alpha - 1} \mathrm{d}x \tag{3.5}$$

The Rényi divergence extends the Kullback-Leibler divergence in the sense that the limit of the Rényi divergence when $\alpha = 1$ equals the Kullback-Leibler divergence :

$$\lim_{\alpha \to 1} D_\alpha(f_P, f_T) = D_{KL}(f_P, f_T) \tag{3.6}$$

3.3 Portfolio Density

In statistics, several methods exist to estimate the density of a distribution. One way of performing would be to consider f_P in the same family as f_T and using the maximum likelihood principle to find the parameters of the density. However, in terms of robustness, we see a main issue. Using parametric estimations implies we

make some assumptions about the shape of the portfolio returns which a priori we have no clue about it. Instead, we rely on non parametric Kernel Density as developed in Silverman [1986]. Generally shortened to KDE, is a technique to estimate a density according to some data. The Kernel density estimator is defined as follows :

$$f_P(x) = \frac{1}{Nh} \sum_{i=1}^{N} K\left(\frac{x - x_i}{h}\right) \tag{3.7}$$

where K represents a Kernel function and can be chosen and N the number of observations. We set the Kernel function as a Gaussian distribution. To put it differently, we put on every observation a Kernel function and then sum them to estimate the density. Non parametric methods do not make assumptions about the functional form of the density. By avoiding the potential functional form of the density, the range of possible shapes for the density is wider and could lead to more accurate fits.

3.3.1 Choice of the bandwidth

The choice of the bandwidth can have a major impact on the density. In case the bandwidth is too small, the smoothness of the density will be really low while a large bandwidth will over-smooth the function. We rely on Silverman [1986] rule-of-thumb which corresponds to the bandwidth that minimizes the integrated mean squared error. The bandwidth is defined as follows :

$$h = \left(\frac{4\hat{\sigma}}{3N}\right)^{\frac{1}{5}} \tag{3.8}$$

The parameter h varies with the portfolio weights, since the standard deviation will also vary depending on the weights. At first blush, this could unpredictably

affect the optimal portfolio obtained however it seems natural that the variance varies with weights of the portfolio. The other option is to fix the bandwidth as an equally weighted portfolio and then keep it constant but we opt for the first option.

3.4 Density Target

In this section, an approach to specify the target distribution will be developed. We follow the moment-matching principle and its specification can be achieved in two steps :

- 1. We first select a parametric family of densities $\Omega_{f\theta} := \{f|\theta \in \Theta\}$. The number of parameters Θ must equal the number of investor's constraints J meaning that if $J = 2$, the investor could use a Gaussian distribution for the density target.
- 2. Once the constraints are defined, we need to find $(\theta_1, ..., \theta_J) \in \Theta$ by moment-matching such that $\mathbb{E}[\phi_j(T(\theta))] = c_j, j = 1, ..., J$.

A main advantage of this approach is the flexibility that it offers in terms of distribution. In fact, a lot of parametric distributions exist and correctly fit the financial data. Theodossiou [1998] developed a skewed generalized T-student distribution that is highly flexible and can then be used for different cases.

3.4.1 The Gaussian Case

The first natural choice as target is the Gaussian distribution. It is the most studied distribution in probability and can be a good approximate for a lot of phenomena when the number of experiments increase. Choosing the Gaussian

distribution would lead an investor to only have 2 constraints. For example, if an investor has a Gaussian target with the following moments, $\mu_T = r$ and $\sigma_T = \sigma$, where r is a specific value of return, then target density is then defined as :

$$f_T(x) = \frac{1}{\sigma_T}\phi\left(\frac{x-\mu_T}{\sigma_T}\right) \tag{3.9}$$

where ϕ is the Gaussian probability density function. We remind that targeting a Gaussian distribution is implicitly setting the skewness and kurtosis preferences to 0 and 3 respectively.

3.4.2 Skew T-student Case

The Skew T-student distribution has four parameters and is very attractive in terms of distribution theory. First, it generalizes, among others, the skewed normal T-student and Gaussian distribution. Moreover, with four parameters an investor has a large flexibility in terms of targets. Arellano-Valle and Azzalini [2013] derived the parameterization of the distributions. If $T \sim ST(\mu_T, \sigma_T, \lambda_T, v_T)$ then f_T and its first four cumulants are as follows :

$$f_T(x) = \frac{2}{\sigma}d(z; v_T)D\left(\lambda_T z\sqrt{\frac{v_T+1}{v_T+z^2}}; v_T+1\right), \tag{3.10}$$

$$\begin{aligned}
m_1 &= b_v\delta \\
m_2 &= \frac{v_T}{v_T-2} - b_{v_T}^2\delta^2 \\
m_3 &= b_v\delta\left(\frac{3v_T}{(v_T-2)(v_T-3)} - \delta^2(\frac{v_T}{v_T-3} - 2b_v^2)\right) \\
m_4 &= \frac{3v_T^2}{(v_T-2)(v_T-4)} - 6\delta^2 b_v^2\frac{v_T({}_Tv-1}{(v_T-2)(v_T-3)} + \delta^4 b_v^2(\frac{4v_T}{(v_T-3)} - b_v^2)
\end{aligned} \tag{3.11}$$

where v_T must be larger than the order of the cumulant and :

$$\begin{aligned} \delta &= \frac{\lambda_T}{\sqrt{(1+\lambda_T^2)}} \\ b_v &= \sqrt{\frac{v_T}{\pi}}\frac{\Gamma(\frac{(v_T-1}{2})}{\Gamma(\frac{v_T}{2})} \end{aligned} \tag{3.12}$$

In equation (4.10), $z = \frac{(x-\mu_T)}{\sigma_T}$ and d and D are respectively the probability density function and cumulative distribution function of the standardized T-Student. Lambda is the parameter that influences the skewness. In fact when $\lambda_T < 0$, the distribution is left-skewed and inversely. The parameter v_T controls the kurtosis of the distribution. Lower values of v_T brings to more leptokurtic distribution. The first four moments are then defined as :

$$\begin{aligned} \mathbb{E}[T] &= \mu_T + \sigma_T m_1 \\ Var(T) &= \sigma_T^2 m_2 \\ S(T) &= \frac{m_3}{m_2^{\frac{3}{2}}} \\ K(T) &= \frac{m_4}{m_2^2} \end{aligned} \tag{3.13}$$

Since the Skew T-student has four parameters we can perform the moment-matching for the first four parameters. From equation (3.13), an investor can define the moments he wants to target according to his preferences. For example, if an investor would like to target the following moments : mean = 5.974%,standard deviation = 3.04%, skewness = 0.4 and kurtosis = 4, we just have to solve the non linear system to find the direct parameters of the distribution. In the example aforementioned, the parameters would be as follows : $\mu = 0.038$, $\sigma = 0.033$, $\lambda = 1.08$, $v = 12.083$. We also underline that not every combination of the four moments are possible. For example, with the same first two moments and a

skewness = 0.7 and kurtosis = 3.5, the density is unattainable. We rely on the R package *cp2dp* to transform centered moments into parameters.

3.5 Moments selection process

One of the main challenges of the minimum divergence portfolio is to choose the different targets of the portfolio. Naturally, every investor would prefer a return distribution with a high mean and right skewed as much as possible. Obviously, this method can not reach the unattainable, meaning that we have to deal with market features. For example, using the method in the treasury bond market and looking for a target with very high return provides inconsistent results.

To this end, we first rely on the approach proposed by Lassance and Vrins [2019] where the authors target a reference portfolio and then fix the first two moments of investor's target as the same as the reference portfolio. The first reference portfolio would be the mean-variance efficient portfolio which could be the minimum variance portfolio or the one that maximizes the Sharpe ratio.

This approach can be considered efficient for different reasons. First, it allows the investor to choose an achievable target in terms of return and volatility. Moreover, by using this method, the investor already chooses an optimal point in the sense that no other portfolio dominates the one targeted in terms of mean and variance.

We propose a new starting point based on the same method but where the reference portfolio is first the modified value at risk portfolio and then the portfolio based on a constant relative risk aversion utility function respectively defined in equation (2.27) and equation (2.24). This method also allows us to verify if first targeting a distribution instead of downside risk measure leads to different results and secondly to find out if targeting a distribution leads to an improvement in

the higher moments even when the reference portfolios are portfolios that already account for higher orders.

Finally, we also asses the equally weighted portfolio that is a naive but well-known strategy in investment management where an investor have the proportion of each stock in his portfolio. Investors are still using this type of strategy to allocate their wealth. Hence, the equally-weighted portfolio may appear a simple strategy to define the starting point of our minimum divergence portfolio but can be realistic. Besides, it has been proved by DeMiguel et al. [2009] that this simple strategy can outperform some other optimal allocation policies.

Chapter 4

Data and Methodology

In this chapter, we describe the methodology that has been pursued to conduct our study of the minimum divergence portfolio presented in the previous chapter and ultimately answer the questions that have been raised in chapter 1. To this aim, we will first describe the data that have been collected. Secondly, we will present the procedure that we followed in order to achieve our analysis. The majority of our analysis has been performed with Rstudio.

4.1 Description of Data

In this thesis, we based our analysis on two different data-sets from K. French's library : five industry U.S. portfolios and six portfolios of firms sorted by size and book-to-market. We collected a set of daily and monthly portfolio returns. Both data-sets are collected during the time-period 1978-2018 providing 40 years of data. The data-set sorted by size and book-to-market is collected on a monthly basis providing 492 observations. The other data-set sorted by industry is collected on a daily basis providing 10089 observations. We respectively call these

data-sets *6BTM* and *5Ind*. As mentioned in Jagannathan and Ma [2003], daily return provides better estimates than weekly or monthly returns. Nevertheless, for illustration and in our framework, testing monthly is also interesting.

The data-sets displayed in table 4.1 and table 4.2 clearly confirm the non-normality of the markets exposed in section 2.5. We can observe that the skewness of all markets are negative and all of them have an excess kurtosis higher than zero whereas daily returns exhibit high values of excess kurtosis.

4.1.1 Key Statistics

	Cnsmr	Manuf	HiTec	Hlth	Other
Minimum(%)	-17.28	-18.23	-18.89	-17.89	-15.25
Maximum(%)	9.88	14.52	14.25	11.10	11.27
Mean (annual in %)	13.79	12.87	12.76	14.49	13.08
Volatility (annual in %)	15.61	16.75	21.02	17.75	19.38
Skewness	-0.62	-0.63	-0.10	-0.50	-0.29
Excess kurtosis	16.24	21.88	11.13	11.63	13.98

Table 4.1: Descriptive statistics of *5Ind* based on daily observations. All values are expressed in daily terms except for the return and volatility that are expressed in annual terms

	Small LoBM	ME1 BM2	Small HiBM	Big LoBM	ME2 BM2	Big HiBM
Minimum(%)	-32.37	-27.98	-27.72	-23.18	-20.23	-22.27
Maximum(%)	27.08	16.64	17.29	14.44	13.4	17.66
Mean (annual in %)	11.13	15.49	16.37	12.38	12.07	12.9
Volatility (annual in %)	23.13	17.94	18.34	15.79	14.91	16.67
Skewness	-0.54	-0.89	-1.03	-0.50	-0.62	-0.82
Excess kurtosis	2.04	3.09	3.29	1.77	2.51	2.79

Table 4.2: Descriptive statistics of *6BTM* based on monthly observations. All values are expressed in monthly terms except for the return and volatility that are expressed in annual terms

4.2 Methodology

The proposed methodology is based first on the performance analysis of the different portfolio selection techniques. This is analysed through an in-sample and out-of-sample analysis. Then, we further analyse the impact of different targets moments on the Kullback-Leibler function. We continue by investigating if the shape of the target really matters in the minimum divergence portfolio framework. After , we compare the choice of the divergence measures and assess the impact of the α parameter in the Rényi divergence. Finally, we also develop the integration of the Kullback-Leibler divergence under a Gaussian quadrature. We now explain in detail the methodology :

In and out-of-sample analysis

The minimum divergence framework being relatively new, we evaluate the performance of the different considered portfolios, listed in 5.8, in and out-of-sample. This allows us to better understand how the minimum divergence portfolio behaves even though we clearly acknowledge that the out-of-sample is more important. To perform the out-of-sample analysis, we used a rolling window strategy meaning that we estimate the portfolio's weights on window τ smaller than N and then keep our weights constant under the next window. Keeping the weights constant is also called a buy-and-hold strategy.

Comparison of portfolios

Once the computations are done, we compare the Markowitz portfolios with portfolios under higher orders. This comparison is done by analysing the different moments of the portfolios and then the different indicators presented earlier in this thesis namely the modified value-at-risk, the Sharpe ratio and and the adjusted Sharpe ratio. After, we compare the Markowitz portfolios and portfolios under higher orders with the minimum divergence portfolios. We also compare the dif-

ferent obtained minimum divergence portfolios when the reference is switched to a constant relative risk aversion utility function or a modified value-at-risk respectively developed in equations 2.24 and 2.27. This allows us to verify if using other reference portfolio than the Markowitz portfolio leads to other results.

Sensitivity Analysis

For the sensitivity analysis, we analyse how the Kullback-Leibler function evolves according to target distributions that have three same moments and one moment that is changing. For example, we will target a Skew-T distribution with the same mean,variance and kurtosis but with different values of skewness. We will graphically analyse how the Kullback-Leibler function evolves.

Importance of the shape

To assess the importance of the target shape in the minimum divergence framework, we target Skew-T distributions that all have the same specific modified value-at-risk but with different values of parameters. This allows us to assess if the obtained results are similar or if the shape of the target influence the result. This will also advance the idea that targeting a specific measure and targeting an entire distribution does not lead to the same results.

Rényi and Kullback-Leibler Divergence

We also present the minimum divergence portfolio under the Rényi divergence measure. In this section, we compare the obtained result of the minimum divergence portfolio under the two divergence measures. We first see if the results differ and then analyse the effect of the α parameter in the Rényi divergence. In order to perform this analysis, different values of α will be chosen and results will be compared.

Kullback-Leibler integration and Gaussian quadrature

In order to approximate the integration of the Kullback-Leibler function, we present the Gaussian quadrature and assess if using the Gaussian quadrature leads to the same result than under the integration.

Optimization

All the computations of the portfolios are achieved by using the non-linear solver fmincon and DEoptim in r. The fmincon function in r is derived from the package pracma and is another way to write the function Nlcoptim.

Portfolio denominations

We list here below the different abbreviations of the portfolios that will be compared in this thesis :

Type	Portfolio	Abbreviation	Section
Traditional Portfolios	Minimum Variance Portfolio	MV	2.1
	Maximum Sharpe Ratio Portfolio	MSR	2.11
	Equally-Weighted Portfolio	EQ	3.5
Portfolios Under Higher Orders	Constant Relative Risk Aversion Portfolio($\lambda = 100$)	CRRA	3.4
	Modified Value-at-Risk Portfolio	M-VaRPort	3.5
	3 equal moments	3Mom	3.4
Minimum Divergence Portfolios	Gaussian target return and MV as reference portfolio	MDivGMV	4.5
	Gaussian target return and MSR as reference portfolio	MDivGSR	3.4.1
	Gaussian target return and M-VaR as reference portfolio	MDivGMVAR	3.4.1
	Gaussian target return and CRRA as reference portfolio	MDivGCRRA	3.4.1
	Gaussian target return and EQ as reference portfolio	MDivGEQ	3.4
	Skew-T target return and MV as reference portfolio	MDivSTMV	3.4.2
	Skew-T target return and MSR as reference portfolio	MDivSTSR	3.4.2
	Skew-T target return and MVaR as reference portfolio	MDivSTMVAR	3.4.2
	Skew-T target return and CRRA as reference portfolio	MDivSTCRRA	3.4.2
	Skew-T target return and EQ as reference portfolio	MDivSTEQ	3.4.2

Table 4.3: Presentation and abbreviations of the different considered portfolios.

Chapter 5

Results and Further Analysis

Based on the methodology presented in chapter 4, we now report and analyse our results to provide some answers about the minimum divergence portfolio. First, we present the performance analysis. Secondly, the sensitivity of the Kullback-Leibler function under different targets will be presented. We then assess the importance of the shape of the target. After, we provide some insight regarding the order of the Rényi divergence and a small comparison of the two divergence measures. Finally, we approach the Kullback-Leibler divergence with a Gaussian quadrature.

5.1 Portfolio Comparison

We first present the results an investor would have obtained if he had set his wealth allocation according to Markowitz portfolio selection technique and also under portfolio strategies that takes higher moments into account. We then compare them with minimum divergence portfolios with different starting points

as presented in 4.3. Finally, the minimum divergence portfolios are computed under the Gaussian distribution and under the Skew-T distribution.

5.1.1 In-sample analysis of the considered portfolios

The in-sample analysis allows us to see if the wealth allocation under the minimum divergence portfolio would have outperformed the other portfolio selection techniques. The first two moments for the investor's targets are chosen as in Lassance and Vrins [2019] where the reference portfolio is first the Markowitz portfolios which are then switched to a modified value-at-risk and then to constant relative risk aversion utility function and finally to an equally-weighted portfolio. Under the Skew-T distribution, we chose the skewness and the kurtosis being equal to 0.2 and 4 respectively. From an investor's perspective, we would like to reach a higher skewness and we are ready to accept moderately more kurtosis than under a Gaussian distribution. All the following analysis is based on table 5.1 where the obtained results reflect the whole data-set meaning the period 1978-2018. We now discuss the results :

- Comparison of **traditional portfolios and portfolios under higher moments**

    1. **Mean-variance comparison** : At first blush, the Markowitz portfolios seem to be a good strategy for asset allocation. They offer a competitive mean for a reasonable standard deviation where this can be observed on both data-sets. The portfolios under higher moments offer relatively the same mean variance trade-off where they generally offer a higher mean and higher volatility than the MV but a lower mean and lower volatility than the MSR. The EQ portfolio might offer a higher mean but generally with way more volatility or inversely

which makes it less desirable for the investors.

2. **Skewness and Kurtosis** : When we look at the higher moments in more depth, we can see that the skewness is highly negative and that the excess kurtosis is largely positive for Markowitz portfolios. This leads the investor to a position that is not quite attractive according to investor preferences. The portfolios under higher moments generally lead to an enhancement of higher orders and especially in *5Ind*. The CRRA portfolio enhances higher orders but we note that it behaves relatively similarly to the MV portfolio even though we took a large value of lambda. In fact, this portfolio is also closely related to the 3Mom portfolio where we consider equal weights for the first three moments. Surprisingly, the EQ portfolio can sometimes lead to better higher orders than more advanced techniques.

3. **Modified value-at-risk, Sharpe Ratio and Adjusted Sharpe Ratio** : On an in-sample basis, the M-VaRPort portfolio leads to the smallest modified value-at-risk and the MSR to the maximum Sharpe ratio. This seems natural since the aim of these portfolios are respectively to minimize and maximize the modified value-at-risk and the Sharpe ratio. We note that the adjusted Sharpe ratio is generally in favour of portfolios under higher orders but in a low proportion especially in *6BTM*.

- Comparison of **traditional portfolios and portfolios under higher moments with the minimum divergence portfolios under a Gaussian target** :

1. **Mean-variance comparison** : If we now turn to the minimum divergence portfolios under a Gaussian target, we note that the mean-

variance trade-off is relatively similar but tends to be higher than the traditional portfolios and portfolios under higher orders.

2. **Skewness and Kurtosis** : When we look at higher orders, results are striking. In fact, the minimum divergence portfolio always leads to much better higher orders. The enhancement of higher orders is particularly high when we compare with Markowitz portfolios and the MDivGMV and MDivGMSR. The minimum divergence portfolios based on Markowitz portfolios outperform portfolios under higher orders. When we compare with portfolios under higher orders, the minimum divergence portfolio based on portfolios under higher orders also enhance the skewness and excess kurtosis. Moreover, in *5Ind*, the MDivGMVAR reaches skewness and excess kurtosis that are even better than MDivGMV and MDivGMSR.

3. **Modified value-at-risk, Sharpe Ratio and Adjusted Sharpe Ratio** : The modified value-at-risk is generally in favour of the minimum divergence portfolios and sometimes in favour of the traditional portfolios and portfolios under higher orders. Since the minimum divergence portfolios can lead to a higher volatility, the gain in higher orders is counterbalanced by the increase in volatility. The adjusted Sharpe ratio is almost always in favour of the minimum divergence portfolio and especially in *5Ind*. The Sharpe ratios are relatively similar.

- Comparison of **traditional portfolios and portfolios under higher moments with the minimum divergence portfolio under a Skew-T target** :

1. **Mean-variance comparison** : Under the Skew-T distribution, the

same observation as before can be done. In fact, the mean and variance of the minimum divergence portfolios are relatively similar are those under a Gaussian. In *6BTM* the mean and variance under the Skew-T distribution is generally higher than the reference portfolio but smaller than under a Gaussian distribution. On the contrary, in *5Ind*, the mean and variance are generally higher than the reference portfolio and than under a Gaussian target. Globally, the mean-variance trade-off is similar to the reference portfolios.

2. **Skewness and Kurtosis** : In *5Ind*, the minimum divergence portfolios under a Skew-T distribution lead to a significant enhancement of higher orders in comparison with the reference portfolio and small enhancement in comparison with the minimum divergence portfolio under a Gaussian target. The excess kurotsis is also improved under the Skew-T distribution. The improvement is really significant in comparison with the Markowitz portfolios and slightly less significant with portfolios under higher orders. We outline that targeting a distribution with a higher skewness and larger excess kurtosis does not necessarily lead to a better skewness and larger excess kurotis. In fact, in *6BTM*, the skewness under a skew-T distribution is lower than under a Gaussian distribution. However, in *5Ind*, the skewness is better than under a Gaussian with also smaller excess kurtosis. This can be explained by the fact that the minimum divergence portfolio tries to be as close as possible to the target and is thus impacted by the complete shape of the distribution.

3. **Modified value-at-risk, Sharpe Ratio and Adjusted Sharpe Ratio** : The modified value-at-risk is generally in favour of the minimum divergence portfolios and sometimes in favour of the reference

portfolios. However, the adjusted Sharpe ratio is almost always in favour of the minimum divergence portfolios. Again, the Sharpe ratio can sometimes be in favour of the traditional portfolios and sometimes in favour of the minimum divergence portfolios.

- Comparison of **starting points for the minimum divergence portfolios**

  1. **Markowitz portfolios** : Choosing Markowitz portfolios as reference to define the investor's target seems to be a good starting point. The enhancement of higher orders is significant for almost the same mean-variance trade-off. Turning to the adjusted Sharpe ratio, it is always in favour of the minimum divergence portfolio while the Sharpe ratio can sometimes be in favour of the reference portfolio and sometimes in favour of the minimum divergence portfolio.

  2. **Portfolios under higher orders** : Choosing portfolios under higher orders to define investor's target seem to be a nice starting point as well. In fact, choosing the M-VaRPort as starting point can lead to even smaller higher orders, smaller modified value-at-risk and higher adjusted Sharpe ratio. Choosing the CRRA portfolio as reference is extremely similar as choosing the MV as reference.

To further confirm these first results, we need to perform the same analysis but on an out-of-sample basis. This framework being a new method, both analyses are recommendable. However, if an investor has pursued his wealth allocation according to the minimum divergence on the considered period, he would have been in a much more desirable position.

Portfolio	Mean	SD	Skewness	Excess Kurtosis	M-VaR	Sharpe Ratio	Adjusted SR
6BTM							
MV	12.04%	14.58%	-0.6225	2.431	12.49%	0.8259	0.6981
MSR	14.84%	16.27%	-1.058	3.335	15.02%	0.9125	0.6601
CRRA	12.12%	14.63%	-0.6207	2.316	12.42%	0.8284	0.7025
3Mom	12.11%	14.63%	-0.6218	2.367	12.47%	0.8275	0.7007
M-VaRPort	12.14%	14.86%	-0.5812	2.132	12.39%	0.8171	0.704
EQ	13.39%	16.44%	-0.942	2.921	14.87%	0.8141	0.6444
MDivGMV	12.33%	15.38%	-0.5397	1.923	12.57%	0.8016	0.7025
MDivGMSR	12.38%	15.79%	-0.5085	1.778	12.73%	0.7838	0.6961
MDivGCRRA	12.34%	15.47%	-0.5328	1.888	12.6%	0.7979	0.7014
MDivGMVAR	12.34%	15.47%	-0.5324	1.886	12.6%	0.7977	0.7013
MDivGEQ	12.38%	15.79%	-0.5085	1.778	12.73%	0.7838	0.6961
MDivSTMV	12.25%	14.91%	-0.5864	2.19	12.49%	0.8214	0.7049
MDivSTMSR	12.38%	15.79%	-0.5085	1.778	12.73%	0.7838	0.6961
MDivSTCRRA	12.21%	14.8%	-0.6026	2.297	12.52%	0.8254	0.7032
MDivSTMVAR	12.23%	14.83%	-0.5966	2.256	12.51%	0.8242	0.704
MDivSTEQ	12.38%	15.79%	-0.5085	1.778	12.73%	0.7838	0.6961
5Ind							
MV	13.53%	15.09%	-0.7307	19.77	6.871%	0.8963	0.2055
MSR	13.85%	15.31%	-0.6918	17.81	6.524%	0.9046	0.2609
CRRA	13.66%	15.49%	-0.6574	16.27	6.239%	0.8819	0.3316
3Mom	13.69%	15.21%	-0.6859	17.88	6.494%	0.8998	0.2644
M-VaRPort	14.06%	17.03%	-0.549	12.22	5.817%	0.8257	0.4766
EQ	13.39%	16.3%	-0.6394	16.55	6.635%	0.8213	0.3673
MDivGMV	14.23%	16.23%	-0.6148	14.97	6.217%	0.8765	0.3778
MDivGMSR	14.26%	16.43%	-0.5975	14.44	6.162%	0.8679	0.3995
MDivGCRRA	14.28%	16.53 %	-0.5895	14.2	6.138%	0.8639	0.4091
MDivGMVAR	14.4%	17.64%	-0.5103	11.92	5.931%	0.8163	0.4895
MDivGEQ	14.35%	17.16%	-0.5427	12.82	6.009%	0.836	0.4608
MDivSTMV	14.29%	16.62%	-0.5822	13.98	6.116%	0.86	0.4177
MDivSTMSR	14.32%	16.79%	-0.5683	13.58	6.075%	0.8529	0.4329
MDivSTCRRA	14.34%	16.93%	-0.5571	13.26	6.044%	0.8471	0.4445
MDivSTMVAR	14.45%	17.74%	-0.4972	11.64	5.89%	0.8144	0.4974
MDivSTEQ	12.73%	21.01%	-0.09645	11.14	6.567%	0.6056	0.4966

Table 5.1: In-sample performance of all the considered portfolios where the mean and standard deviation are presented on an annual basis.

5.1.2 Out-of-sample analysis of the considered portfolios

We now turn to the out-of-sample analysis that better reflects the reality of asset allocation. In this section, we follow the same strategy as in the previous section but the only difference is that we estimate the portfolios on a sample period that is smaller than the entire data-set and we then keep the weights constant for the rest of the period of the data-set. The rolling window used in this thesis is such that $\tau = 456$ observations in *6BTM* which corresponds to 38 years of monthly data and keep the weights constant for the next window of two years. In *5Ind*, $\tau = 9324$ observations which corresponds to 37 years of daily data and keeps the weights constant for the next three years where τ equals the period of estimation. In other words, the analysis of the performance is based on a window of respectively two and three years. All the following analysis is based on table 5.2.

- Comparison of **traditional portfolios and portfolios under higher moments** :

    1. **Mean-variance comparison** : As in the in-sample analysis, the mean-variance trade-off between the traditional portfolios and portfolios under higher orders is relatively similar. On the considered data-sets, the portfolios under higher orders generally lead to higher mean but also higher volatility. The EQ portfolio generally leads to higher mean but also a higher volatility than the Markowitz portfolios.

    2. **Skewness and Kurtosis** : Surprisingly, the skewness is better in the traditional portfolios than in portfolios under higher orders except for the M-VaRPort in *5Ind*. The excess kurtosis is relatively enhanced in portfolios under higher orders. However, we note that that the higher orders are extremely close. The higher orders of the EQ can be lower

or higher than other portfolios.

3. **Modified value-at-risk, Sharpe Ratio and Adjusted Sharpe Ratio** : In almost every case, the modified value-at-risk, Sharpe ratio and adjusted Sharpe ratios are in favour of portfolios under higher orders. Since the higher orders are extremely close, this can be explained by the fact that the mean in portfolio under higher orders is generally higher for almost the same volatility.

- Comparison of **traditional portfolios and portfolios under higher moments with the minimum divergence portfolio under a Gaussian target** :

1. **Mean-variance comparison** : It is difficult to detect any particular trend in terms of mean and variance. In *6BTM* the mean is always higher than the traditional portfolios and than portfolios under higher orders while in *5Ind* the inverse can be observed. However, the reader will notice that the mean and variance of the minimum divergence portfolios are generally closely related to the reference portfolios.

2. **Skewness and Kurtosis** : In almost every case, the minimum divergence portfolios lead to better higher orders. The skewness is always improved in comparison with the reference portfolio. Naturally, the proportion of improvement can differ depending on the reference portfolio but the minimum divergence portfolios almost always reach a better skewness. The same observation can be done with the excess kurtosis where it is significantly improved. In terms of higher orders, the investors are almost always in a more desirable position than under traditional portfolios or portfolios under higher orders.

3. **Modified value-at-risk, Sharpe Ratio and Adjusted Sharpe**

Ratio : The modified value-at-risk of the minimum divergence portfolios is generally relatively close to the one of the reference portfolios. As aforementioned, the enhancement of higher orders can be counterbalanced by an increase in volatility or decrease of the mean. However, in *6BTM* the Sharpe ratio and adjusted Sharpe ratio are always in favour of the minimum divergence meaning that the minimum divergence portfolio is more attractive for the investor. On the contrary, in *5Ind*, the minimum divergence portfolios lead to poorer results. Naturally, we did not expect the minimum divergence portfolio to always lead to better results than traditional portfolio selection techniques and this confirms that sometimes traditional portfolio selection techniques can be better.

- Comparison of **traditional portfolios, portfolios under higher moments,minimum divergence portfolio under a Gaussian target with the minimum divergence portfolio under a Skew-T target** :

  1. **Mean-variance comparison** : Under the Skew-T distribution, the mean and variance of the minimum divergence portfolios are relatively similar than under a Gaussian distribution and lower than the reference portfolios. Globally, the mean-variance trade-off is similar to the reference portfolios.

  2. **Skewness and Kurtosis** : In *5Ind*, the minimum divergence portfolios under a Skew-T distribution lead to a significant enhancement of skewness in comparison with the reference portfolio and a small enhancement in comparison with the minimum divergence portfolio under a Gaussian target. This seems natural since the Gaussian distribution targets a skewness of zero and the Skew-T a skewness of

0.2. However, we can see that targeting a distribution that has a larger skewness does not automatically lead to a larger skewness as in the in-sample analysis. In *6BTM*, the skewness under the Skew-T distribution is lower than under a Gaussian distribution. The excess kurotsis is also improved under the Skew-T distribution. The improvement is really significant in comparison with the Markowitz portfolios and slightly less significant with portfolios under higher orders.

3. **Modified value-at-risk, Sharpe Ratio and Adjusted Sharpe Ratio** : The modified value-at-risk can sometimes be in favour of the traditional portfolios and sometimes in favour of the minimum divergence portfolios. However, in *6BTM* the Sharpe ratio and adjusted Sharpe ratio is almost always in favour of the minimum divergence portfolios. Globally, in *6BTM*, the minimum divergence portfolios under a Skew-T distribution also outperform the traditional portfolios and portfolios under higher orders. On the contrary, in *5Ind*, the minimum divergence portfolios only enhance the higher orders.

- Comparison of **starting points for the minimum divergence portfolios**

  1. **Markowitz portfolios** : Choosing Markowitz portfolios as reference to define the investor's target seems to be a good starting point. The enhancement of higher orders is significant for almost the same mean-variance trade-off.

  2. **Portfolios under higher orders** : Choosing portfolios under higher orders to define the investor's targets seems to be a nice starting point as well. In fact, choosing the M-VaRPort as starting point can lead to even smaller higher orders.

Portfolio	Mean	SD	Skewness	Excess Kurtosis	M-VaR	Sharpe Ratio	Adjusted SR
6BTM							
MV	8.417%	10.74%	-1.152	2.099	9.111%	0.7836	0.6237
MSR	9.755%	14.37%	-0.5034	2.021	11.93%	0.679	0.614
CRRA	9.548%	11.03%	-1.173	1.933	9.144%	0.866	0.6671
3Mom	10.25%	11.04%	-1.159	1.83	9.03%	0.9282	0.7007
M-VaRPort	10.21%	11.1%	-1.157	1.849	9.098%	0.9204	0.697
EQ	9.49%	13.91%	-0.9543	1.816	11.7%	0.6822	0.5841
MDivGMV	11.32%	11.4%	-1.122	1.683	9.166%	0.993	0.74
MDivGMSR	12.33%	11.79%	-1.08	1.575	9.352%	1.046	0.7738
MDivGCRRA	11.61%	11.51%	-1.111	1.65	9.211%	1.009	0.7499
MDivGMVAR	11.71%	11.54%	-1.107	1.638	9.228%	1.015	0.7534
MDivGEQ	12.33%	11.79%	-1.08	1.575	9.352%	1.046	0.7738
MDivSTMV	9.471%	10.92%	-1.16	1.932	9.057%	0.8676	0.6695
MDivSTMSR	12.33%	11.79%	-1.08	1.575	9.352%	1.046	0.7738
MDivSTCRRA	9.481%	10.92%	-1.16	1.93	9.057%	0.8683	0.6699
MDivSTMVAR	9.481%	10.92%	-1.16	1.93	9.057%	0.8683	0.6699
MDivSTEQ	12.33%	11.79%	-1.08	1.575	9.352%	1.046	0.7738
5Ind							
MV	10.7%	11.55%	-0.2972	2.749	2.254%	0.9256	0.7923
MSR	10.54%	11.99%	-0.2794	2.761	2.336%	0.8786	0.7646
CRRA	11.17%	11.48%	-0.306	2.68	2.228%	0.9735	0.8222
3Mom	11.31%	11.39%	-0.3079	2.609	2.199%	0.9929	0.8359
M-VaRPort	9.697%	14.76%	-0.2027	2.151	2.716%	0.6571	0.6171
EQ	12.39%	12.57%	-0.3265	2.96	2.5%	0.9855	0.8146
MDivGMV	9.499%	13.6%	-0.2204	2.277	2.534%	0.6986	0.6484
MDivGMSR	9.315%	13.94%	-0.2136	2.197	2.579%	0.6683	0.6251
MDivGCRRA	9.377%	13.82%	-0.2158	2.224	2.564%	0.6785	0.633
MDivGMVAR	8.888%	15.52%	-0.1935	1.95	2.811%	0.5725	0.5467
MDivGEQ	8.926%	14.96%	-0.1998	2.02	2.726%	0.5967	0.567
MDivSTMV	9.187%	14.34%	-0.2066	2.118	2.636%	0.6405	0.6031
MDivSTMSR	9.025%	14.76%	-0.2014	2.048	2.696%	0.6114	0.5793
MDivSTCRRA	8.42%	15.85%	-0.1965	1.926	2.868%	0.5313	0.51
MDivSTMVAR	8.964%	14.92%	-0.1996	2.025	2.72%	0.6006	0.5704
MDivSTEQ	8.52%	15.78%	-0.1958	1.93	2.855%	0.54	0.5178

Table 5.2: Out-of-sample performance of all the considered portfolios where the mean and standard deviation are presented on an annual basis.

After having analysed the performance of the minimum divergence it becomes quite clear that the method developed by Lassance and Vrins [2019] can outperform traditional portfolios and portfolios under higher orders. The method leads generally to a similar mean-variance trade-off - without any special trend - as in the chosen reference portfolio combined with significant enhancement in higher orders.

However, we also saw that the method does not always outperform traditional portfolios and portfolios under higher orders. In *5Ind*, the minimum divergence portfolios lead to poorer results. The choice of the distribution leads to different results but we outline the fact that using a Gaussian distribution as target is already a good starting point given the general market conditions.

The choice of the reference portfolio having been assessed, we can say that choosing the Markowitz portfolios as reference leads to significant improvement of higher orders in comparison to the reference portfolios but also that choosing the M-VaRPort as reference can lead to nice higher orders and desirable portfolios for an investor. Finally, choosing the CRRA portfolio as reference is almost similar as choosing the MV. In the end, whatever the choice of the reference portfolio, using the minimum divergence portfolio as strategy leads to an enhancement of higher orders but in different proportion depending on the investor's reference portfolio.

5.2 Kullback-Leibler function sensitivity

In this section, we assess how the Kullback-Leibler function evolves according to different moments. In order to answer this question, we first assume our portfolio is distributed according to a Skew-T distribution that has the same first four moments as the MV portfolio. We then assess the evolution of the Kullback-Leibler function when we target a Skew-T with another Skew-T that has one particular moment that is changing. The first case illustrates the point where the portfolio density is exactly the same as the target. In the following tables, the mean and SD are presented on an annual basis but we have used a monthly basis in our optimization. We also plot the evolution of the Kullback-Leibler function under every considered moments.

5.2.1 Skewness and Kurtosis sensitivity

The table 5.3 represents how our portfolio is distributed and the different targets we would like to reach. The table is as follows :

Skew-T target	Mean	SD	**Skewness**	Excess Kurtosis	$KL(x)$ value
Skewness Varying					
Portfolio	12.04%	14.58%	**-0.6225**	2.431	-
Target1	12.04%	14.58%	**-0.6**	2.431	0
Target2	12.04%	14.58%	**-0.3**	2.431	0.3345
Target3	12.04%	14.58%	**0**	2.431	0.9858
Target4	12.04%	14.58%	**0.3**	2.431	1.981
Target5	12.04%	14.58%	**0.6**	2.431	3.792
Target6	12.04%	14.58%	**0.9**	2.431	8.052
Excess Kurtosis Varying					
Portfolio	12.04%	14.58%	-0.6225	**4.5**	-
Target1	12.04%	14.58%	-0.6225	**4.5**	0
Target2	12.04%	14.58%	-0.6225	**3.5**	0.01516
Target3	12.04%	14.58%	-0.6225	**2.5**	0.1373
Target4	12.04%	14.58%	-0.6225	**1.5**	0.6595
Target5	12.04%	14.58%	-0.6225	**1**	1.496
Target6	12.04%	14.58%	-0.6225	**0.5**	5.796

Table 5.3: The first two moments are chosen to be in line with the in-sample MV portfolio in *6BTM*

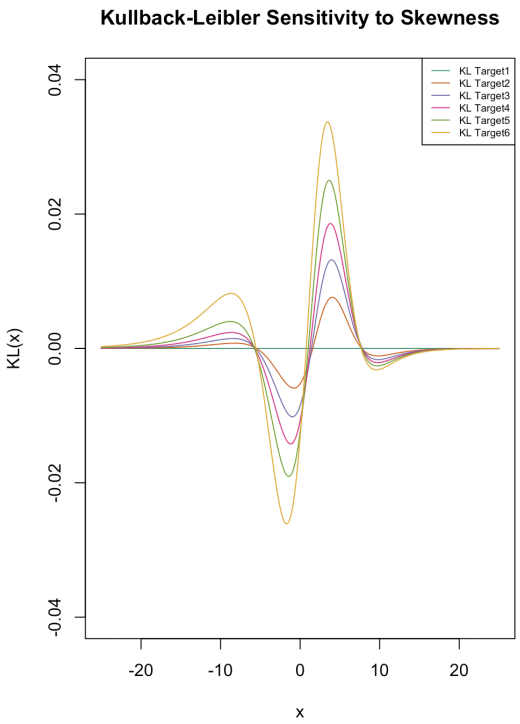

Figure 5.1: Evolution of the Kullback-Leibler function when a constant portfolio density approaches targets with the skewness that is varying

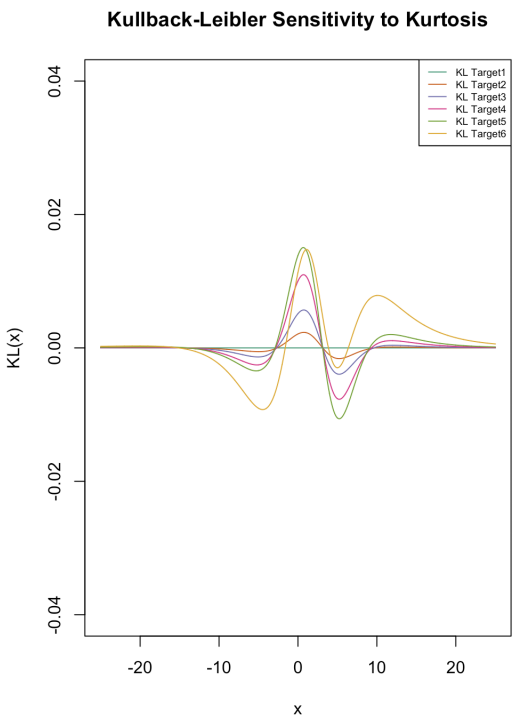

Figure 5.2: Evolution of the Kullback-Leibler function when a constant portfolio density approaches targets with the kurtosis that is varying

As our first target has exactly the same properties as our portfolio distribution, the sum of the Kullback-Leibler function equals 0. When we increase the skewness we note that the sum of the KL increases. When the skewness becomes positive we can see that the sum of the KL function is rapidly increasing. Moreover, when the skewness of the target is far from the distribution skewness, the sum of the KL becomes largely positive. This can also be observed in figure 5.1 where the area under the function is increasing at each step. Regarding the kurtosis, we can observe the same phenomenon but in a smaller proportion both numerically and graphically in figure 5.2. This allows us to say that the KL function is more sensitive to the skewness but also fairly impacted by the kurtosis.

5.2.2 Mean and Variance sensitivity

In the mean-variance sensitivity, we change our portfolio distribution where in the variance sensitivity our portfolio is distributed according to a Skew-T distribution that has the same first four moments as in the in-sample MSR of *6BTM* and we then decrease the variances of the targets. If we take the portfolio moments as in the MV we could obviously not decrease the variance. Concerning the mean analysis, we define our portfolio distribution as a Skew-T with the same moments as in the in-sample MV portfolio in *6BTM* and increase the mean of the targets. This allows us to respect investor's preferences. We also present unreasonable mean and variance targets in order to underline the significance of the target choice.

Skew-T target	Mean	SD	**Skewness**	Excess Kurtosis	$KL(x)$ value
Variance Varying					
Portfolio	12.05%	**16.27%**	-1.058	3.35	-
Target1	12.05%	**14.86%**	-1.058	3.35	0.7245
Target2	12.05%	**14.79%**	-1.058	3.35	0.7808
Target3	12.05%	**14.58 %**	-1.058	3.35	1.076
Target4	12.05%	**12.12%**	-1.058	3.35	8.576
Mean Varying					
Portfolio	**12.05%**	14.58%	-0.6225	2.431	-
Target1	**12.05%**	4.209%	-0.6225	2.431	0
Target2	**12.14%**	4.209%	-0.6225	2.431	0.000193
Target3	**14.84%**	4.209%	-0.6225	2.431	0.1795
Target4	**16.37%**	4.209%	-0.6225	2.431	0.4273
Target5	**20.4%**	4.209%	-0.6225	2.431	1.586

Table 5.4: The first two moments are chosen to be in line first with the in-sample MSR portfolio in *6BTM* and secondly to be in line with the in-sample MV in *6BTM*

As it can be seen on fig 5.3, the KL function is more sensitive to the variance than to the mean. Naturally both moments impact the KL function but for a small change in the standard deviation from the portfolio distribution, the KL function tremendously increases. Decreasing the standard deviation of 4 percentage points leads to an increase of the KL function from 0 to 8.57 while an increase of 8

percentage points in the mean only leads to an increase of the KL function of 1.58. From this observation, we can conclude that the choice of the standard deviation of the target highly impacts the KL function.

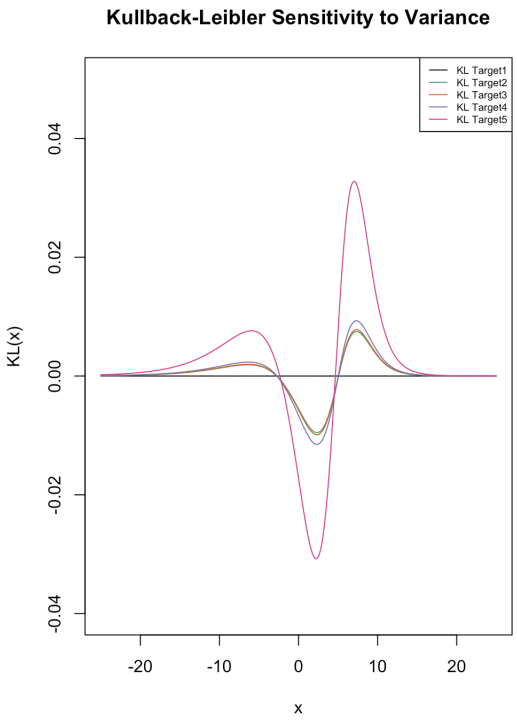

Figure 5.3: Evolution of the Kullback-Leibler function when a constant portfolio density approaches targets with the standard deviation that is varying

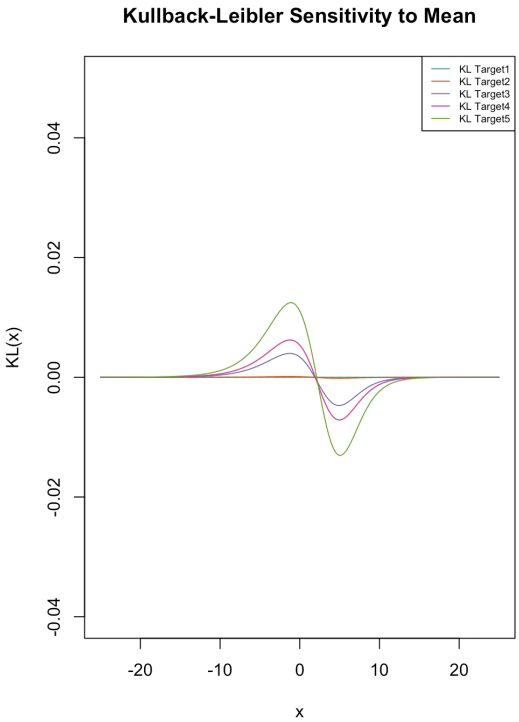

Figure 5.4: Evolution of the Kullback-Leibler function when a constant portfolio density approaches targets with the mean that is varying

5.3 Does the shape of the target really matter?

In this section, we assess if targeting a specific value of modified value-at-risk but with different moments always leads to the same weight specification or if the shape of the distribution influences the result. Naturally, we expect that the shape of the target highly influences the result. The reference portfolio in this case would be the M-VaRPort. We target the following Skew-T distributions that all have the same modified value-at-risk but with different moments where the first target has exactly the same first four moments as the in-sample M-VaRPortf in *6BTM*. We obtain the following :

Skew-T target	Mean	SD	**Skewness**	**Excess Kurtosis**	M-VaR
Skewness and Kurtosis Varying					
M-VaRPort	12.14%	14.86%	-0.5812	2.132	12.4%
Target1	12.14%	14.86%	-0.5812	2.132	12.4%
Target2	12.14%	14.86%	-0.3	2.63	12.4%
Target3	12.14%	14.86%	0.1	3.752	12.4%
Target4	12.05%	12.12%	0.4	4.938	12.4%
Skew-T target	**Mean**	**SD**	Skewness	Excess Kurtosis	M-VaR
Mean and Variance Varying					
M-VaRPort	12.14%	14.86%	-0.5812	2.132	12.4%
Target1	12.12%	14.85%	-0.5812	2.132	12.4%
Target2	14.84%	14.86%	-0.5812	2.132	12.4%
Target3	18%	15.41%	-0.5812	2.132	12.4%
Target4	9.6%	14.63%	-0.5812	2.132	12.4%

Table 5.5: The reference portfolio is the M-VaRPort and targets are chosen to all have the same modified value-at-risk but with different moments

Portfolio	Mean	SD	Skewness	Excess Kurtosis	M-VaR	Sharpe Ratio	Adjusted SR
Skewness and Kurtosis Varying							
M-VaRPort	12.14%	14.86%	-0.5812	2.132	12.39%	0.8171	0.704
Portfolio1	13.09%	15.17%	-0.8561	2.722	13.43%	0.8626	0.6836
Portfolio2	12.41%	14.91%	-0.6778	2.272	12.66%	0.8327	0.6997
Portfolio3	12.18%	14.74%	-0.6156	2.389	12.58%	0.8263	0.7001
Portfolio4	12.18%	14.75%	-0.6127	2.368	12.56%	0.8264	0.701
Mean and Variance Varying							
M-VaRPort	12.14%	14.86%	-0.5812	2.132	12.39%	0.8171	0.704
Portfolio1	13.09%	15.13%	-0.8527	2.715	13.39%	0.8651	0.6855
Portfolio2	13.56%	15.42%	-0.9166	2.902	13.84%	0.8793	0.679
Portfolio3	14%	15.81%	-0.9725	3.122	14.42%	0.8853	0.668
Portfolio4	12.68%	15.01%	-0.7554	2.46	12.99%	0.8444	0.693

Table 5.6: Portfolio characteristics when all the targets have the same modified value-at-risk but with different moments

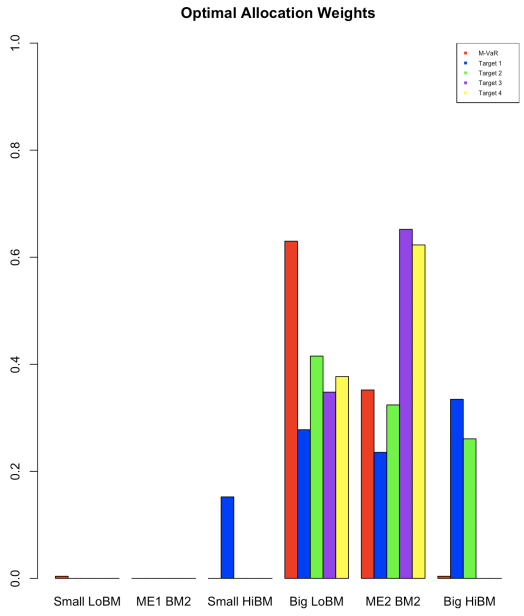

Figure 5.5: Weight of the different targets and the M-VaRPort portfolio when the skewness and the kurtosis of the targets are varying

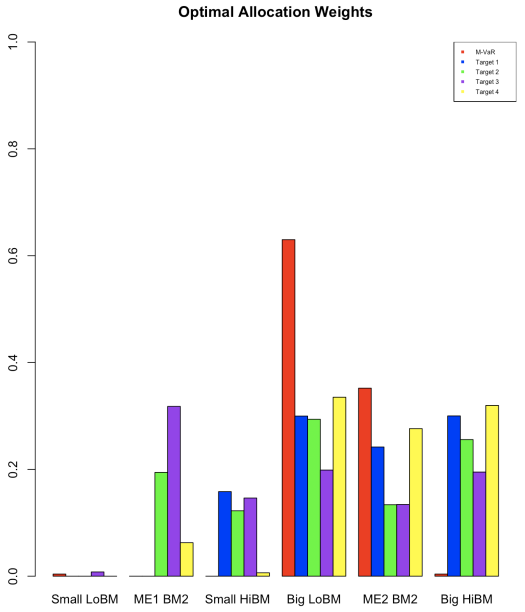

Figure 5.6: Weight of the different targets and the M-VaRPort portfolio when the mean and the variance of the targets are varying

- As expected, targeting a specific modified value-at-risk or targeting a return distribution that has the same modified value-at-risk is clearly not the same approach. We obtain significant different results according to the target. This first observation confirms that the way an investor defines his target is important.

- Moreover, we outline that targeting a distribution that has exactly the same first four moments as the M-VaRPort portfolio would give the investor an unsatisfactory performance. He would end up with a larger volatility, a larger negative skewness and a higher excess kurtosis (Portfolio 1). Targets with other specific moments lead to other portfolio moments but still a poorer performance than the M-VaRPort. This also confirms that the way an investor specifies his target is crucial. Targeting a Skew-T distribution

with a skewness being equal to -0.58 leads, in this case, to a skewness of -0.85.

- Finally, the reader can easily see that fixing the target as a distribution that has a specific value of modified value-at-risk is not an efficient way of doing. Almost every portfolio has a larger volatility and excess kurtosis, a lower skewness and a lower adjusted Sharpe ratio.

5.4 Rényi Divergence and Kullback-Leibler Divergence Comparison

In this section, we compare the two measures presented in this thesis respectively the Kullback-Leibler and the Rényi divergence. Taking an $\alpha = 0.99$ for the Rényi divergence would lead to the same results as under the Kullback-Leibler divergence. We performed this analysis to confirm our results but do not present it in this thesis.

Hence, we first present the optimization where we target the minimum variance portfolio under a Gaussian distribution performed under the Kullback-Leibler and then the Rényi divergence with $\alpha = 0.8$.

Portfolios	Mean	SD	Skewness	Excess Kurtosis	M-VaR	Sharpe Ratio	Adjusted SR
Target1	1.004%	4.209%	0	0			
Portfolio under Kullback-Leibler	12.33%	15.38%	-0.5397	**1.923**	12.57%	0.8016	0.7025
Portfolio under Rényi ($\alpha = 0.8$)	12.34%	15.47%	-0.5327	**1.888**	12.6%	0.7978	0.7014

Table 5.7: Presentation of the obtained portfolios under the Kullback-Leibler divergence and the Rényi divergence

We observe that the portfolio density under the Kullback-Leibler divergence and the Rényi divergence are extremely close. This is relatively straightforward since the objective of the divergence function is to be as close as possible to the density target. Considerable differences would be uncanny. Nevertheless, we note that the portfolio under the KL divergence is slightly more leptokurtic. This is mainly due to the effect of the α parameter which we will assess now.

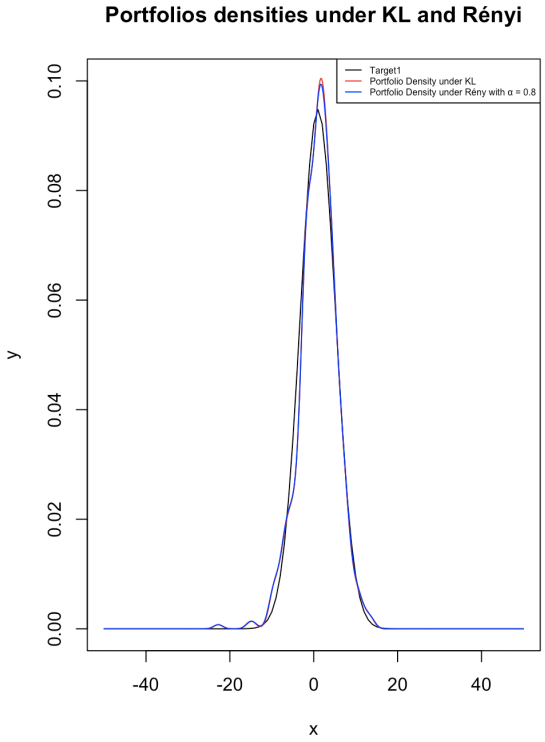

Figure 5.7: Densities of the obtained portfolio under Kullback-Leibler and Rényi divergence

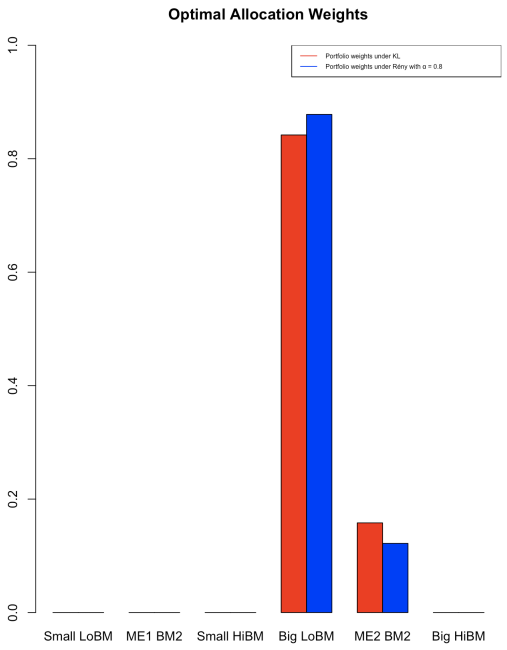

Figure 5.8: Present the different allocation weights under different divergence measures

5.4.1 Influence of α parameter

The question is : How does the α parameter in the Rényi divergence influence the result of the minimum divergence portfolio? In order to answer this question, we target three times the same distribution but with different values of α as follows:

Portfolios	Mean	SD	Skewness	Excess Kurtosis	M-VaR	Sharpe Ratio	Adjusted SR
Target1	1.004%	4.209%	0	0			
Portfolio under Rényi ($\alpha = 0.8$)	12.34%	15.47%	-0.5327	**1.888**	12.6%	0.7978	0.7014
Portfolio under Rényi ($\alpha = 1.2$)	12.35%	14.85%	-0.6472	**2.251**	12.57%	0.8313	0.7028
Portfolio under Rényi ($\alpha = 1.6$)	12.18%	14.89%	-0.6614	**2.509**	12.89%	0.8179	0.687

Table 5.8: Presentation of the obtained portfolio moments for different values of alpha parameter

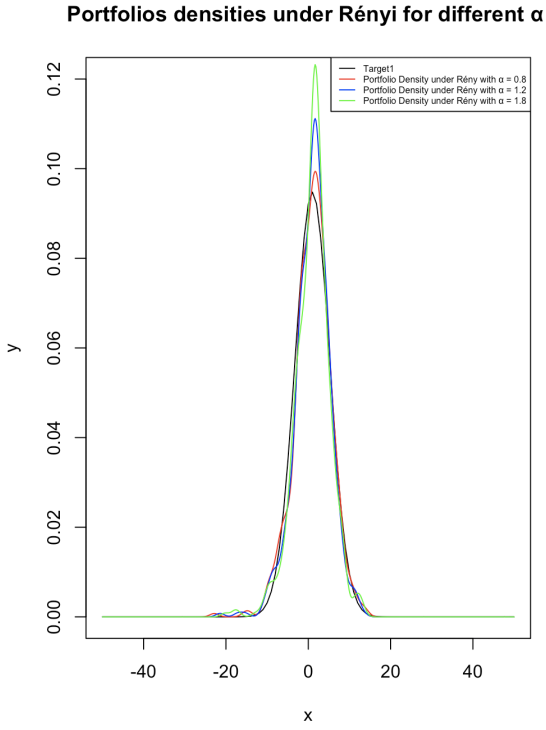

Figure 5.9: Portfolios densities under Rényi divergence for different values of alpha

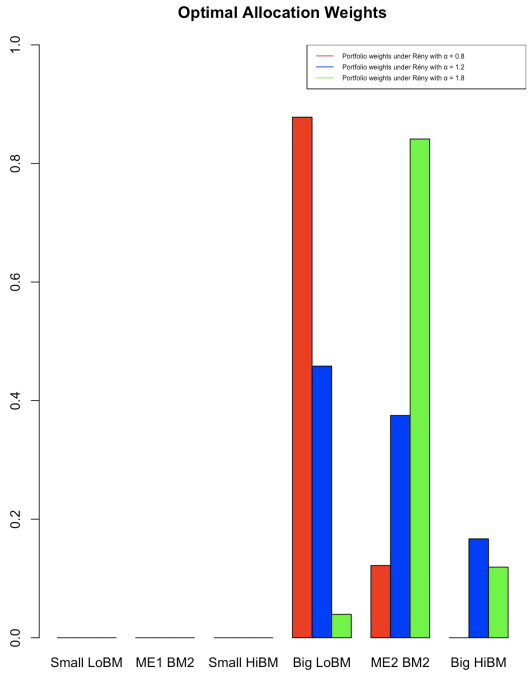

Figure 5.10: Optimal portfolio weight allocation for different values of alpha under the Rényi divergence

We observe that taking larger values of α leads to more leptokurtic distributions meaning that in the tail region the distribution of a portfolio with a large α will have fatter tails than a distribution with lower values of α. We can observe on figure 5.9 that the green distribution has fatter tails in the tail region. This is also what we can observe in the portfolio moments where the portfolio with a large value of α leads to more excess kurtosis. Concerning the other moments, smaller values of α also lead to better performance in terms of mean and skewness but not in terms of volatility. This is what can be observed on *6BTM* data-set. We also performed the same optimization on *5Ind* where results can differ in terms of the first three moments but in line with the observation about the excess kurtosis.

Recall from 3.5 that when $\alpha > 1$ the divergence measure is an increasing function of $\mathbb{E}\left((\frac{f_P}{f_T})^{\alpha-1}\right)$ and when $\alpha < 1$ the function becomes $\left(\mathbb{E}\left[(\frac{f_P}{f_T})^{\alpha-1}\right]\right)^{-1}$.

- Increasing the value of α leads to a decreasing contribution of the regions where $f_P(x) < f_T(x)$
- Decreasing the value of α leads to an increasing contribution of the regions where $f_P(x) < f_T(x)$

Practically, lower values of α would lead to a portfolio density that has lower probability mass in the tails than larger values of α. This means that if the minimum divergence is computed under the Rényi divergence, two options are possible :

1. Select a lower value of α that will certainly lead to a portfolio with less excess kurtosis.
2. Select a large value of α that will lead to a portfolio density with a larger excess kurtosis

5.5 Numerical integration of the Kullback-Leibler Divergence

This last part aims at presenting the numerical integration of the minimum divergence portfolio. Theoretically, the integration of the divergence should be possible on interval $\mathbb{R}$. However, we face a problem in the integration where the divergence can be non-finite when the bounds of the integrals are too wide. In fact, recall from equation 4.5 that we use a ratio of logarithms and that when we look too far in the tails, the values can become increasingly small leading to a divergence measure that is not convergent.

However, the integration can be avoided using the Gaussian quadrature. For a given random variable Z distributed as Gaussian distribution, the expectation of a function $g(Z)$ can be approached as follows :

$$\mathbb{E}[g(Z)] = \int g(z)\phi(z)\mathrm{d}z \approx \sum_{i=1}^{n} \omega_i g(x_i), \tag{5.1}$$

where the ω_i's and x_i's are the weights and nodes associated to the quadrature, which are independent of g and n represents the number of nodes considered. Now in order to approach the Kullback-Leibler divergence with a Gaussian quadrature we need to rewrite the Kullback-Leibler function. Recall from 3.4 that the Kullback-Leibler divergence is defined as :

$$\langle f_P | f_T \rangle = \int f_P(z) \ln \frac{f_P(z)}{f_T(z)} \mathrm{d}z, \tag{5.2}$$

If we define a new function that corresponds to the the Kullback-Leibler function

divided by $\phi(z)$ and call this function KL_ϕ we obtain the following :

$$KL_\phi(z) = \int \frac{f_P(z) \ln \frac{f_P(z)}{f_T(z)}}{\phi(z)} \mathrm{d}z, \tag{5.3}$$

Using KL_ϕ, we can now approximate the Kullback-Leibler divergence by a Gaussian quadrature as follows :

$$\langle f_P | f_T \rangle = \mathbb{E}[KL_\phi(Z)] = \int \frac{f_P(z) \ln \frac{f_P(z)}{f_T(z)}}{\phi(z)} \phi(z) \approx \sum_{i=1}^{n} \omega_i KL_\phi(x_i), \tag{5.4}$$

where the reader can notice that by multiplying and dividing by ϕ,the Kullback-Leibler function is retrieved.

For example, for the following target $f_T \sim N(1.004\%, 4.209\%)$, the results of the numerical integration and the Gaussian quadrature for different n points are as follows : As we can see, we obtain the same result than under the integration

Type of integration	n points	w_1	w_1	w_1	w_1	w_1	w_1	KL Value
Under integration	-	0	0	0	0.8419	0.1581	0	0.03824
Gaussian Quadrature 1	64	0.1033	0.009552	0	0.4814	0.4057	0	0.007659
Gaussian Quadrature 2	128	0	0.02546	0.02811	0.9464	0	0	0.01905
Gaussian Quadrature 3	256	0	0	0	0.8419	0.1581	0	0.03824

Table 5.9: Portfolio weights under numerical integration and Gaussian quadrature for different number of points

when we consider a sufficient number of points. This validates the obtained results even though the bounds are limited. In figure 5.11, we plot the Kullback-Leibler function between a portfolio distributed as an equally-weighted portfolio and a Gaussian distribution with the first two moments as in the in-sample MV portfolio in *6BTM*. We can see that when the bounds are wide enough all the variations of the Kullback-Leibler function are taken into account and that the function is not impacted anymore :

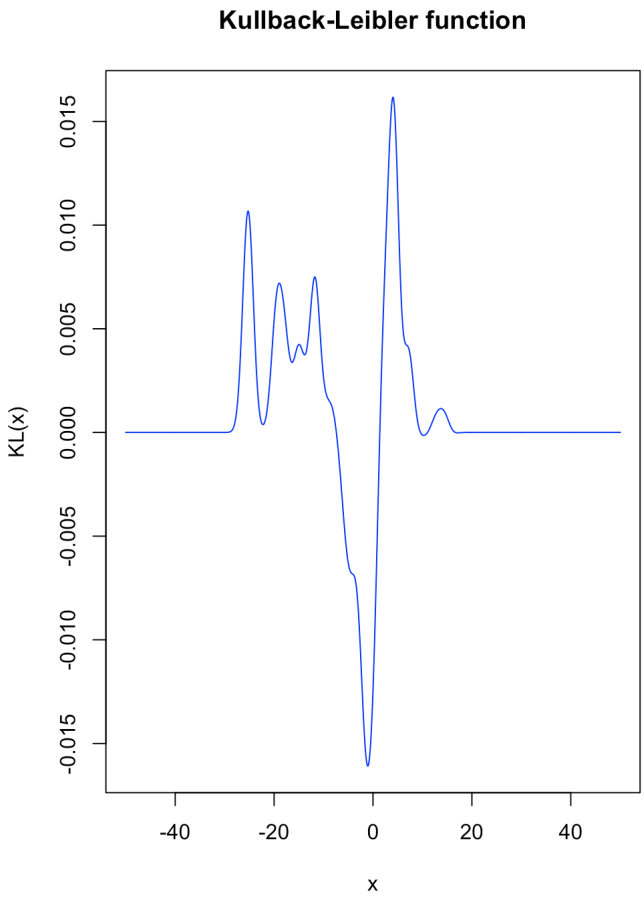

Figure 5.11: Kullback-Leibler function

Chapter 6

Conclusion

This thesis wished to provide a study of the minimum divergence portfolio selection framework. We start our conclusion by briefly summarizing our results, then address the encountered research limitations and conclude by proposing suggestions for future research. The purpose of this thesis was to evaluate the performance of the minimum divergence portfolio against traditional portfolio selection techniques and assess the impact of the target choice as well as the choice of the divergence measure. We compare the performance of the minimum divergence portfolio developed by Lassance and Vrins [2019] following the same methodology of the authors against classical portfolio selection techniques. We developed our methodology based on targets that are distributed under a Gaussian and a Skew-T distribution. Furthermore, we evaluate the choice of new reference portfolios based on the minimization of the modified value-at-risk portfolio, the portfolio selection framework based on constant relative risk aversion and the equally weighted portfolio. We also compared if targeting a return distribution that has a specific value of modified value-at-risk is similar to minimizing the modified value-at-risk to judge if the shape of the target really matters in the

minimum divergence portfolio. In terms of divergence measure, we empirically evaluate the sensitivity of the Kullback-Leibler function to different moments. Then, we compare the choice of the divergence measure that can be used under this methodology and assess the impact of the α parameter in the Rényi divergence. Finally, we approached the numerical integration of the Kullback-Leibler function by a Gaussian quadrature in order to see if we obtain the same results and to confirm that limiting the bounds of the integral is not an issue.

The results found in the empirical analysis tells us that wealth allocation based on the minimum divergence portfolio generally leads to the same mean-variance trade-off but also to a significant improvement in higher orders especially in comparison with Markowitz portfolios. This improvement in higher orders generally make the modified value-at-risk and the adjusted Sharpe ratio in favour of the minimum divergence portfolios. Hence, the minimum divergence portfolios can outperform the traditional portfolios and portfolios under higher orders. However, this observation is not always true since on one data-set the Markowitz portfolios outperformed the minimum divergence portfolio. Furthermore, the minimum divergence portfolios also enhance higher orders in comparison with portfolios that already account for higher orders. We also notice that the minimum divergence portfolio based on a modified value-at-risk portfolio can lead to an improvement in higher orders where the attained skewness is less negative and the excess kurtosis smaller than the minimum divergence portfolio with the Markowitz portfolios as reference. Taking the constant relative risk aversion utility function as reference generally leads to similar results as the minimum divergence portfolio based on the minimum variance portfolio. The shape of the distribution is also crucial in the minimum divergence framework. Targeting distributions that all have the same modified value-at-risk but with different parameters leads to highly different results. This highlight the fact that the shape of the target highly influences

the results. Turning to the divergence measure, we observed that the Kullback-Leibler function is fairly impacted by all the target moments but is especially sensitive to the variance and to the skewness. Finally, comparing the use of the Rényi divergence and the Kullback-Leibler divergence, we find out that when the target is properly defined both divergence measures lead to similar results. The main difference lays in the choice of the α parameter where we find out that larger values of α lead to more leptokurtic distributions. Hence, for an investor that based his wealth allocation according to the minimum divergence portfolio but under the Rényi divergence, we advice him to select a low value of α if he does not want to face a portfolio distribution that has higher probability mass in the tails than his target. To conclude, we approached the Kullback-Leibler integration under a Gaussian quadrature and found that we obtained the same results when we considered a sufficient number of points under the quadrature. From this observation, we can conclude that restricting the bounds of the integral of the Kullback-Leibler function is not an issue.

To conclude, a careful reader should remember the following : The minimum divergence portfolio framework can outperform the Markowitz portfolios and portfolios under higher orders whatever the choice of the reference portfolio. Targeting a return distribution is not the same as targeting specific moments of a distribution. The Kullback-Leibler function is highly influenced by the skewness and variance and fairly impacted by the mean and kurtosis. Under the Rényi divergence, an investor should select a low value of alpha if he does not want to face a portfolio distribution that has higher probability mass in the tails than his target.

6.1 Research limitations

Every research has its limitations and this section aims at pointing out the main ones we encountered. A careful reader should consider the following aspects :

1. We performed only one re-balancing period which might not be realistic for an investor and makes our results really influenced by the rolling window. In fact, the choice of another window could lead to better or poorer results. Furthermore, we did not take into consideration any transaction costs while in reality they can not be ignored.

2. The second one comes from the fact that all our analyses are based on only two data-sets and therefore the obtained results are specific to these data-sets and could vary depending on other data.

3. We also limited our analysis of other portfolio selection techniques while obviously in the literature there are plenty of other techniques. We tried to limit our analysis to the most relevant portfolios but in practice other comparisons could be done.

This being said, the reader should keep in mind these limitations while considering the obtained results about the minimum divergence portfolio.

6.2 Suggestions for further research

We conclude this thesis by suggesting several future research proposals that are worth investigating :

1. We only assess the minimum divergence portfolio under the Gaussian distribution and the Skew-T distribution. Other distributions can be tested

and could lead to an enhancement of the results.

2. The approach of a reference portfolio proposed in Lassance and Vrins [2019] is the main approach we followed where we investigate new references portfolios. An investigation on another approach to fix the target could be worth investigating.

3. Finally, the authors analysed the performance of the minimum divergence during the financial crisis. A comparison on how the different portfolios behave during special macroeconomic conditions is worth investigating too.

References

Alexander, G. and Baptista, A. (2002). Economic implications of using a mean-var model for portfolio selection : A comparison with mean-variance analysis. *Journal of Economics Dynamics Control*, 26:1159–1193. 19

Arellano-Valle, R. and Azzalini, A. (2013). The Centred Parameterization and Related Quantities of the Skew-t Distribution. *Journal of Multivariate Analysis*, 113:73–90. 29

Baumol, W. (1963). An expected gain-confidence limit criterion for portfolio selection. *Management Science*, 10:174–182. 19

Boudt, K., Peterson, B., and Croux, C. (2008). Estimation and decomposition of downside risk for portfolios with non-normal returns. *The Journal of Risk*, 11(2):79–103. 20

Cavenaile, L. and Lejeune, T. (2012). A note on the use of modified value-at-risk. *Journal of Alternative Investments*, 4(14):79–83. 21

Chalabi, Y. and Wurtz, D. (2012-2014). Portfolio optimization based on divergence measures. *ETH Econohysics Working and White Papers Series*. 25

Chopra, V. and Ziemba, W. (1993). The effect of errors in means, variances,

and covariances on optimal portfolio choice. *Journal of Portfolio Management*, 119(1):6–11. 14

Cont, R. (2001). Empirical properties of asset returns: stylized facts and statistical issues. *Quantitative Finance*, 1(3):223–23. 16

DeMiguel, V., Lorenzo, G., and Raman, U. (2009). Optimal versus naive diversification: How inefficient is the 1-n portfolio strategy? *Review of Financial Studies*, 22(5):1915–1953. 32

DeMiguel, V. and Nogales, F. (2009). Portfolio selection with robust estimation. *Operations Research*, 55(5):798–812. 2

Favre, L. and Galeano, J. (2002). Mean-modified value-at-risk optimization with hedge funds. *Journal of Alternative Investments*, 5(2):21–25. 3, 21

Jagannathan, R. and Ma, T. (2003). Risk reduction in large portfolios: Why imposing the wrong constraints helps. *Journal of Finance*, 58(4):1651–1684. 34

Jondeau, E. and Rockfinger, M. (2011). Optimal portfolio allocation under higher moments. *Operations Research Letters*, 39(3):163–171. 17

Kallberg, J.G.and Ziemba, W. (1984). Mis-specifications in portfolio selection problems. In G., B. and K., S., editors, *Risk and Capital. Lecture Notes in Economics and Mathematical Systems*, pages 74–87. Springer, Berlin, Heidelberg. 14

Kimball, M. (1993). Standard risk aversion. *Econometrica*, 61(3):589–611.

Kullback, S. and Leibler, R. (1951). On information and sufficiency. *The Annals of Mathematical Statistics*, 2(1):79–86. 3, 24, 25

Lassance, N. and Vrins, F. (2019). Portfolio selection with higher-order moments: A target-distribution approach. Paris, France. 9th General AMaMeF Conference. 3, 24, 31, 39, 50, 65, 69

Ledoit, O. and Wolf, M. (2004). A well-conditioned estimator for large-dimensional covariance matrices. *Journal of Multivariate Analysis*, 88(2):365–411. 2

Lim, A., Shanthikumar, J., and Vahn, G. (2011). Conditional value-at-risk in portfolio optimization: Coherent but fragile. *Operations Research Letters*, 39(3):163–171. 3

Mahalanobis, P. (1936). On divergences and informations in statistics and iformation theory. *Proceedings of the National Institute of Sciences of India*, 2:49–55. 24

Mandelbrot, B. (1963). The variation of certain speculative prices. *Journal of Business*, 36(4):394–419. 16

Markowitz, H. (1952). Portfolio selection. *Journal of Finance*, 7:77–91. 2, 4, 6, 7

Martellini, L. and Ziemann, V. (2010). Improved estimates of higher-order comoments and implica- tions for portfolio selection. *Review of Financial Studies*, 23(4):1467–1502. 3, 17, 18, 19

Merton, R. (1972). An analytical derivation of the efficient portfolio frontier. *Journal of Finance and Quantitative Analysis*, 7(4):1851–1872. 11

Pardo, L. (2006). *Statistical Inference Based On Divergence Meausures*. Taylor-Francis Group, 6000 Broken Sound Parkway NW, Suite 300 Boca Raton, FL 33487-2742. 25

Pézier, J. (2004). Risk and risk aversion. In Alexander, C. and Sheedy, E. N., editors, *The professional risk managers' handbook.* PRMIA Publications, Wilmington DE. 21

Rényi, A. (1951). On measures of entropy and information. *Fourth Berkeley Symposium on Mathematical Statistics and Probability*, (1):547–561. 24, 26

Scott, R. and Horvath, P. (1980). On the direction of Preference for moments of higher orders than the variance. *Journal of Finance*, 35(4):915–919. 2, 17

Sharpe, W. (1994). The Sharpe Ratio. *The Journal of Portfolio Management*, 21(1):49–58. 12

Silverman, B. (1986). *Density Estimation for Statistics and Data Analysis*. Chapman Hall, London. 27

Theodossiou, P. (1998). Financial data and the skewed generalized t distribution. *Management Science*, 44(12):1650–1661. 28

Ullah, A. (1996). Entropy, divrgence and distance measures with econometric applications. *Journal of Statistical Planning and Inference*, 49(1):137–162. 25

Xiong, J. and Idzorek, T. (2011). The impact of skewness and fat tails on the asset allocation decision. *Financial Analysts Journal*, 67(2):23–35. 17

Zvi, B., Alex, K., and J.M, A. (2014). *Investment.* McGraw-Hill Education, Penn Plaza, New York. 7