Asset Management Project

Quantitative Asset and Risk Management

Portfolio optimization with Expected Shortfall

Master of Science (MSc) in Finance

Habib Joe
Borel Maxime
Bandelier Axel
Joye Yoann

Prof. Eric Jondeau

TA. Alexandre Pauli

University of Lausanne - 21.04.2020

UNIL | Université de Lausanne
HEC Lausanne

Contents

List of Figures

List of Tables

Abstract

Our study is based on the paper *Portfolio optimization with conditional Value-at-Risk objective and constraints* from Pavlo Krokhmal, Jonas Palmquist and Stanislav Uryasev. Hence, we followed their suggestion of optimization of Condition Value-at-Risk (CVaR), to obtain an optimal portfolio in terms of trade-off between their returns and their protection against large downturns (or potential shortfall). In the paper of Krokmal and al., they confront their model, which is a maximization of returns under a constraint of CVaR, with a Mean-Variance optimal portfolio. From there, we extend the study by creating two other models that include the CVaR plus the one of the paper. Then, we confront them with several optimal portfolios, such as the Global Minimum Variance Portfolio, the Market Portfolio and the Mean-Variance portfolio. Finally, our case study is the Fama and French's 17 industries portfolio.

1 Introduction

Portfolio optimization theory have known large improvement since Markowitz (1952). Tones of paper propose new approach and improvements of this basic methodology which is still largely use. Indeed, The classic optimization theory proposed by Markowitz use restrictive assumptions. In the real world implementation, these assumptions are obviously often violated like the iid feature and normality of returns. Moreover, it may be the case that the investor care not only about mean and variance but has also preference for asymmetric measures. The classical optimization do not take into consideration large events in the tail distribution of returns. In these days, people could be very interested to construct a portfolio that take into account these extreme risks.

This report aims to show an implementation of a portfolio optimization under constraint of a downside risk measure. We use the Conditional Value-at-Risk and construct a portfolio around this measure. In this study, we make no assumptions about the distribution of returns and therefore do not assume normality.

As regards our study, we will first go through the data chosen to perform the models. Then, we will move on to the different models' parameters and constraints used. Next, as in the paper, we will compare the CVaR optimal portfolio with the mean-variance portfolio, as well as with other models. These being the Global Minimum Variance (GMVP) and the Market portfolios. After that, we will perform a sensitivity analysis across all our three CVaR models.

The Value-at-Risk (VaR) is the maximum loss that can be reached over a certain period of time. This threshold is represented by the percentile α. For example, A VaR of 95% is the upper estimate of losses which will exceed the threshold with a probability of 5%. In the finance industry, the VaR is the most widely used benchmark to measure financial downturn risk.

The Conditional Value-at-Risk (CVaR) is also a measure of extreme negative returns. Indeed, it is the weighted average of the losses strictly higher than the VaR. Contrary to the VaR, the CVaR is widely used in the insurance industry (the CVar in not a standard in financial world). In fact, the CVaR represents an expected shortfall value. Hence, both CVaR and VaR are measures of potential downside returns.

Then, the approach of the paper by Krokhmal and al., was to implement a CVaR constraint to a portfolio optimization process. Then, they chose to compare the results with a standard Markowitz's Mean-Variance (MV) optimal portfolio. One of the challenges to compare them was that Mean-variance portfolio includes both negative and positive values, whereas the CVaR optimal portfolio reshapes its distribution, by summarizing results with different time frames, in order to have all its possible values in the negative area of the possible returns' values.

To conclude with the paper, they found that the Mean-Variance optimal portfolio has a higher efficient frontier than the CVaR optimal portfolio, in that sense the former performs better. Still, it is worth to mention that with a lower significance level, the gap between the MV optimal portfolio and the CVaR one, widens.

1.1 Definition: Value at Risk

The Value-at-Risk (VaR) is the loss associated with a particular confidence level of a cumulative loss function. In other words, it is the maximum loss that we can have according to a certain period of time and a certain probability. The VaR is in fact a threshold in the negative tail of the return after which (in absolute value) we consider the negative returns to unlikely. it gives a simple interpretation of the downside risk in probabilistic terms. Mathematically, it is defined as follows with α a probability and $\Psi(x, \zeta)$, the cumulative distribution function of the loss function:

$$\zeta_\alpha(x) = min\{\zeta \in R : \Psi(x, \zeta) \geqslant \alpha\} \tag{1}$$

1.2 Definition: Expected Shortfall

The Expected Shortfall also known as the Conditional Value-at-Risk is a downside risk measure as well. In fact, the CVaR and the VaR complement one another, in the sense that the VaR will consider the entire negative tail and the CVaR will measure the average loss in the area bounded by the VaR. As the VaR, this risk measure is defined for a certain confidence level such as 90%, 95% or 99%. Moreover, it is also defined for a certain time horizon. As shown below in Figure 1, the CVaR represents the average loss which is larger than or equal to the VaR (in absolute term). If we define for example $CVaR_\alpha$ with $\alpha = 95\%$, it means that we are 95% sure that the loss will not exceed this amount. A larger loss could intervene on average every 20 months in this case.

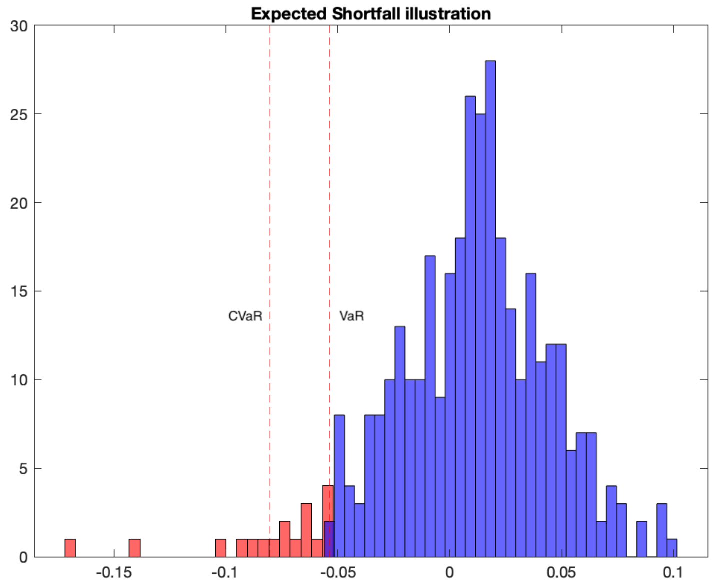

Figure 1: VaR and CVaR illustration

We have to keep in mind that in Figure 1, we graphically represent the distribution of returns. Hence, the VaR and the CVaR have a negative sign.

Now, we develop the mathematics behind the risk measure. In their paper, Krokhmal, Palmquist Uryasev (2002) define the Conditional Value-at-Risk $\phi_\alpha(x)$ as in the equation (2) with $f(x, y)$ being the loss function associated with the vector of weights x and $p(y)$, the probability density function of the returns. Meaning that, we will take the average loss function conditionally to being greater than or equal to the threshold ζ which is the VaR. The whole expression is for a given confidence level α:

$$\phi_\alpha(x) = (1-\alpha)^{-1} \int_{f(x,y) \geqslant \zeta} f(x, y) p(y) \, \mathrm{d}y \tag{2}$$

Optimizing under the CVaR condition suppose that we have to compute first the VaR which is difficult to handle. In the paper, they proposed a function of the CVaR that is independent of the VaR. Hence, the optimization problem becomes less difficult to solve. They define the function $F_\alpha(x, \zeta)$ as follow:

$$F_\alpha(x, \zeta) = \zeta + (1-\alpha)^{-1} \int_{y \in R^n} [f(x, y) - \zeta]^+ \, p(y) \, \mathrm{d}y \tag{3}$$

Krokhmal, Palmquist and Uryasev show that it is equivalent to solve the function $\phi_\alpha(x, \zeta)$ or to minimize $F_\alpha(x, \zeta)$ as in the equation (4). This equation is crucial as it will be used for the implementation of the portfolio optimization:

$$\phi_\alpha(x,\zeta) = \min_{\zeta \in R} F_\alpha(x,\zeta) \tag{4}$$

The main implication of this equation is that in addition to finding the optimal weights for the portfolio assets, we have to find an optimal ζ. They proved in a previous paper that minimizing over the CVaR of the portfolio, we would get an almost optimal VaR.
The function $F_\alpha(x,\zeta)$can be discretized as follow :

$$\tilde{F}_\alpha(x,\zeta) = \zeta + (1-\alpha)^{-1} \sum_{j=1}^{J} \pi_j \left[f(x,y_j) - \zeta\right]^+ \tag{5}$$

We can approximate the integral by sampling the probability distribution where π_j is the probability of occurrence of the scenario y_j. We will explain how we treat the π_j in the section 3.1. Then, the final step is to use a dummy variable z_j, j=1,...,J to replace the max function. The function $F_\alpha(x,\zeta)$ can be rewritten as a function under some constraints :

$$\zeta + (1-\alpha)^{-1} \sum_{j=1}^{J} \pi_j z_j \leqslant w \tag{6}$$

$$z_j \geqslant f(x,y_j) - \zeta, z_j \geqslant 0, j = 1, ..., J \tag{7}$$

The dummy variable z_j will take the value $f(x,y_j) - \zeta$ when it is greater than 0 and 0 otherwise. In addition, we consider several constraints that transform the non-linear problem in a linear one. Moreover, we have as optimization variables, the weights of the portfolio but also the parameter ζ and all the dummy variables z_j for each scenario we consider. Here, we understand that the CVaR is computed as follows : it is the VaR defined by ζ and we add the average losses measured by $(1-\alpha)^{-1} \sum_{j=1}^{J} \pi_j z_j$, which must be smaller than the objective of CVaR.

2 Data

To begin with, for the data we based our project on, we used the Fama and French's 17 industries portfolio. For the sake of comparing our results, we used several benchmarks such as:

- U.S. 10 Years Treasury Constant Maturity Rate monthly basis: it will be our risk-free rate benchmark.

- The excess return on the market, value-weight return of all CRSP firms incorporated in the US and listed on the NYSE, AMEX or NASDAQ and we add the risk free rate above the excess return: this will be an approximation of the market.

We retrieve data from several sources. Regarding the Risk free rate, we get it from the Federal Reserve of St.Louis Fed. The market return is from the Fama-French three-factor model. The reasons why we chose those data were:

- For the risk-free rate, a 10 years on a monthly basis rate enabled to match the time lag between two re-balancing events of the portfolio. Moreover, it also match the classical long term horizon of an investment. Indeed, by choosing 10 years rate, we made sure we had the right risk free rate for every sampled period.

- Regarding the estimation of the market, we used the portfolio of industries instead of a portfolio of assets in a sake of stability for the returns. Practically, as the market returns (R_m) were not explicitly given, but instead the excess returns of the market ($R_m - R_f$), we just added back the risk-free rate (R_f), that is $R_m - R_f + R_f = R_m$.

Also, both the 1-month US Treasury Bond and the Fama-French are expressed in returns term. Furthermore, the Fama-French 17 industries are composed of the following sectors: Mines, Oil, Food, Clothes, Durables, Chemicals, Consumables, Construction, Steel, Fabricated Products, Machinery and Business Equipment, Automobiles, Transportation, Utilities, Retail Stores, Banking and Financials, Others. We can therefore assume, from this extensive list, that it covers a large landscape of American industries.

One other reason why we used Fama-French database is that they are free of missing data. If it were to happen that some data were missing, we would have replaced this one by zero in order not to affect the aggregated returns.

Concerning the time frame of our study, we used the most recent years to have more data available, and also, we assumed that not many funds existing nowadays started before 1960. Moreover, we decided to start our analysis in 1990 in order to have 30 years of monthly data, which we assume to be necessary, in terms of number of observations, to obtain a decent level of accuracy in our estimates.

3 Methodology

In this section, we are going to describe the assumptions and the constraints that we have made for the optimization process and how we have implemented these ones in terms of linear programming. Then, we are going to describe the various models that we are going to estimate with the differences between them and their various implications.

3.1 Assumptions and optimization setting

To start with, we are going to discuss the assumptions relative to the returns, the CVaR, the scenarios and the optimization settings that are important for the optimization. We start with the optimization settings as it explains sometimes certain of the assumptions that we will discuss thereafter.

We start the optimization process in 1990. As our data has begun in 1960, we are using a rolling-window with 30 years length with monthly data. We are proceeding this way in order to have enough data to estimate, the CVaR, the co-variance matrix and the expected returns. Moreover, it is not common to have some investment funds that start in 1970 and still operating. Hence, we define the optimization period from January 1990 to January 2020.

Then, regarding the returns, as we use the sample average over the rolling-window to predict the expected returns, we assume that the returns are mean-reverting. Hence, we supposed that past returns are good predictors for the next period. This practice have several implications, that we will discuss later. By the previous assumption, we also assume that returns are stationary and identically distributed without assuming any distribution. We also assume that market is efficient in the weak form as there is no predictability in return.

Regarding the CVaR, using the methodology of Krokhmal, Palmquist and Uryasev, it imposes no assumption on the distribution of the returns. Hence, we are able to take into account the non-normality of returns. To compute it, we will use the scenarios. We assume that the investor uses a confidence level of 95%, meaning we compute the $CVaR_\alpha$ with $1-\alpha = 5\%$. We could have used the $CVaR_\alpha$ with $1-\alpha = 1\%$ but it would have required much more data to have a good approximation of the tail of the distribution. We use for the principal optimization a coefficient $\lambda = 1$. As we consider the optimization regarding the CVaR risk measure, we assume that the investor only cares about the downside risk of the distribution. Moreover, we do not impose any restriction on the volatility as we expect that maximizing the expected returns, such that the CVaR is smaller than a threshold, will give a low variance. Finally, as we have monthly returns, the validity of the CVaR is of one month.

To compute the CVaR, we need to have some scenarios. The approach of Krokhmal, Palmquist Uryasev allows to use historical data without assuming any distribution as mention above. In our case, we use historical data. Having a large rolling-window is also important for the estimation of the CVaR. With the rolling window of 30 years, it gives us a good approximation of the empirical distribution.
As mention previously, in equation (6), we have to assign probability π_j to each scenario. In this case,

we assume that the scenarios are equally-weighted over the past. Therefore, $\pi_j = 1/J$. Basically, this is the average of the scenarios.
We can rewrite Equation (6) as follow:

$$\zeta + \frac{1}{J(1-\alpha)} \sum_{j=1}^{J} z_j \leqslant w \tag{8}$$

Finally, we need to precise how we have computed the CVaR of the portfolio over time. We compute it in a simple way. As we consider the confidence level of $\alpha = 95\%$ for the optimization, we also use this assumption for the historical CVaR. We compute it as follows:

$$VaR_p = quantile(R_{p,t}, 1-\alpha) \tag{9}$$

$$CVaR_p = \frac{1}{T} \sum_{t=1}^{T} r_{p,t} \quad \text{with } r_{p,t} \geqslant VaR_p \tag{10}$$

3.2 Models

3.2.1 Model 1

Our first model aims at maximizing the returns with a dynamic change of weights for each asset during each period of the rolling-window (every month). We set the usual constraints on the weights : it has to sum to 1 (no leverage allowed) and each weight must be between 0 and 0.2. On top of these, we add a third constraint : the CVaR cannot overtake a given threshold, that we have fixed at 9.5% (see sensitivity analysis for effect of other thresholds), which means that no more than 9.5% of the assets within the 5% tail after the VaR can reach higher losses than the CVaR. This model is similar to the the paper's from Krokhmal, Palmquist and Ursayev.

We can notice that, in model 1 and model 3, we do not need to compute the co-variance matrix as plug-in estimator. The estimation error is concentrated in the expected returns.
The optimization problem can be formulated as follow :

$$\begin{aligned}
\min_{w_t,\alpha,z_1,\ldots,z_j} \quad & -w_t'\mu_t \\
\text{s.t.} \quad & \zeta + \frac{1}{J(1-\alpha)} \sum_{j=1}^{J} z_j \leqslant 0.095 \\
& z_j \geqslant 0 \quad j = 1,\ldots,J \\
& w_t' r_j + \zeta + z_j \geqslant 0 \quad j = 1,\ldots,J \\
& w_t' 1 = 1 \\
& 0 \leqslant w_t \leqslant 0.2
\end{aligned} \tag{11}$$

3.2.2 Model 2

The second model has as objective function the Sharpe ratio with a constraint of expected shortfall. The Sharpe ratio is a measure of performance of the returns relative to the variance: $SR = \frac{R_p - r_f}{\sigma_p}$

We aim at maximizing this Sharpe ratio. This approach adds implicitly a restriction which is minimizing the expected standard deviation of the portfolio. As we are maximizing a performance ratio, we expect this portfolio to get a better risk adjusted performance. Compared to the two other models, we need to compute the co-variance matrix. It leads to certain problems of estimation errors such that, there is a large number of assets, it is badly estimated.

$$
\begin{aligned}
\min_{w_t,\alpha,z_1,\ldots,z_j} \quad & -\frac{w_t'\mu_t - r_f}{\sqrt{w_t'\Sigma_t w_t}} \\
\text{s.t.} \quad & \zeta + \frac{1}{J(1-\alpha)}\sum_{j=1}^{J} z_j \leqslant 0.095 \\
& z_j \geqslant 0 \quad j = 1,\ldots,J \\
& w_t' r_j + \zeta + z_j \geqslant 0 \quad j = 1,\ldots,J \\
& w_t' 1 = 1 \\
& 0 \leqslant w_t \leqslant 0.2
\end{aligned} \tag{12}
$$

3.2.3 Model 3

Finally, our last model is a Mean-Variance optimization portfolio with the CVaR replacing the variance. The optimization process aims at finding the optimal level of CVaR that maximizes the utility of an investor under the constraint of a given level of risk aversion specific to this very investor. By default, we set a risk aversion criterion of $\lambda = 1$, which we could say is a relatively low risk aversion and therefore, it is an aggressive portfolio in favour of the maximization of returns. Further in our report, in the sensitivity analysis section, we test a more conservative setting in order to have a better model than the paper, both in terms of simple returns and downside protection.

$$
\begin{aligned}
\min_{w_t,\alpha,z_1,\ldots,z_j} \quad & -(w_t'\mu_t - \lambda(\zeta + \frac{1}{J(1-\alpha)}\sum_{j=1}^{J} z_j)) \\
\text{s.t.} \quad & z_j \geqslant 0 \quad j = 1,\ldots,J \\
& w_t' r_j + \zeta + z_j \geqslant 0 \quad j = 1,\ldots,J \\
& w_t' 1 = 1 \\
& 0 \leqslant w_t \leqslant 0.2
\end{aligned} \tag{13}
$$

3.3 Implementation

In this section, we discuss how we have implemented these portfolio optimizations in terms of linear programming. Indeed, the way that Krokhmal, Palmquist and Ursayev propose to compute the CVaR allows us to do it in a linear way as mentioned before. The implementation is similar for Model 1 and Model 2 as they have the same constraints and they only differ by their objective function. The third Model is different in the sense that we do not define any constraint on the expected shortfall. Basically, we are maximizing the returns while minimizing the CVaR. Hence, the first constraint disappears.

We have to keep in mind that we are not only optimizing with respect to the weights w_t but also for the VaR ζ and all the z_j. Therefore, the vector over which we are optimizing contains all the parameters. We estimate in total $N+1+J$ parameters for each period excluding expected return and co-variance matrix.

Now, we show how we are implementing the set of restrictions in a linear way. For Model 1 and 2, the restrictions are the following:

$$\begin{aligned} \zeta + \frac{1}{J(1-\alpha)}\sum_{j=1}^{J} z_j &\leqslant w \\ z_j \geqslant 0 \quad j &= 1,\ldots,J \\ w_t' r_j + \zeta + z_j \geqslant 0 \quad j &= 1,\ldots,J \\ w_t' 1 &= 1 \\ 0 \leqslant w_t &\leqslant 0.2 \end{aligned} \tag{14}$$

First, we focus on the inequality constraints. Matlab requires an implementation in such a way that $A * w \leqslant b$. Hence, we have A and b :

$$A = \begin{pmatrix} 0 & 0 & \cdots & 0 & 1 & \frac{1}{J(1-\alpha)} & \frac{1}{J(1-\alpha)} & \cdots & \frac{1}{J(1-\alpha)} \\ -r_{1,1} & -r_{1,2} & \cdots & -r_{1,N} & -1 & -1 & 0 & \cdots & 0 \\ -r_{2,1} & -r_{2,2} & \cdots & -r_{2,N} & -1 & 0 & -1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ -r_{J,1} & -r_{J,2} & \cdots & -r_{J,N} & -1 & 0 & 0 & \cdots & -1 \end{pmatrix} \quad b = \begin{pmatrix} w \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

According to this representation, we are implementing the restrictions 1 and 2. The first line corresponds to the first restriction. We have replaced the 0.095 by w because we will use this formulation also in the sensitivity analysis. We see this matrix in three parts. The first N columns concern the weights. Then, the next column is for the VaR and the last block is the auxiliary variables. We replace the positive sign in the restriction 3 by a negative one as we have to express the condition in term of smaller or equal.

Regarding the constraint 2 and the last one, we define them as upper and lower bound of the vector w_t:

$$UB = \begin{pmatrix} ub_1 & ub_2 & \cdots & ub_N & \inf & \inf & \cdots & \inf \end{pmatrix} \tag{15}$$

$$LB = \begin{pmatrix} lb_1 & lb_2 & \cdots & lb_N & 0 & 0 & \cdots & 0 \end{pmatrix} \tag{16}$$

We decide to cap the maximal weight per asset at 20%. It is quite high but we think that it is acceptable since we use industry portfolios and not individual assets for the optimization. In addition, in their paper, Krokhmal, Palmquist and Ursayev use the same upper bond. It appears to be a good starting point.

Furthermore, we compute the turnover T_i for each portfolio as follow:

$$T_i = \sum_{t=2}^{T} \left| \alpha_{i,t} - \alpha_{i,t-1} \frac{1 + r_{i,t}}{\sum_{i=1}^{N} \alpha_{i,t}(1 + r_{i,t})} \right| \tag{17}$$

The calculation of turnovers is very important to have an estimate of the transaction costs generated by our models. Indeed, the more our models involve frequent re-balancing of our portfolio, the more transaction costs increase and consequently our models become less competitive compared to a passive management portfolio (market portfolio).

Finally, we have to implement the last constraint which is that the sum of the weights must be equal to one. Thus, there is no leverage in the portfolio. This is quite intuitive, we have simply used a vector that have one for the first N terms and then 0. This vector will be multiplied to the weights' vector.

4 Results

In this section, we analyze and compare the different results generated by our three models with other common models used in finance, namely the Global Minimum Variance Portfolio (GMVP), the market portfolio and the Mean-Variance portfolio. In addition, we will interpret our results to ensure that our models meet the purpose of the investment strategy, i.e. to limit large losses in the event of crises. Our performance analysis will focus primarily on the following measures: the return, the volatility, the Sharpe Ratio, the Conditional Value-At-Risk (expected shortfall) and the turnover. Finally, we will also interpret the weights changes in the asset classes for the different models over time.

4.1 Graphical representation of the performance

First, we will graphically represent the cumulative performance of our models in order to have an overview of our results. Our performance analysis will break down into 3 decades. Indeed, since we optimize our models over a period from 1990 to 2020, we decompose this period into three following sub-periods: 1990 to 2000, 2000 to 2010 and 2010 to 2020.

4.1.1 First decade (1990 - 2000)

In this first decade, we can see in Figure 2 that our three models perform less well than the market. This is particularly the case during more volatile periods (since 1998). On the other hand, we see less pronounced falls during bearish market trends. This is due to the fact that our models strongly penalize large losses. Of course, we will later complete the analysis of the results with the performance measures calculated during this period.

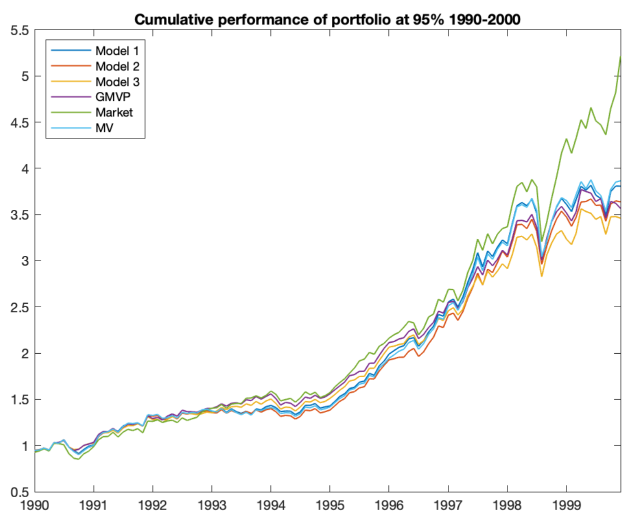

Figure 2: Cumulative performance - First decade (1990 - 2000)

4.1.2 Second decade (2000 - 2010)

We are now looking at the performance of our portfolios over the next decade (2000 to 2010). Of course, we evaluate our portfolios at the same level starting at time 0 in order to compare our models adequately. As we can see in Figure 3, it would seem that our models systematically outperformed the market during this period (in particular during the 2008 crisis). We observe that the model 1 behaves quite similarly to the Mean-Variance portfolio. It could be explained by the fact that those two are constructed in a similar fashion. In the next sub-section, we will investigate further whether the additional financial returns of the three models are accompanied by additional risk.

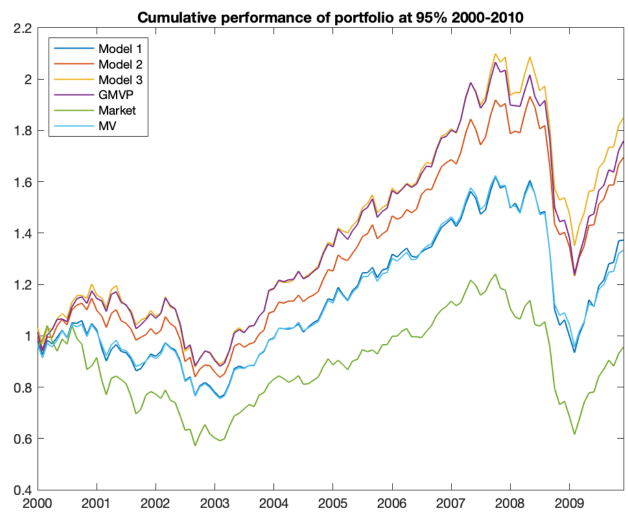

Figure 3: Cumulative performance - Second decade (2000 - 2010)

4.1.3 Last decade (2010 - 2020)

Finally, we can look at the last decade, from 2010 to 2020. We can see in Figure 4 that during upward market trends, our models tend to underperform the market. Indeed, as our models penalize large losses during crises, this must be compensated by lower returns during uptrends. In the next section concerning the sensitivity analysis, we will see the impact on financial performance if we increase the weight on the expected shortfall (CVaR) relative to returns.

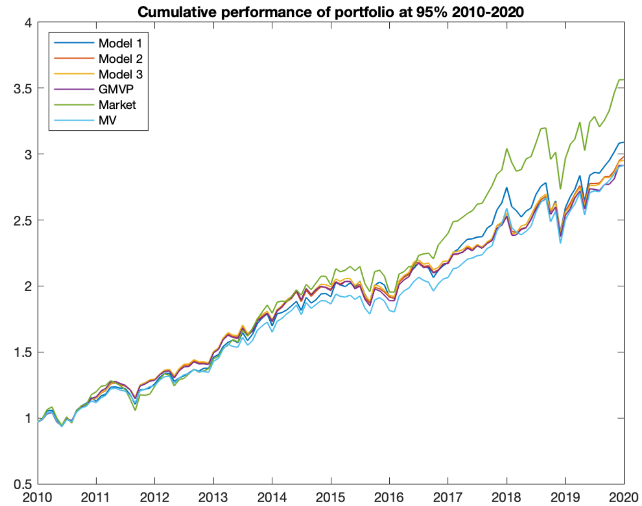

Figure 4: Cumulative performance - Third decade (2010 - 2020)

4.1.4 Overall period (1990 - 2020)

The Figure 5 represents the cumulative performance over the entire period (1990 to 2020). The only observation we can make is that our models tend to perform better during more bearish market trends.

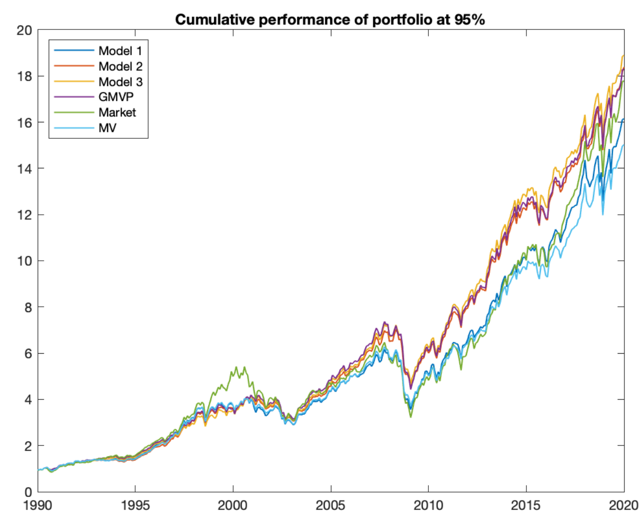

Figure 5: Cumulative performance - Overall period (1990 - 2020)

4.2 Performance measurement

In this sub-section, we will present, analyze and compare the different performance measures (return, volatility, Sharpe Ratio, expected shortfall, turnover) of our portfolios in order to identify which of our models is the most efficient in order to limit large losses during downtrends. We will proceed in the same way as in the previous sub-section, i.e. we will first analyze performance measures over the 3 decades (1990 to 2000, 2000 to 2010, 2010 to 2020) and then analyze the results over the entire period (1990 to 2020). In addition, we will compare the measures that we gathered between our different models as well as with the Market, the Global Minimum Variance Portfolio (GMVP) and the Mean-Variance portfolio (MV).

4.2.1 First decade (1990 - 2000)

The Table 1 shows different performance measures for our three models as well as for the market, GMVP and Mean-Variance portfolio in the first decade (1990 to 2000). The first thing we can see is that the expected shortfall of the model 3 is the smallest of any portfolio. On the other hand, the measure of adjusted risk (Sharpe ratio) is the lowest (which is partly due to the low returns of this model). This result shows that when we want to accept smaller losses in negative stock market events, then we have to forego a larger return. Furthermore, we can see that the returns of our three models are smaller than the market return. On the other hand, the market's expected shortfall is much higher because we are not interested in the large losses that could occur in the event of a recession or adverse events. Another important point we can highlight is that our three models are very close to the GMVP and Mean-Variance (MV) portfolio. This is due to the fact that when we penalize the variance in the GMVP and MV models, then we also penalize in some way the large losses in a crisis. However, our three models are designed in such a way that it penalizes only large unfavorable losses and not the variance in general (which can be beneficial if the variance increases in a favorable way during an economic boom).

Performance Measurement (1990 - 2000)

	Model 1	Model 2	Model 3	Market	GMVP	MV
Return (%)	14,303	13,784	13,217	17,955	13,561	14,477
Volatility (%)	12,437	11,925	11,235	13,820	11,406	12,659
Sharpe Ratio	0,774	0,764	0,760	0,961	0,779	0,774
Expected Shortfall (%)	6,764	6,220	6,204	8,008	6,513	6,874
Turnover	10,836	11,466	6,114	-	217,898	100,283

Table 1: Performance Measurement - First decade (1990 - 2000)

4.2.2 Second decade (2000 - 2010)

In this second decade (during which the subprime crisis unfolded), we can see in Table 2 that our three models have much higher returns than the market or the Mean-Variance portfolio returns. Indeed, the expected shortfall are lower than the expected shortfall of the market portfolio. This clearly shows that during very unfavorable stock market events, our three models respond favorably to the initial goal of this strategy, which is to protect ourselves against market crashes. Our strategy can be seen as a risk management tool in the event of very unfavorable market events. Similar to the previous decade, our models behave similarly to the GMVP portfolio. Finally, we can see that the Sharpe ratios are negative, indicating that returns on risk-free investments were higher than risky investments (most likely due to the 2008 crisis).

Performance Measurement (2000 - 2010)

	Model 1	**Model 2**	**Model 3**	**Market**	**GMVP**	**MV**
Return (%)	3,220	5,417	6,340	-0,446	5,808	2,920
Volatility (%)	13,888	12,323	12,608	16,611	13,044	13,030
Sharpe Ratio	-0,308	-0,169	-0,092	-0,478	-0,130	-0,351
Expected Shortfall (%)	9,697	9,016	9,151	11,070	9,628	9,346
Turnover	15,098	9,913	8,136	-	217,738	119,271

Table 2: Performance Measurement - Second decade (2000 - 2010)

4.2.3 Third decade (2010 - 2020)

For this third period from 2010 to 2020, we can see in Table 3 that our model 2 (maximizes the Sharpe ratio under the expected shortfall constraint) seems to be the most efficient among all 3 models. Indeed, this model has the highest Sharpe ratio while having the lowest expected shortfall. Moreover, we can see that during this period, which was a bull market, the market performed better than all our models (also the market has the highest Sharpe ratio). This shows the trade-off between taking advantage of the upward movements of the market and protecting against large losses in case of recession or adverse events.

Performance Measurement (2010 - 2020)						
	Model 1	**Model 2**	**Model 3**	**Market**	**GMVP**	**MV**
Return (%)	11,839	11,446	11,338	13,434	11,189	11,193
Volatility (%)	11,908	10,299	10,356	12,916	10,499	11,031
Sharpe Ratio	0,488	0,527	0,513	0,574	0,492	0,469
Expected Shortfall (%)	7,066	5,999	6,141	7,538	6,172	6,641
Turnover	14,514	6,738	5,764	-	5,695	14,616

Table 3: Performance Measurement - Third decade (2010 - 2020)

4.2.4 Overall period (1990 - 2020)

Now we are looking at the entire optimization period (1990 to 2020). We can see in Table 4 that our model 2 and 3 perform better than the market, the GMVP and the Mean-Variance portfolio. Moreover, the adjusted risks measures of models 2 and 3 are higher than the market while having a much lower expected shortfall. Model 2 is the model that best protects in a crisis without sacrificing returns relative to the market. Finally, we can also conclude that our three models behave very similarly to the GMVP.

We can also take a closer look at the turnover of our various portfolios. This notion is fundamental because it will have a strong impact on our transaction costs. Indeed, we must take into account the fact that a strategy can be very successful but could involve transaction costs that would not compensate for the additional performance of our portfolio compared to portfolios with lower transaction costs (passive investment as the market portfolio). We can see that in the first two decades, our turnovers are much lower than the GMVP or Mean-Variance portfolio. This result means that our models not only protect us from large losses while having a comparable return, but also imply much lower transaction costs. If we look at the whole period (from 1990 to 2020), we can see that the turnovers between the different portfolios are comparable, i.e. our transaction costs will also be similar compared to the first decade where the difference between our models and both the GMVP and Mean-Variance is huge.

Performance Measurement (1990 - 2020)						
	Model 1	**Model 2**	**Model 3**	**Market**	**GMVP**	**MV**
Return (%)	9,688	10,162	10,262	10,038	10,141	9,426
Volatility (%)	12,795	11,550	11,427	14,633	11,691	12,301
Sharpe Ratio	0,285	0,357	0,369	0,273	0,350	0,275
Expected Shortfall (%)	8,021	7,125	7,349	9,577	7,627	7,681
Turnover	31,562	19,827	12,470	-	10,416	32,831

Table 4: Performance Measurement - Overall period (1990 - 2020)

4.3 Financial interpretation of the evolution of weights

In this sub-section, we will look at the evolution of weights in the different 17 sectors composing our data for the three models that we constructed. We will focus particularly on the decade from 2000 to 2010, which contains the dot-com bubble and the subprime crisis in 2008.

Before analyzing the weights of the different models, it is important to note that we do not invest in every portfolio comprising the 17 industries in our data. At first glance, this seems counter-intuitive because according to the notion of diversification, the more we invest in a large number of assets, the more we reduce the idiosyncratic risk of our portfolio. Let's take a simple example to understand the reason for this phenomenon. If we have two perfectly positively correlated assets, then investing a certain proportion in these two assets will not reduce volatility. In this specific case, it is more efficient to invest only in the asset with the highest return. This explains why our models warn us not to invest in every industry portfolio from the data. In fact, some of the industries making up our data are strongly positively correlated with each other.

4.3.1 Model 1

We can see in Figure 6 that before the crisis of the 2000s, several sectors disappeared from our model, notably the Oil sector, Textiles, and Chemicals. The Financial sector was only slightly exposed between 2000 and 2008. It is only after the 2008 crisis that our model exposes this same sector. It is interesting to note that the Mining sector remains constant throughout the period (1990 to 2020). Moreover, we can see that the weight of the Mining sector has been forced to the upper-bound (at 20%).

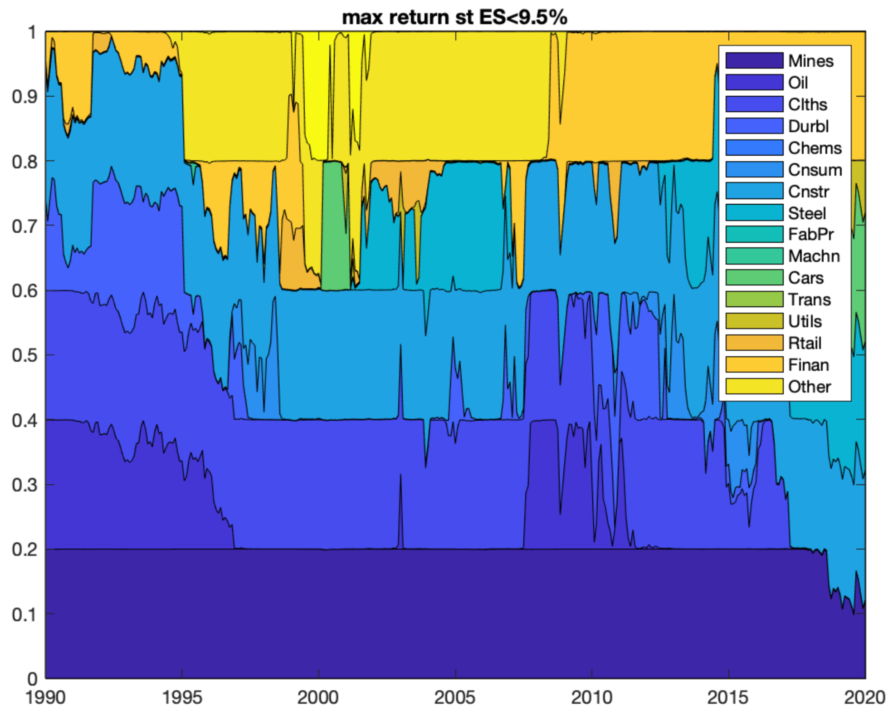

Figure 6: Evolution of weights - Model 1 (1990 - 2020)

4.3.2 Model 2

Now, we can look at weight changes over time for our second model. We can see in Figure 7 that the weights seem much more stable than our first model (in fact, in the previous sub section, we saw that the turnover of the model 2 is much lower than model 1). Similarly to model 1, many sectors disappeared from our portfolio before the crisis of the 2000s (Oil, Textiles, Consumer durables and Construction). The financial sector was also largely exposed at the end of the 2008 crisis (as in the previous model) whereas the Mining sector remained perfectly stable (constrained to the upper-bound). On the other hand, in this model we can see that the Retail sector was heavily exposed from the 2000s until the end of our period. The overexposure of the Mining sector during the entire period of our study (1990 to 2020) can be explained by the fact that it is less sensitive to recessions and consequently to adverse events.

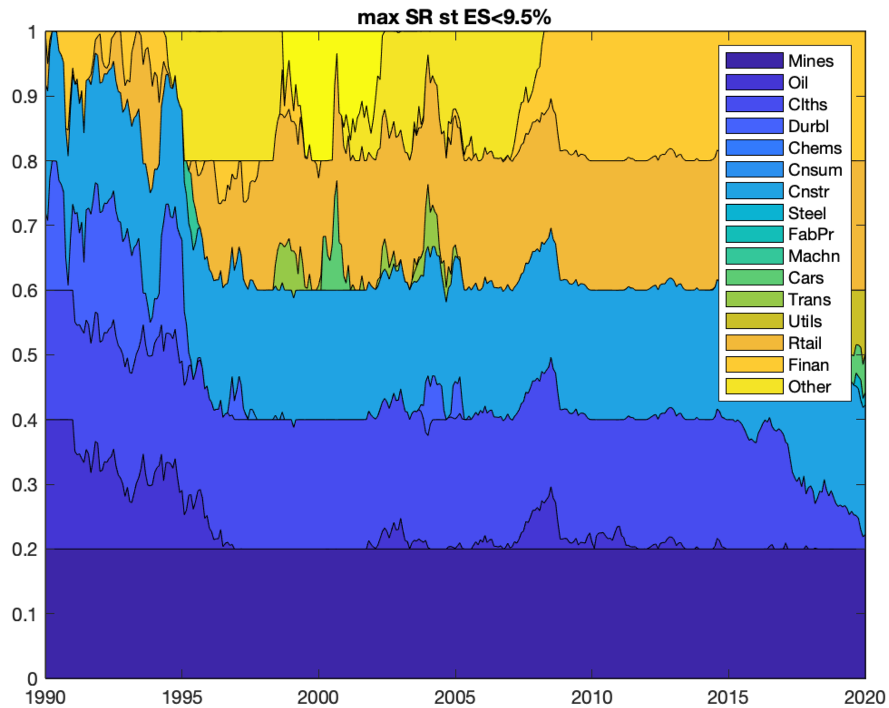

Figure 7: Evolution of weights - Model 2 (1990 - 2020)

4.3.3 Model 3

We can see in Figure 8 that throughout the period, our portfolio is mainly composed of the Mines, Oil, Textile, Construction and Retail sectors. On the other hand, as in our two previous models, the financial sector appears in 2010 just after the subprime crisis. Before the emergence of the financial sector, our portfolio was mainly composed of weak systematic portfolios that are less sensitive to market downturns, which is the aim of our investment strategy (limiting significant losses in the event of unfavorable events). As we saw in model 2, the weights seems very stable compared to model 1. As before, this is confirmed by the low turnover in this third model.

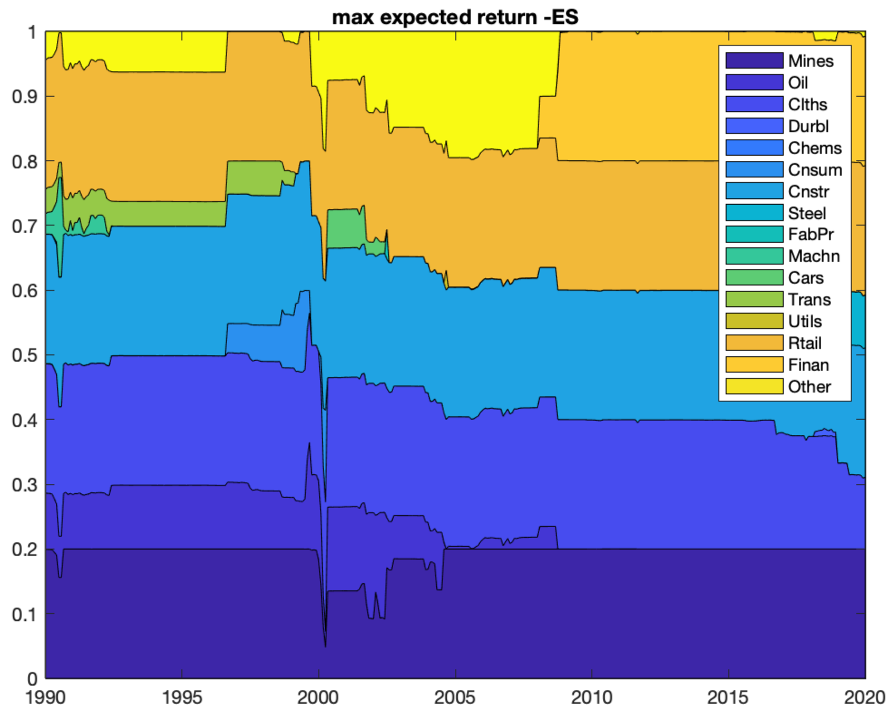

Figure 8: Evolution of weights - Model 3 (1990 - 2020)

5 Sensitivity analysis

In this section, we will perform a sensitivity analysis for both Model 1 and Model 3. We will vary some parameters of our models and see the impact on our performance measures. We focus on the following parameters presented in Table 5. For each of these analyses, we will consider the entire period of our data (1990 – 2020). It is also important to note that we keep the Value-At-Risk constant at 95%. Indeed, we do not take this parameter into account in the sensitivity analysis for two main reasons: the first reason is that there is no economic sense in comparing a certain model with a different Value-At-Risk. Secondly, if we take a 99% Value-At-Risk and calculate the expected shortfall, then due to the low number of data we have (monthly returns), the expected shortfall will be calculated as the minimum of the return's distribution, which is not the purpose of our paper.

Sensitivity Analysis - Model 1 & Model 3			
Model 1			**Model 3**
Upper-bound constraint for weights (%)	*Expected Shortfall constraint (%)*	*Rolling window (months)*	*Change in lambda parameter*
15%	9,5% (initial case)	120	1 (initial case)
20% (initial case)	12%	240	2
25%	15%	360 (initial case)	5

Table 5: Sensitivity Analysis Model 1, Model 3

5.1 Model 1

5.1.1 Upper-Bound constraint for weights

As part of this analysis, we will first compare the impact on our performance measures (return, volatility, Sharpe Ratio, expected shortfall, turnover) of relaxing the upper constraint on weights at 25% and, conversely, of constraining our weights further to 15% compared to our initial case of 20%. From a purely intuitive point of view, by relaxing the constraint (at 25%) we should not perform any worse than if we constrained our weights more (15% and 20% respectively). Indeed, we can see in Table 6, that by relaxing the constraint (at 25%), we have a better performance as well as a higher Sharpe ratio than in the other two cases. We can also see that the expected shortfall value is lower for the less constraining case (25%) than the more constraining cases (15% and 20%). We can also see that the upper constraint at 15% has a higher expected shortfall due to restrictions but it performs better than our initial model at 20%. Nevertheless, it has the highest volatility among our models. It is also important to stipulate that the market portfolio, the GMVP and the Mean-Variance portfolio have not been impacted by the change in the parameter concerned (upper limit of the weight constraint).

Sensitivity Analysis - Model 1 Upper-bound (1990 - 2020)						
	UB - 20%	**UB - 15%**	**UB - 25%**	**Market**	**GMVP**	**MV**
Return (%)	9,688	9,894	10,242	10,038	10,141	9,426
Volatility (%)	12,795	12,909	12,289	14,633	11,691	12,301
Sharpe Ratio	0,285	0,298	0,342	0,273	0,350	0,275
Expected Shortfall (%)	8,021	8,488	7,697	9,577	7,627	7,681
Turnover	31,562	28,819	31,219	-	10,416	32,831

Table 6: Sensitivity Analysis Model 1 Upper-bound (1990 - 2020)

5.1.2 Expected Shortfall constraint

We will now focus on varying the expected shortfall constraint to make it less restrictive (12%, 14.5% compared to our initial case of 9.5%). We can see in Table 7 that, by relaxing the constraint to 12%, we observe a slight increase in our returns and Sharpe ratio compared to the initial case (9.5%). The expected shortfall has slightly increased (this result is obviously expected). On the other hand, when we constrain less our model (14.5%), we can see that the model does not compensate by a higher return which makes this version less efficient compared to the first two cases (9.5% and 12%). The empirical CVaR is not increasing as expected while the constraint optimization was always biding. Finally, we can see that the turnover of the model with an expected shortfall constraint (12%) increases compared to the initial case (9.5%). Consequently, it is important to take into account the additional cost (transaction costs) provided by this model compared to the initial case. The transaction costs could be higher than the benefit of using this less restrictive model and therefore, could be less efficient.

Sensitivity Analysis - Model 1 Expected Shortfall Constraint (1990 - 2020)						
	ES - 9,5%	**ES - 12%**	**ES - 14,5%**	**Market**	**GMVP**	**MV**
Return (%)	9,688	9,926	9,770	10,038	10,141	9,426
Volatility (%)	12,795	13,190	13,222	14,633	11,691	12,301
Sharpe Ratio	0,285	0,294	0,282	0,273	0,350	0,275
Expected Shortfall (%)	8,021	8,167	8,171	9,577	7,627	7,681
Turnover	31,562	38,672	39,612	-	10,416	32,831

Table 7: Sensitivity Analysis Model 1 Expected Shortfall constraint (1990 - 2020)

5.1.3 Rolling-window

Lastly, to complete the sensitivity analysis of the first model, we can focus on changing the length of the rolling-window. Indeed, we will decrease the length to 240 and 120 months compared to our initial case of 360 months. The first thing we can see in Table 8 is that when we decrease the length of the rolling-window, then the turnover doubles compared to the initial case. This result can be explained by the fact that our model very often re-balances our portfolio to mitigate the downside risk in case of an adverse event. From a monetary point of view, this sharp increase in turnover translates into a large increase in transaction costs. We can see that the returns generated are also lower and the variance higher. As a result, these versions of the model are very inefficient compared to the initial case (360 months).

Sensitivity Analysis - Model 1 Rolling Window (1990 - 2020)

	360 months	120 months	240 months	Market	GMVP	MV
Return (%)	9,688	9,561	8,704	10,038	10,141	9,426
Volatility (%)	12,795	15,289	13,556	14,633	11,691	12,301
Sharpe Ratio	0,285	0,230	0,196	0,273	0,350	0,275
Expected Shortfall (%)	8,021	10,698	8,907	9,577	7,627	7,681
Turnover	31,562	62,371	55,169	-	10,416	32,831

Table 8: Sensitivity Analysis Model 1 Rolling Window (1990 - 2020)

5.2 Model 3

5.2.1 Changes in λ parameter

As part of this analysis, we will vary the parameter λ. This parameter makes it possible to give more or less weight to the expected shortfall related to the returns. Indeed, a higher λ does penalize more the increase of the expected shortfall and vice versa. By doing this, we are generating the efficient frontier for the CVaR criterion. We can therefore analyze the increase of this parameter on our performance measurements by increasing it to $\lambda = 2$ and $\lambda = 5$ compared to our initial case being $\lambda = 1$. Intuitively, we can expect that when the parameter increases, the returns and the expected shortfall should decrease. However, we can see in Table 9 that all performance measures (return, volatility, Sharpe Ratio, expected shortfall, turnover) remain stable despite the variation of the parameter concerned. This result can be explained by the fact that our initial data do not allow to have a much lower expected shortfall value even by increasing the weight given to the minimization of the expected shortfall.

Sensitivity Analysis - Model 3 Lambda = 1, 2, 5 (1990 - 2020)						
	Lambda = 1	**Lambda = 2**	**Lambda = 5**	**Market**	**GMVP**	**MV**
Return (%)	10,262	10,316	10,273	10,038	10,141	9,426
Volatility (%)	11,427	11,449	11,492	14,633	11,691	12,301
Sharpe Ratio	0,369	0,373	0,368	0,273	0,350	0,275
Expected Shortfall (%)	7,349	7,389	7,448	9,577	7,627	7,681
Turnover	12,470	12,036	12,570	-	10,416	32,831

Table 9: Sensitivity Analysis Model 3 $\lambda = 1, 2, 5$ (1990 - 2020)

6 Discussion

6.1 Expected returns

In our model, one of the biggest limitation is the estimation of the expected return. For remind, we only take the sample mean. Indeed, the sample mean is an unbiased estimator of the true value but it is very noisy. Hence, it is very badly estimated. It leads the optimization to over-allocate the expected returns that are overestimated and under-allocate the assets that the average returns underestimate.

There are multiple ways to improve their estimation. We explain two of them that could be implemented. Once again, we do not expect that we have to improve the estimation of the co-variance matrix because we do not have too many assets and we use a long rolling-window which should give a relatively good estimation. Moreover, we only need it in the Model 2. Nevertheless, it could be useful to mention a technique to improve the estimation of this one in the case someone would optimize with additional assets or would add a constraint over the portfolio variance.

Firstly, one could use a VAR model in order to make some predictions of the expected returns using past returns and cross-sectional relations. It is relatively a simple way to have an other approximation of the sample mean.

Secondly, one could implement the Black-Litterman approach in order to have, in our sense, the best estimation of the expected returns. Indeed, the Bayesian approach gives very good results for allocation portfolio with the expected returns. To implement such a strategy, one could follow the methodology proposed by Idzorek (2002), who propose a guide to implement such an approach. This way is quite interesting because it allows to implement some view and some others information obtain outside from the data and incorporate them in the expected returns.

Finally, regarding the co-variance matrix, in the case mentioned above, one may need to improve the estimation of the latter. One could use the approach develop by Ledoit and Wolf (2004). They shrink the co-variance matrix through a target. They proposed, in one hand, to use the constant correlation matrix and in the other hand the single-index covariance matrix.

If one of these to latter technique is use, one should be aware that these ones introduce a bias in the estimators but it reduces massively the uncertainty.

6.2 Other Limitations

First of all, we have found that we cannot reach any limit for the expected shortfall value. Indeed, we saw in the sensitivity analysis that by implementing a higher coefficient λ (and thus further penalizing the increase of the CVaR), then our model 3 remains stable. Moreover, in Model 1 and 2, we saw that it is hard to optimize for lower value restriction on the CVaR. Each time we try to optimize below 9.5%, we have problems with the Hessian matrix (the matrix became singular, hence, no inversion is possible).

Another limit that we have is the probability of scenarios (historical data). For this paper, we assume that all scenarios have the same probability. In other words, they are equally-weighted over the past periods and it simplifies the implementation for our computations. But in reality, we think that the last decade of the rolling-window should have a larger probability than the first decade. Indeed, the last ten years have probably a larger influence on today's returns and, therefore, future ones as well, than the older rolling-window's periods.

In the case of our study, we have assumed that our returns are stationary, which is generally the case in finance. However, in order to assert this assertion on returns, statistical tests of stationary would have to be carried out. We must bear in mind that if our returns are not stationary, it would have changed our results.

Furthermore, we take monthly returns data. Instead, if we choose daily measures, it would improve the precision of the estimation of the return distribution. However, it makes the implementation much more challenging. As we consider monthly returns instead of daily returns, the empirical distribution tends more to a Normal distribution, hence we have a smaller Skewness (in absolute value) and a smaller excess Kurtosis. Consequently, the CVaR computed is smaller than the CVaR computed with daily returns.

The final limitation of our project concerns variance. Indeed, we decided not to impose variance constraints in our second model. Therefore, as the expected shortfall, we could have added a restriction in the optimization problem, regarding a target for the portfolio standard deviation.

7 Conclusion

As part of this project, we implemented and analysed a portfolio optimisation with expected shortfall, which is an asset allocation strategy whose main purpose is to protect us against downside risks. We all first introduced our study by formally explaining the concepts of VaR and CVaR (expected shortfall) which are fundamental to the construction of our portfolio. We, then, set out the methodology used to build three separate models incorporating the CVaR concept. We analyzed and compared our three models with several standard portfolios (market portfolio, GMVP, Mean-variance portfolio) over several decades (1990-2000, 2000-2010, 2010-2020) as well as over the entire period (1990-2020). In addition, we also analyzed the impact of parameter changes in our models (models 1 and 3) on the performance measures as part of the sensitivity analysis, which allowed us to better understand the role of parameters in our results. Finally, we discussed the possible limitations of our approach.

In light of our various results presented in this study, we can conclude that, in general, our three models are consistent with the goal of the investment strategy, i.e., to protect the investor from negative events. In particular, we found that the expected shortfall values were lower than the value of the market portfolio. In addition, we found our models 2 and 3 to be very attractive compared to the market portfolio and the Mean-Variance portfolio. Indeed, not only do our models have a lower expected shortfall than the benchmark portfolios (market and MV) but they also have similar returns to those of the market while recording a lower turnover (comparable to those of the GMVP) and therefore, to lower transaction costs. We have seen that our investment strategy is very successful (relatively with the market portfolio) especially during the period from 2000 to 2010 when two major crises took place (Dot-com and 2008 crisis) which shows that our models (especially model 2 and 3) protect us effectively against downturns. It is also important to consider the trade-off between being protected during an economic downturn by minimizing the expected shortfall and accepting a lower return in the event of an uptrend in the market. Indeed, we have seen that in the period from 2000 to 2010, our models were better protected during the crisis (compared to the market) but performed less well during the uptrend period from 2010 to 2020.

In the current context of bearish market trends due to the drastic restrictions imposed by the world's major governments in order to fight the COVID-19 virus, an investor, wishing to protect himself more effectively against this risk, will benefit from our portfolio optimisations with expected shortfall, which should limit losses compared to passive management (market portfolio). In this context, we recommend investors to develop a strategy according to the second model.

References

[1] Pavlo Krokhmal, Jonas Palmquist, and Stanislav Uryasev. *Portfolio Optimization with Conditional Value-at-Risk Objective and Constraints*. 2002.

[2] Rockafellar, R. T. and Uryasev, S. *Optimization of conditional value-at-risk*. Journal of Risk, 2:21–42, 2000.

[3] Kull, Matthias, *Portfolio Optimization for Constrained Shortfall Risk: Implementation and IT Architecture Considerations*. Master Thesis, ETH Zurich, 2017.

[4] Kenneth R. French, Data Library *Fama-French 3 Factors and 17 Industry Portfolios*. Data extraction, 2020.
$http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html$