FACULTY
OF MATHEMATICS
AND PHYSICS
Charles University

MASTER THESIS

Vojtěch Novotný

Multi-objective portfolio optimization

Department of Probability and Mathematical Statistics

Supervisor of the master thesis: doc. RNDr. Ing. Miloš Kopa, Ph.D
Study programme: Financial and Insurance Mathematics

Prague 2025

I declare that I carried out this master thesis on my own, and only with the cited sources, literature and other professional sources. I understand that my work relates to the rights and obligations under the Act No. 121/2000 Sb., the Copyright Act, as amended, in particular the fact that the Charles University has the right to conclude a license agreement on the use of this work as a school work pursuant to Section 60 subsection 1 of the Copyright Act.

In .... date ....

Author's signature

My deepest gratitude goes to my supervisor, Associate Professor Miloš Kopa, for whose guidance it was much more pleasant to develop this work. His advice and recommendations were invaluable, making the writing process significantly smoother. I would also like to thank, first and foremost, my parents, as well as my friends, who supported me through the writing of this thesis and throughout my studies at the Faculty of Mathematics and Physics. Finally, a thank you to Anna Benešová for her feedback from an outside perspective, which helped me improve this thesis.

Title: Multi-objective portfolio optimization

Author: Vojtěch Novotný

Department: Department of Probability and Mathematical Statistics

Supervisor: doc. RNDr. Ing. Miloš Kopa, Ph.D, Department of Probability and Mathematical Statistics

Abstract: This thesis addresses portfolio optimization using mean-CVaR models. We establish the connection between CVaR and second-order stochastic dominance, introduce core concepts of portfolio optimization, and present suitable optimization methods. We propose extensions of the standard mean-CVaR model by incorporating additional objective criteria. Both additional CVaR levels and entirely new criteria are considered. We analyze the impact of these extensions on the efficient set. A possible reformulation of the presented problem using spectral risk measures is explored. Lastly, we present a numerical study using real-world data. The results validate previously presented theoretical properties of the efficient set and show a performance comparison of constructed portfolios.

Keywords: portfolio optimization, risk measures, efficient sets

Název práce: Vícekriteriální optimalizace portfolia

Autor: Vojtěch Novotný

Katedra: Katedra pravděpodobnosti a matematické statistiky

Vedoucí diplomové práce: doc. RNDr. Ing. Miloš Kopa, Ph.D, Katedra pravděpodobnosti a matematické statistiky

Abstrakt: Tato práce se zabývá optimalizací portfolia založenou na modelech mean–CVaR. Navazujeme na spojitost mezi CVaR a stochastickou dominancí druhého řádu, představujeme základní koncepty optimalizace portfolia a vhodné optimalizační metody. Navrhujeme rozšíření standardního modelu mean–CVaR o další cílová kritéria. Zahrnujeme dodatečné hladiny CVaR i zcela nová kritéria. Analyzujeme vliv těchto rozšíření na množinu eficientních portfolií. Dále zkoumáme možné přeformulování problému pomocí spektrálních rizikových měr. Nakonec představujeme numerickou studii na reálných datech. Výsledky potvrzují prezentované teoretické vlastnosti eficientní množiny a porovnávají výkonnost sestavených portfolií.

Klíčová slova: optimalizace portfolia, míra rizika, eficientní množiny

Contents

Introduction **2**

1 Stochastic dominance and risk measures **3**
1.1 First-order stochastic dominance . . . 4
1.2 Second-order stochastic dominance . . . 6
1.3 Conditional value-at-risk . . . 8

2 Multi-objective optimization **11**
2.1 Portfolio efficiency . . . 11
2.2 Optimization methods . . . 13
2.2.1 The Weighted sum method . . . 13
2.2.2 The ε-constraint method . . . 14
2.2.3 Other methods . . . 15

3 Extended mean-risk models **16**
3.1 CVaR constraints approach . . . 17
3.2 Extreme-tail CVaR constraint . . . 18
3.3 Efficient mean–CVaR frontier . . . 19
3.4 Spectral risk measure reformulation . . . 20

4 Incorporation of additional criteria **23**
4.1 Mean-risk model with transaction cost . . . 23

5 Numerical study **26**
5.1 Data . . . 26
5.2 Benchmark model . . . 30
5.3 CVaR constraints model . . . 32
5.4 Extreme-tail CVaR model . . . 32
5.5 General mean-CVaR model . . . 36
5.5.1 Model M3 . . . 36
5.5.2 Model M4 . . . 39
5.6 Model with transaction cost . . . 42
5.7 Out-of-sample performance . . . 44

Conclusion **47**

Bibliography **48**

A Additional theorems **50**
A.1 Highest CVaR portfolios . . . 50

List of Figures **52**

List of Tables **53**

List of Abbreviations and Notation **54**

Introduction

In today's world of investing, the modern portfolio selection, a mean-variance model introduced by Markowitz [1], remains popular despite its limitations. Notably, its risk measure, variance, treats very high returns and very high losses equivalently, and it is not a coherent risk measure in the sense of Artzner et al. [2], a generally desired property.

To improve on this method, Rockafellar and Uryasev proposed a new approach [3] that retains the two-criteria mean-risk model structure, but replaces variance with conditional value-at-risk (CVaR), a risk measure shown to be coherent [4], which focuses only on losses.

This thesis investigates the possibilities of extending mean-CVaR models. Specifically, we focus on models that employ three or more optimization criteria, considering both pure mean-CVaR options and the addition of some unrelated criterion. We will study the convexity of the problems and changes to the set of efficient solutions, as well as comparing the performance to other portfolio selection methods.

The thesis will be organized as follows. First, we examine the conditional value-at-risk as a risk measure and explore its connection to stochastic dominance. Next, we present the theory of multi-objective optimization, focusing on concepts such as dominance and efficiency of solutions, along with methods of finding the optimal solutions (portfolios). Then, we extend the mean-CVaR models by considering the conditional value at risk at multiple confidence levels as additional criteria, and evaluate the impact on the set of efficient solutions and the performance of the resulting portfolios. Subsequently, we examine the option of widening the scope to mean-CVaR models with completely new added criteria, specifically transaction costs.

Finally, we will perform a numerical study on ten stocks from the S&P500 index. We will present the results of optimization problems, comparing the assigned weights, in-sample and out-of-sample performance of the previously studied models and discuss the replicability of our results, as well as advantages and disadvantages of each option for the investor.

Overall, this thesis aims to offer a detailed analysis of portfolios achieved by multi-objective optimization that incorporate alternative risk measures and additional decision criteria.

1 Stochastic dominance and risk measures

Generally speaking, the goal of portfolio optimization is to identify the portfolio which maximizes the investor's expected utility, that is a portfolio which provides the best possible trade-off between risk and return given the investor's preferences. To identify this portfolio, we need a utility function $u : \boldsymbol{W} \to \mathbb{R}$, which assigns the utility to a given level of wealth from $\boldsymbol{W}$, a set of possible wealth levels.

Definition 1. *Mapping* $u : \boldsymbol{W} \to \mathbb{R}$ *is a* ***(cardinal) utility function****, if it is continuous and non-decreasing.*

The first condition (continuity) is a technical assumption, while the second reflects an economical assumption that investors prefer more wealth. The concept of utility of a portfolio is very useful, since the function is shaped according to the investor's risk attitude. For example, every risk-averse investor will have a concave utility function because losses of wealth diminish utility more than equivalent gains increase it. Because of that, a portfolio which maximizes the expected value of utility is not necessarily the portfolio with the highest expected return, but rather one that provides the best balance between risk and return based on investor preferences.

Once we have a utility function, we can compare two investments (portfolios), to see which one is preferred by the investor. However, it is virtually impossible to know exact utility function of an individual investor, since that would require the knowledge of utility for infinitely many levels of wealth. Additionally, rather than finding one portfolio for every individual investor, we would like one that is preferable for a whole set of investors, characterized by a similarity of their utility function.

Definition 2 (Levy [5] (pp. 42), modified for non-specific generator)**.** *We say that random variable (investment) X* ***strictly dominates*** *random variable Y w.r.t. generator U (set of utility functions) if for all utility functions such that $u \in U$, $\mathsf{E}_X\, u(x) \geq \mathsf{E}_Y\, u(x)$ and for at least one utility function $u_0 \in U$, there is a strict inequality. An investment is included in the* ***efficient set*** *if there is no other investment that dominates it. The efficient set includes all undominated investments.*

With this tool, we can now identify a set of optimal (efficient) portfolios for a specific type of investor. For example, if we focus only on risk-averse investors, the efficient set for this group includes only portfolios that cannot be improved without any downside.

We consider a portfolio comprising K assets. Let $\boldsymbol{r} \in \boldsymbol{R} \subset \mathbb{R}^K$ denote a random vector of future values (returns) governed by a Borel probability measure P on $\boldsymbol{R}$. A portfolio is characterized by a decision (weight) vector $\boldsymbol{w} \in \mathcal{W} \subset \mathbb{R}^K$, where $\mathcal{W}$ represents a set of feasible decisions under some constraints (e.g., budget or the unavailability of short positions).

The resulting portfolio is given by the random variable of returns $X = \boldsymbol{w}^T\boldsymbol{r}$. We denote the set of such portfolio returns by $\mathcal{X}$, a set of essentially bounded random variables from probability space $(\Omega, \mathcal{F}, \mathsf{P})$. Throughout, we use $X, Y \in \mathcal{X}$ for (returns of) investments and consider the loss variable $L_X = -X = -\boldsymbol{w}^T\boldsymbol{r}$. If X is essentially bounded, L_X is also essentially bounded and $L_X \in \mathcal{X}$.

1.1 First-order stochastic dominance

First, we consider the generator U_1, defined as the set of all utility functions. In other words, we consider investments that are optimal for any investor, regardless of the risk attitude. This approach is very broad, we might want to consider investments that would not satisfy risk-seeking investors, who adopt high-risk strategies without regard for potential losses. However, it can still be useful to study the properties of these investments, and the optimality of a portfolio for any investor should be noted.

Definition 3. *Let X, Y be investments with cumulative distribution functions F_X and F_Y, respectively. We say that investment X strictly dominates investment Y by **first-order stochastic dominance (FSD)**, denoted $F_X \succ_{FSD} F_Y$, if for all $u \in U_1$ (set of all utility functions) we have $\mathsf{E}_X\, u(x) \geq \mathsf{E}_Y\, u(x)$ and there exists $u_0 \in U_1$ such that $\mathsf{E}_X\, u_0(x) > \mathsf{E}_Y\, u_0(x)$.*

To distill a risk measure from FSD, we take a quantile approach. First, we define a quantile.

Definition 4. *Let F_X be the cumulative distribution function of a random variable X. The v**-th quantile** of the distribution of X is defined as:*

$$F_X^{(-1)}(v) = \min\{s\colon F_X(s) \geq v\},\ v \in [0,1].$$

To show the quantile approach to FSD, we need the following necessary and sufficient condition of FSD:

Theorem 1 (Hanoch and Levy [6])**.** *Let F_X, F_Y be two (cumulative) distributions. Then:*

$$F_X \succ_{FSD} F_Y \iff \Big\{F_X(x) \leq F_Y(x), \forall x \in \mathbb{R}\ \ \&\ \ \exists x_0 \in \mathbb{R} : F_X(x_0) < F_Y(x_0)\Big\}.$$

The proof of the theorem is included in [6]. To interpret this theorem for investments, we say an investment X with the return distribution defined by F_X dominates an investment Y corresponding to F_Y, if and only if, the probability that the return of X is lower or equal to an arbitrary real number x is smaller or equal to the same probability for Y, with strict inequality for at least one x. In this scenario, any reasonable investor would prefer X over Y.

Theorem 2 (Levy [5] (p. 130))**.** *Let F_X and F_Y be the cumulative distributions of the return on two investments. Then $F_X \succ_{FSD} F_Y$ if and only if:*

$$F_X^{(-1)}(v) \geq F_Y^{(-1)}(v),\ \forall v \in [0,1]$$

and there is at least one value v_0 for which a strict inequality holds.

Proof. We prove two equivalences:

1. $F_X^{(-1)}(v) \geq F_Y^{(-1)}(v),\ \forall v \in [0,1]$ if and only if $F_X(x) \leq F_Y(x), \forall x \in \mathbb{R}$. (1.1)

$\Leftarrow$ Suppose $F_X(x) \leq F_Y(x)$ for all x. Let $v_1^* \in [0,1]$. If there exists $x^* \in \mathbb{R}$:

$$x^* = \mathrm{argmin}_{x\in\mathbb{R}}\{v_1^* = F_X(x)\},\ \text{then } v_1^* = F_X(x^*) \leq F_Y(x^*) = v_2^*,$$

and for any $v_1^* \in [0,1]$:

$$\begin{aligned} F_X^{(-1)}(v_1^*) = x^* &\geq \inf\{s : F_Y(s) \geq F_Y(x^*)\} = \inf\{s : F_Y(s) \geq v_2^*\} \\ &\geq \inf\{s : F_Y(s) \geq v_1^*\} = F_Y^{(-1)}(v_1^*). \end{aligned}$$

Such x^* does not exist only if $v_1^* = 1$ and $F_X^{(-1)}(1) = \infty$ (then the implication clearly holds), or if $v_1^* = 0$ and $\{s : F_X(s) \geq 0\} = \mathbb{R}$ (then we have $\{s : F_Y(s) \geq 0\} \supseteq \{s : F_X(s) \geq 0\} = \mathbb{R}$ and the implication holds as well).

$\Rightarrow$ Suppose $F_X^{(-1)}(v) \geq F_Y^{(-1)}(v)$ for all v. Let $x^* \in \mathbb{R}$. Set:

$$v^* = F_X(x^*),\ X_1^* = \{x \in \mathbb{R} : F_X(x) = v^*\},\ X_2^* = \{x \in \mathbb{R} : F_Y(x) = v^*\}.$$

Then, using the fact X_1^* and X_2^* are left-closed:

$$\min(X_1^*) = F_X^{(-1)}(v^*) \geq F_Y^{(-1)}(v^*) = \min(X_2^*),$$

$$F_X(\min(X_1^*)) = v^* = F_Y(\min(X_2^*)) \leq F_Y(\min(X_1^*)).$$

Since F_X is constant on X_1^*, F_Y is by definition non-decreasing and $x^* \in X_1^*$, $F_X(\min(X_1^*)) \leq F_Y(\min(X_1^*))$ implies $F_X(x^*) \leq F_Y(x^*)$ and the equivalence (1.1) does hold.

2. There is a strict inequality for some v_0 in the first statement of (1.1) if and only if there is a strict inequality at some x_0 in the second.

$\Leftarrow$ Suppose that for some $x_0 \in \mathbb{R} : F_X(x_0) < F_Y(x_0)$. Then there exists $x_0' \leq x_0$ such that $F_Y(x_0') = F_Y(x_0)$ and $\forall x < x_0' : F_Y(x) < F_Y(x_0)$. Since F_X is non-decreasing, $F_X(x_0') \leq F_X(x_0) < F_Y(x_0')$ and from properties of CFD exists $x_0'' = \inf\{x : F_X(x) = F_Y(x_0')\}$, $x_0'' > x_0'$. Set $v_0 = F_X(x_0'') = F_Y(x_0')$, for which $F_X^{(-1)}(v_0) = F_X^{(-1)}(F_X(x_0'')) = x_0'' > x_0' = F_Y^{(-1)}(F_Y(x_0')) = F_Y^{(-1)}(v_0)$.

$\Rightarrow$ Suppose that for some $v_0 \in [0,1] : F_X^{(-1)}(v_0) > F_Y^{(-1)}(v_0)$.

- If $F_X(F_Y^{(-1)}(v_0)) = F_X(F_X^{(-1)}(v_0))$, then $F_Y^{(-1)}(v_0) \geq F_X^{(-1)}(v_0)$ as $F_X^{(-1)}(v_0) = \inf\{s : F_X(s) \geq v_0\}$, which contradicts our assumption.
- If $F_X(F_Y^{(-1)}(v_0)) < F_X(F_X^{(-1)}(v_0))$, then $F_X(F_Y^{(-1)}(v_0)) < v_0$, $v_0 \leq F_Y(F_Y^{(-1)}(v_0))$ and the equivalence of the strict inequality holds.

□

Using this theorem, we can show the connection between FSD and a popular risk measure, value-at-risk (VaR). Before defining the risk measure, we need to introduce uniform notation. Generally, risk measure can be any mapping $\rho \colon \mathcal{X} \to \mathbb{R}$.

Definition 5. *The **VaR (value-at-risk) at level** $\alpha \in [0,1]$ of a loss random variable L is the value:*

$$\mathrm{VaR}_\alpha(L) = \min\{z \in \mathbb{R} \colon\ P[L \le z] \ge \alpha\}.$$

*Additionally, **upper VaR** is defined as:*

$$\mathrm{VaR}^+_\alpha(L) = \inf\{z \in \mathbb{R} \colon\ P[L \le z] > \alpha\}.$$

The interpretation of the value-at-risk is that if a portfolio has the loss L, then $\mathrm{VaR}_\alpha(L)$ is equal to loss level that will not be exceeded with probability α, which is usually chosen very high, e.g. 0.95 or 0.99. It always holds $\mathrm{VaR}_\alpha(L) \le \mathrm{VaR}^+_\alpha(L)$ and the values are equal except when F_L is constant on a certain interval.

Corollary. Let X, Y denote returns of two portfolios with continuous, strictly increasing distribution functions F_X, F_Y. Then $F_X \succ_{FSD} F_Y$ if and only if $\mathrm{VaR}_\alpha(-X) \le \mathrm{VaR}_\alpha(-Y)$, for each $\alpha \in [0,1]$ and exists α_0 such that the inequality is strict.

Proof. We have:

$$\begin{aligned}
\mathrm{VaR}_\alpha(-X) &= \min\{z \in \mathbb{R} \colon\ \mathsf{P}[-X \le z] \ge \alpha\} \\
&= \min\{z \in \mathbb{R} \colon\ \mathsf{P}[X \ge -z] \ge \alpha\} \\
&= \min\{z \in \mathbb{R} \colon 1 - F_X(-z) \ge \alpha\} \\
&= \inf\{z \in \mathbb{R} \colon F_X(-z) \le 1 - \alpha\} \\
&= -\sup\{x \in \mathbb{R} \colon F_X(x) \le 1 - \alpha\} \\
&= -\inf\{x \in \mathbb{R} \colon F_X(x) > 1 - \alpha\} \\
&= -\min\{x \in \mathbb{R} \colon F_X(x) \ge 1 - \alpha\} \\
&= -F_X^{(-1)}(1 - \alpha),
\end{aligned} \tag{1.2}$$

where the fourth equivalence holds for any F_X continuous and the seventh equivalence holds for any F_X strictly increasing. The proposition holds based on the combination of (1.2) and Theorem 2. □

This connection between value-at-risk and first-order stochastic dominance is only one of multiple reasons why it is a widely adopted risk measure. However, it also has multiple flaws, which will be shown later.

1.2 Second-order stochastic dominance

In the previous section, we considered the set of all possible utility functions. Although that approach could be useful sometimes, it is often overly broad for practical use. If we focus on risk-averse investors, who are characterized by concave utility functions, we can reduce the set of utility functions without excluding most real-world investor preferences. This restriction might provide a risk measure with significant advantages compared to value-at-risk (VaR).

Definition 6. *Let X, Y be investments with cumulative distribution functions F_X and F_Y, respectively. We say that investment X strictly dominates investment Y by **second-order stochastic dominance (SSD)**, denoted by $F_X \succ_{SSD} F_Y$, if for all $u \in U_2$ (set of all concave utility functions) we have $E_X\, u(x) \ge E_Y\, u(x)$ and there exists $u_0 \in U_2$ such that $E_X\, u_0(x) > E_Y\, u_0(x)$.*

In this definition, we do not require strict concavity, therefore risk-neutral investors are also included. From here, we can proceed very similarly to FSD, but we use integrated distribution and quantile functions.

Definition 7. *Let X be a random variable with cumulative distribution function F_X. The* ***integrated distribution function*** *is defined as $F_X^{(2)}(t) = \int_{-\infty}^{t} F_X(s)ds$.*

Definition 8. *Let X be a random variable with cumulative distribution function F_X and quantile function $F_X^{(-1)}$. The* ***integrated quatile function*** *is defined as $F_X^{(-2)}(p) = \int_0^p F_X^{(-1)}(q)dq,\ p \in (0,1];\ F_X^{(-2)}(0) = 0$.*

Similarly to the FSD case, it can be shown (see Hanoch and Levy [6]) that $F_X^{(2)}(t) \leq F_Y^{(2)}(t)$ for every $t \in \mathbb{R}$, with a strict inequality for some $t_0 \in \mathbb{R}$, is a necessary and sufficient condition for second-order stochastic dominance $F_X \succ_{SSD} F_Y$. Furthermore, according to Levy [5, pp. 133] $F_X \succ_{SSD} F_Y$ if and only if $F_X^{(-2)}(p) \geq F_Y^{(-2)}(p)$ for every $p \in [0,1]$ and there exists $p_0 \in [0,1]$ such that the inequality is strict.

The simplicity of the VaR design can be an advantage, but it can also be its biggest flaw. Being a quantile of a loss distribution, it tells the investor what is the value of the worst outcome in most cases, e.g. with 99% probability, but it does not contain any information about the remaining one percent other than its lower bound. That is one of the reasons why an alternative risk measure, conditional value-at-risk (CVaR) is often preferred.

Definition 9. *The* ***CVaR (conditional value-at-risk) at level*** *$\alpha \in [0,1]$ of a loss random variable L is the mean of its α-tail distribution F_α:*

$$\mathrm{CVaR}_\alpha(L) = \mathsf{E}_{F_\alpha}[L], \quad F_\alpha(x) = \begin{cases} \frac{F_L(x)-\alpha}{1-\alpha}, & \text{if } x \geq \mathrm{VaR}_\alpha(L), \\ 0, & \text{if } x < \mathrm{VaR}_\alpha(L). \end{cases}$$

Remark. This definition of CVaR works for general loss distribution. When considering continuous loss distribution, according to Rockafellar and Uryasev [3], the definition is equivalent to the following optimization formula:

$$\mathrm{CVaR}_\alpha(L) = \min_{z\in\mathbb{R}} f_\alpha(z,L) = \min_{z\in\mathbb{R}}\Big(z + \tfrac{1}{1-\alpha}\,\mathsf{E}[\max(L-z,0)]\Big),\ \alpha \in (0,1), \quad (1.3)$$

and moreover $\mathrm{argmin}_{z\in\mathbb{R}} f_\alpha(z,L) = \Big[\mathrm{VaR}_\alpha(L), \mathrm{VaR}_\alpha^+(L)\Big]$. Persistence of this formula for general loss distributions was later shown by Rockafellar and Uryasev [4].

If we again consider portfolios with continuous, strictly increasing distribution function, using (1.2) and an alternative expression of CVaR of loss $L_X = -X$ from Acerbi and Tasche [7], we can show:

$$\begin{aligned} \mathrm{CVaR}_\alpha(L_X) &= \frac{1}{1-\alpha}\int_\alpha^1 \mathrm{VaR}_t(L_X)dt \\ &= \frac{1}{1-\alpha}\int_\alpha^1 -F_X^{(-1)}(1-t)dt \\ &= -\frac{1}{1-\alpha}\int_0^{1-\alpha} F_X^{(-1)}(s)ds\Big|_{s=1-t} \\ &= -\frac{F_X^{(-2)}(1-\alpha)}{1-\alpha}. \end{aligned} \quad (1.4)$$

Therefore, similarly to FSD and VaR, for continuous, strictly increasing distribution functions, $F_X \succ_{SSD} F_Y$, if and only if $\mathrm{CVaR}_\alpha(-X) \leq \mathrm{CVaR}_\alpha(-Y)$, for each $\alpha \in [0,1]$ and there exists α_0 such that the inequality is strict.

This shows us that CVaR might be a useful risk measure, as its minimization should benefit any risk-averse investor. In the next section, we will further examine its properties outside of the stochastic dominance framework and show why we use it as a risk measure in our portfolio optimization.

1.3 Conditional value-at-risk

In the previous section, we defined a risk measure called conditional value-at-risk (CVaR) and shown its connection to second-order stochastic dominance. When applied to a portfolio, CVaR at level α tells us the mean value of losses that exceed VaR at level α. In other words, if we set $\alpha = 0.95$, then VaR_α is the loss value that will not be exceeded with probability 95 percent and CVaR_α is the mean value of losses that match or exceed this threshold, i.e., the worst 5 percent of losses. We will denote CVaR at level $\alpha = 0.95$ as CVaR_{95}, using the percentile index (95) instead of the probability level (0.95) for improved readability.

The described difference could be considered the first reason why CVaR is often considered superior to VaR. An artificial but useful example would be two investments, where X always returns one percent per week and Y returns one percent per week in 96 weeks out of 100, but loses fifty percent in the remaining four. In terms of VaR_{95}, both investments are equivalent, while CVaR_{95} can distinctly identify the preferred investment. This is related to a more theoretical, but very important property of risk measures, coherence.

Definition 10 (Artzner et al. [2]). *A risk measure $\rho\colon \mathcal{X} \to \mathbb{R}$ for random losses from $\mathcal{X}$ is said to be **coherent** if it follows these four axioms:*

(i) Translation invariance: For all $L_X \in \mathcal{X}$ and real numbers c, we have:

$$\rho(L_X + c) = \rho(L_X) + c.$$

(ii) Subadditivity: For all L_X and $L_Y \in \mathcal{X}$:

$$\rho(L_X + L_Y) \leq \rho(L_X) + \rho(L_Y).$$

(iii) Positive homogeneity: For all $\lambda \geq 0$ and all $L_X \in \mathcal{X}$:

$$\rho(\lambda L_X) = \lambda \rho(L_X).$$

(iv) Monotonicity: For all L_X and $L_Y \in \mathcal{X}$ with $L_X \leq L_Y$, we have:

$$\rho(L_X) \leq \rho(L_Y).$$

Coherence can also be defined for random returns instead of losses by appropriate adjustment to translation invariance and monotonicity. The four axioms are considered desirable properties of a risk measure. Monotonicity and positive homogeneity are very straightforward, monotonicity requires that if one investment produces higher or equal loss to different one almost surely, it is considered to have higher or equal risk. Positive homogeneity means that if our position in a portfolio doubles, the associated risk doubles as well. This creates the assumption that a linear increase of loss leads to a linear increase in risk.

Translation invariance requires that if loss is increased by a constant, the associated risk should increase by the same value. This is a useful property for determination of the capital requirement, a volume of capital which, when held, decreases our risk to zero. Lastly, subadditivity of a risk measure tells us that when we diversify our investment by combining two different portfolios, our risk does not increase compared to the two individual investments of the same volumes.

Theorem 3 (Rockafellar and Uryasev [4])**.** *CVaR, as defined in 9, is a coherent risk measure, while VaR is not.*

Coherence of CVaR can be proven using the optimization formula (1.3) (see Rockafellar and Uryasev [4]), modified to express CVaR using VaR:

$$\begin{aligned} \mathrm{CVaR}_\alpha(L) &= \min_{z\in\mathbb{R}}\Big(z + \tfrac{1}{1-\alpha}\,\mathsf{E}[\max(L-z,0)]\Big) \\ &= \mathrm{VaR}_\alpha(L) + \tfrac{1}{1-\alpha}\,\mathsf{E}[\max(L-\mathrm{VaR}_\alpha(L),0)]. \end{aligned} \tag{1.5}$$

VaR is not a coherent risk measure because it is not subadditive. This can be shown using a simple example of two bonds.

Example 1. Let X, Y be bonds with nominal value 1, that do not show any loss with probability 0.97 and lose 0.5 with probability 0.03, X, Y are independent. Clearly, $\mathrm{VaR}_{0.95}(L_X) = \mathrm{VaR}_{0.95}(L_Y) = 0$, but the combination:

$$L_X + L_Y = \begin{cases} 0, & \text{with probability } p_1 = 0.97\cdot 0.97 = 0.9409, \\ 0.5, & \text{with probability } p_2 = 0.97\cdot 0.03\cdot 2 = 0.0582, \\ 1, & \text{with probability } p_3 = 0.03\cdot 0.03 = 0.0009, \end{cases}$$

and $\mathrm{VaR}_{0.95}(L_X + L_Y) = 0.5 > \mathrm{VaR}_{0.95}(L_X) + \mathrm{VaR}_{0.95}(L_Y)$. △

Beyond coherence, convexity in portfolio weights further enhances the appeal of CVAR for optimization. It is not a desired property from an economical viewpoint, but rather a computational one. Convexity with respect to loss L is a direct consequence of subadditivity and positive homogeneity, but Rockafellar and Uryasev [4] show its convexity with respect to the decision (weights) vector $\boldsymbol{w}$, without the use of coherence.

Theorem 4. *If loss function L is convex with respect to decision vector $\boldsymbol{w}$, then $\mathrm{CVaR}_\alpha(L)$ is convex with respect to $\boldsymbol{w}$ as well and $f_\alpha(z, L)$ from (1.3) is jointly convex in $(z, \boldsymbol{w})$.*

Likewise, if the loss is sublinear with respect to $\boldsymbol{w}$, then $\mathrm{CVaR}_\alpha(L)$ is sublinear with respect to $\boldsymbol{w}$ and $f_\alpha(z, L)$ is jointly sublinear in $(z, \boldsymbol{w})$.

Remark. Function $h(x)$ is sublinear if $h(x + y) \leq h(x) + h(y)$ and $h(\lambda x) = \lambda h(x)$ for $\lambda > 0$.

Proof. The joint convexity of $f_\alpha(z, L)$ is a consequence of linearity of expectation and convexity if the function $(z, \boldsymbol{w}) \to max(f(\boldsymbol{w}, \boldsymbol{r}) - z, 0)$ when $f(\boldsymbol{w}, \boldsymbol{r})$ is convex in $\boldsymbol{w}$. Convexity of CVaR in $\boldsymbol{w}$ follows from the fact that minimizing jointly convex function $f_\alpha(z, L)$ with respect to z results in convex function with respect to $\boldsymbol{w}$ (see Rockafellar [8, pp. 38]).

According to [8], sublinearity is equivalent to the combination of convexity with positive homogeneity. The convexity of CVaR was already shown and from (1.5), for any $\lambda > 0$:

$$\begin{aligned}\mathrm{CVaR}_\alpha(\lambda L) &= \mathrm{VaR}_\alpha(\lambda L) + \tfrac{1}{1-\alpha}\,\mathsf{E}\Big[max(\lambda L - \mathrm{VaR}_\alpha(\lambda L), 0)\Big] \\ &= \lambda\,\mathrm{VaR}_\alpha(L) + \tfrac{1}{1-\alpha}\lambda\,\mathsf{E}\Big[\max(L - \mathrm{VaR}_\alpha(L), 0)\Big] = \lambda\,\mathrm{CVaR}_\alpha(L).\end{aligned}$$

□

Remark. The chosen loss function $L = -\boldsymbol{w}^T\boldsymbol{r}$ is linear, therefore both convex and sublinear with respect to $\boldsymbol{w}$ and both parts of Theorem 4 apply to it.

2 Multi-objective optimization

In many decision-making scenarios, there are multiple conflicting objectives, and an optimal solution cannot be selected using a single criterion. In portfolio selection investors typically seek to maximize returns while minimizing the associated risk. These are two goals that align very rarely, because the highest-return portfolios are seldom the lowest-risk ones. In Markowitz or simple mean-CVaR models, we have two criteria, return and risk, and we look for solutions that are somehow optimal in that scenario. The set of preferred solutions is called efficient, a set of solutions that cannot improve in one objective without degrading another.

In this thesis, we focus on models that consider multiple risk criteria, which means three or more criteria overall. A model with M criteria can be written using $\boldsymbol{f} = (f_1, \dots, f_M)^T,\ f_m \colon \mathbb{R}^K \to \mathbb{R}$, where f_1 is a function of return, and $f_2, \dots, f_M$ are some risk functions. Therefore, for a feasible set $\mathcal{W}$, we get the model:

$$\begin{aligned} \text{“}\min_{\boldsymbol{w}}\text{”}\ & \boldsymbol{f}(\boldsymbol{w}), \\ \text{s.t. } & \boldsymbol{w} \in \mathcal{W}. \end{aligned} \tag{2.1}$$

This is rather a simplification, for example, profit should be maximized, not minimized, but it can serve well to show the problem we need to solve.

2.1 Portfolio efficiency

We begin by introducing the concept of an ***ideal solution***, a theoretical solution that performs optimally across all objectives. As noted above, cases where one portfolio both maximizes return and minimizes risk are exceedingly rare.

Definition 11. *Vector $\tilde{\boldsymbol{w}} \in \mathcal{W}$ is an* ***ideal solution*** *of the multi-objective problem* (2.1) *if:*

$$\tilde{\boldsymbol{w}} \in \bigcap_{m=1}^{M} \arg\min_{\boldsymbol{w}\in\mathcal{W}} f_m(\boldsymbol{w}).$$

Since an ideal solution often does not exist, we need a more practical attribute that would enable selection of the desired solutions. Therefore, we introduce the concept of dominance and efficient solutions similar to efficiency in random variables from the first chapter.

Definition 12 (Ehrgott [9] (pp. 24))**.** *A feasible solution $\hat{\boldsymbol{w}} \in \mathcal{W}$ is called* ***efficient*** *or* ***Pareto optimal****, if there is no other solution $\boldsymbol{w} \in \mathcal{W}$ such that $\boldsymbol{f}(\boldsymbol{w}) \leq \boldsymbol{f}(\hat{\boldsymbol{w}})$ and $\boldsymbol{f}(\boldsymbol{w}) \neq \boldsymbol{f}(\hat{\boldsymbol{w}})$. For $\boldsymbol{w}^1, \boldsymbol{w}^2 \in \mathcal{W}$ and $\boldsymbol{f}(\boldsymbol{w}^1) \leq \boldsymbol{f}(\boldsymbol{w}^2)$, $\boldsymbol{f}(\boldsymbol{w}^1) \neq \boldsymbol{f}(\boldsymbol{w}^2)$ we say that $\boldsymbol{w}^1$* ***dominates*** *$\boldsymbol{w}^2$. The set of all efficient solutions $\hat{\boldsymbol{w}} \in \mathcal{W}$ is called the* ***efficient set*** *and denoted $\mathcal{W}_E$.*

In the context of mean-risk portfolio selection, a portfolio is efficient if there is no other portfolio with simultaneously lower (or equal) risk and higher (or equal) return. Since every inefficient portfolio is, by definition, dominated by an efficient portfolio, we only consider investments in efficient portfolios.

When considering portfolios without short positions and limited budget, assuming that there are no other constraints, the set of all feasible solutions $\mathcal{W}$ is a bounded, closed, thus compact, and also convex set.

The following single-criterion optimization problem is called a **convex optimization problem**:

$$\begin{aligned} \min_{\boldsymbol{w}} \ & f_1(\boldsymbol{w}), \\ \text{s.t. } & f_i(\boldsymbol{w}) \leq 0, \ i = 2, \ldots, M+K, \\ & \sum_{k=1}^{K} w_k = \text{“}budget\text{”}, \end{aligned} \tag{2.2}$$

where $f_1, \ldots, f_{M+K}$ are convex functions and “budget” sets the sum of all weights. In portfolio optimization, it defines the volume of assets to be acquired.

This formulation is beneficial, as some optimization methods transform all but one function from the utility function $\boldsymbol{f}$ into constraints, and the remaining constraints guarantee $\boldsymbol{w} \in \mathcal{W}$. Convex programming problems have multiple advantageous properties. For example, the set of all feasible solutions is convex, any local minimum is also a global minimum, and the Karush–Kuhn–Tucker optimality conditions are sufficient for a minimum (see Mishra et al. [10]). Because of this, convex programming problems are favorable for computation.

The convexity of a problem can be beneficial beyond single-objective problems. We will define an equivalent property for problems similar to (2.1) and study its consequences.

Definition 13 (Miettinen [11] (pp. 7))**.** *A multi-objective problem with utility function $\boldsymbol{f} = (f_1, \ldots, f_M)$, $f_m \colon \mathbb{R}^K \to \mathbb{R}$ and a set of feasible solutions $\mathcal{W}$ is* ***convex*** *if $\mathcal{W}$ is a convex set and f_m, $m = 1, \ldots, M$ are convex functions.*

If $\mathcal{W}$ is compact and f_m, $m = 1, \ldots, M$ are continuous, the efficient set of a convex multi-objective optimization problem is closed.

Theorem 5. *The efficient set $\mathcal{W}_E$ of a multi-objective optimization problem with continuous utility function and compact set of feasible solutions is closed.*

Proof. If $\mathcal{W}_E$ is empty, then the statement is trivial. For non-empty $\mathcal{W}_E$, we assume that $\boldsymbol{w}^*$ is a feasible solution such that $\boldsymbol{w}^* \in \operatorname{cl} \mathcal{W}_E$ (closure of $\mathcal{W}_E$), but $\boldsymbol{w}^* \notin \mathcal{W}_E$. Then there exists $\hat{\boldsymbol{w}} \in \mathcal{W}_E$, such that $\hat{\boldsymbol{w}}$ dominates $\boldsymbol{w}^*$ and a sequence $\{\boldsymbol{w}_n\}_{n=1}^{\infty} \subset \mathcal{W}_E$ of points converging to $\boldsymbol{w}^*$ for $n \to \infty$. Since $f_1, \ldots, f_M$ are continuous, we have $\lim_{n \to \infty} f_m(\boldsymbol{w}_n) = f_m(\boldsymbol{w}^*)$ for all $m = 1, \ldots, M$. Therefore, for sufficiently large n, $\hat{\boldsymbol{w}}$ dominates $\boldsymbol{w}_n$, contradicting the assumption that $\{\boldsymbol{w}_n\} \subset \mathcal{W}_E$. Hence, $\operatorname{cl} \mathcal{W}_E \subset \mathcal{W}_E$ and $\mathcal{W}_E$ is closed. □

To show additional properties, we need to define the following three properties of a set.

Definition 14 (Ehrgott [9] (pp. 62, 67))**.** *A set $\mathcal{Y} \in \mathbb{R}^p$ is called $\mathbb{R}^p_{\geq}$-**closed** if $\mathcal{Y} + \mathbb{R}^p_{+} := \{\mathbf{y} + \mathbf{r} \colon \ \mathbf{y} \in \mathcal{Y}, \ \mathbf{r} \in \mathbb{R}, \ \mathbf{r} \geq \mathbf{0}, \ \mathbf{r} \neq \mathbf{0}\}$ is a closed set. $\mathcal{Y}$ is called $\mathbb{R}^p_{\geq}$-**convex** if $\mathcal{Y} + \mathbb{R}^p_{\geq} := \{\mathbf{y} + \mathbf{r} \colon \ \mathbf{y} \in \mathcal{Y}, \ \mathbf{r} \in \mathbb{R}, \ \mathbf{r} \geq \mathbf{0}\}$ is a convex set.*

Definition 15 (Ehrgott [9] (pp. 86))**.** *A set $\mathcal{Y} \in \mathbb{R}^p$ is called **not connected** if there exist sets $\mathcal{Y}_1$, $\mathcal{Y}_2 \in \mathbb{R}^p$ such that $\mathcal{Y} = \mathcal{Y}_1 + \mathcal{Y}_2$, where $\mathcal{Y}_1, \mathcal{Y}_2 \neq \emptyset$, $\operatorname{cl} \mathcal{Y}_1 \cap \mathcal{Y}_2 = \mathcal{Y}_1 \cap \operatorname{cl} \mathcal{Y}_2 = \emptyset$. Otherwise, $\mathcal{Y}$ is called **connected**.*

To examine the connectedness of $\mathcal{W}_E$, we first need an auxiliary theorem on the properties of $\mathbf{f}(\mathcal{W})$.

Theorem 6 (Hartley [12])**.** *If $\mathcal{Y}$ is non-empty, $\mathbb{R}^p_{\geq}$-closed and $\mathbb{R}^p_{\geq}$-convex,*

$$\mathcal{S}(\mathcal{Y}) := \bigcup_{\lambda\in\mathbb{R}^K,\ \lambda>0} \mathcal{S}(\lambda,\mathcal{Y}); \qquad \mathcal{S}(\lambda,\mathcal{Y}) := \left\{\hat{\mathbf{y}} \in \mathcal{Y}\colon\ \langle\lambda,\hat{\mathbf{y}}\rangle = \min_{\mathbf{y}\in\mathcal{Y}}\langle\lambda,\mathbf{y}\rangle\right\}$$

and $\mathcal{Y}_N$ denotes the non-dominated points in $\mathcal{Y}$, the following inclusions hold:

$$\mathcal{S}(\mathcal{Y}) \subset \mathcal{Y}_N \subset \operatorname{cl}\mathcal{S}(\mathcal{Y}).$$

Using this theorem, we now get to a key result about the efficient set.

Theorem 7 (Ehrgott [9] (pp. 91))**.** *Let $\mathcal{W} \subset \mathbb{R}^n$ be a convex and compact set. Assume that all objective functions f_k are convex. Then $\mathcal{W}_E$ is connected.*

The proof provided by Ehrgott [9, (pp. 91)] contains a mistake: The author's claim "Because $\mathcal{W}$ is compact and convex and f_k are convex and continuous, $\mathcal{Y} = f(\mathcal{W})$ is also compact and convex." does not hold, however convexity is used only to show that the assumptions of Theorem 6 hold. This theorem does not require convexity, but rather that $\mathcal{Y}$ is non-empty, $\mathbb{R}^p_{\geq}$-closed and $\mathbb{R}^p_{\geq}$-convex, properties that can be shown for any $\mathcal{W} \subset \mathbb{R}^n$ non-empty, convex and compact set, and all objective functions f_k convex:

Let $\boldsymbol{w}_1$ and $\boldsymbol{w}_2$ in $\mathcal{W}$. From convexity of $\mathcal{W}$ and f_k, for each $\lambda \in [0,1]$ holds $\lambda\boldsymbol{w}_1 + (1-\lambda)\boldsymbol{w}_2 \in \mathcal{W}$ and $f_k(\lambda\boldsymbol{w}_1 + (1-\lambda)\boldsymbol{w}_2) \leq \lambda f_k(\boldsymbol{w}_1) + (1-\lambda)f_k(\boldsymbol{w}_2)$ for all $k = 1,\ldots,p$. Therefore exists $\mathbf{c} = (c_1,\ldots,c_p)^T \in \mathbb{R}^p_{\geq}$ such that:

$$\mathbf{f}(\lambda\boldsymbol{w}_1 + (1-\lambda)\boldsymbol{w}_2) + \mathbf{c} = \lambda\mathbf{f}(\boldsymbol{w}_1) + (1-\lambda)\mathbf{f}(\boldsymbol{w}_2),$$

so $\mathbf{f}(\lambda\boldsymbol{w}_1 + (1-\lambda)\boldsymbol{w}_2) \in \mathbf{f}(\mathcal{W}) + \mathbb{R}^p_{\geq}$ and $\mathbf{f}(\mathcal{W})$ is $\mathbb{R}^p_{\geq}$-convex. According to Theorem 5, $\mathbf{f}(\mathcal{W})$ is closed, therefore also $\mathbb{R}^p_{\geq}$-closed. Lastly, if $\mathbf{f}(\mathcal{W})$ is empty, $\mathcal{W}_E$ is connected by definition.

2.2 Optimization methods

Various methods exist for identifying the set of efficient solutions in a multi-objective optimization problem, and the choice should align with the specific characteristics of a problem. This section focuses primarily on two options, the weighted sum method and the ε-constraint method.

2.2.1 The Weighted sum method

According to Ehrgott [9], the "simplest" method to solve multi-objective problems is the weighted sum method. In this method, we solve:

$$\min_{\boldsymbol{w}\in\mathcal{W}} \sum_{i=1}^{M} \lambda_i f_i(\boldsymbol{w}), \tag{2.3}$$

for some vector of weights $\boldsymbol{\lambda} \in \mathbb{R}^M_+$. This method is capable of finding all solutions of convex multi-objective problems, however, for non-convex problems, it may work very poorly, as can be seen in the following theorem.

Theorem 8 (Ehrgott [9] (pp. 97)). *Let* $\hat{\boldsymbol{w}}$ *be an optimal solution of* (2.3) *for some* $\lambda \in \mathbb{R}^M_+$.

- *If* $\lambda > 0$, $\hat{\boldsymbol{w}}$ *is an efficient solution of the corresponding multi-objective problem with the feasible set* $\mathcal{W}$ *and* $\boldsymbol{f} = (f_1, \dots, f_M)$.
- *If the solution* $\hat{\boldsymbol{w}}$ *is unique, it is an efficient solution of the corresponding multi-objective problem.*

Let $\mathcal{W}$ *be a compact, nonempty, convex set and* $f_1, \dots, f_M$ *be convex functions.*

- *If* $\hat{\boldsymbol{w}}$ *is an efficient solution of the corresponding multi-objective problem, there exists* $\lambda \in \mathbb{R}^M_+$, $\lambda \neq 0$ *such that* $\hat{\boldsymbol{w}}$ *is an optimal solution of* (2.3).

We will prove the first part of the theorem. The proof of the second part can be found, for example, in Dupačová et al. [13, pp. 125].

Proof. We assume that there exist $\boldsymbol{w}^* \in \mathcal{W}$ such that $\boldsymbol{f}(\boldsymbol{w}^*) \leq \boldsymbol{f}(\hat{\boldsymbol{w}})$ and $m \in \{1, \dots, M\}$ such that $f_m(\boldsymbol{w}^*) < f_m(\hat{\boldsymbol{w}})$.

- If $\lambda_m > 0$, we have $\lambda_i f_i(\boldsymbol{w}^*) \leq \lambda_i f_i(\hat{\boldsymbol{w}})$ for each $i \in \{1, \dots, M\}$ and $\lambda_m f_m(\boldsymbol{w}^*) < \lambda_m f_m(\hat{\boldsymbol{w}})$, therefore $\sum_{i=1}^M \lambda_i f_i(\boldsymbol{w}^*) < \sum_{i=1}^M \lambda_i f_i(\hat{\boldsymbol{w}})$, which is a contradiction to the assumption that $\hat{\boldsymbol{w}}$ is an optimal solution of (2.3).
- If $\lambda_m = 0$, then similarly $\sum_{i=1}^M \lambda_i f_i(\boldsymbol{w}^*) \leq \sum_{i=1}^M \lambda_i f_i(\hat{\boldsymbol{w}})$ and $\hat{\boldsymbol{w}}$ is not a unique solution of (2.3).

□

This theorem indicates that, under certain conditions, the weighted sum method can effectively identify efficient solutions. However, in non-convex scenarios, it might find only a small subset of all efficient solutions.

Example 2. We can consider a simple two-criteria problem with objective function $\boldsymbol{f}(w_1, w_2) = (w_1, w_2)$ and $\mathcal{W} = \{\boldsymbol{w} \in \mathbb{R}^2_+ \colon \max(w_1 + \frac{w_2}{2}, \frac{w_1}{2} + w_2) \geq 1\}$. The efficient set of this problem is $\mathcal{W}_E = \{\boldsymbol{w} \in \mathbb{R}^2_+ \colon \max(w_1 + \frac{w_2}{2}, \frac{w_1}{2} + w_2) = 1\}$, yet the weighted sum method can find only two efficient solutions, $\hat{\boldsymbol{w}}^1 = (1, 0)$ and $\hat{\boldsymbol{w}}^2 = (0, 1)$. △

2.2.2 The ε-constraint method

Beyond the weighted sum method, the ε-constraint method, introduced by Haimes et al. [14], is a widely adopted technique for solving multi-objective optimization problems. This method transforms the multi-objective problem (2.1) into a single-objective one, which minimizes a single while constraining the rest. It can be formulated as

$$\begin{aligned} &\min_{\boldsymbol{w} \in \mathcal{W}} f_i(\boldsymbol{w}), \\ &\text{s.t. } f_m(\boldsymbol{w}) \leq \varepsilon_m, \quad m = 1, \dots, M, \quad m \neq i, \end{aligned} \tag{2.4}$$

where $\varepsilon \in \mathbb{R}^M$.

An efficient solution of (2.1) is also an optimal solution of (2.4) for some $\varepsilon \in \mathbb{R}^M$, but more importantly, it can be shown that if $\boldsymbol{w}$ is an optimal solution for some $\varepsilon \in \mathbb{R}^M$ regardless of the choice of the minimized criterion f_m, it is an efficient solution of (2.1). Therefore, every solution of the multi-objective problem can be found using (2.4).

Theorem 9 (Ehrgott [9] (pp. 100))**.** *The feasible solution of multi-objective optimization $\hat{\boldsymbol{w}} \in \mathcal{W}$ is efficient if and only if there exists an $\hat{\varepsilon} \in \mathbb{R}^M$ such that $\hat{\boldsymbol{w}}$ is an optimal solution of* (2.4) *for all $i = 1, \ldots, M$.*

Proof. We will prove both implications of the equivalence by contraposition.

$\Rightarrow$ Let $\varepsilon = \boldsymbol{f}(\hat{\boldsymbol{w}})$. Assume $\hat{\boldsymbol{w}}$ is not an optimal solution of (2.4) for some i. Then there exists $\boldsymbol{w} \in \mathcal{W}$ with $f_i(\boldsymbol{w}) < f_i(\hat{\boldsymbol{w}})$ and $f_m(\boldsymbol{w}) \leq \varepsilon_m = f_m(\hat{\boldsymbol{w}})$ for all $m \neq i$. That is, $\hat{\boldsymbol{w}}$ is not efficient.

$\Leftarrow$ Suppose $\hat{\boldsymbol{w}}$ is not efficient. Then there is an index $i \in \{1, \ldots, M\}$ and feasible solution $\boldsymbol{w} \in \mathcal{W}$ such that $f_i(\boldsymbol{w}) < f_i(\hat{\boldsymbol{w}})$ and $f_m(\boldsymbol{w}) \leq f_m(\hat{\boldsymbol{w}})$ for $m \neq i$. Therefore $\hat{\boldsymbol{w}}$ cannot be an optimal solution of (2.4) for any ε for which it is feasible.

□

This theorem provides a check of efficiency of a feasible solution. However, the proof of the theorem uses $\boldsymbol{f}(\hat{\boldsymbol{w}})$ to set the values of ε, and therefore cannot be directly used to find the whole set of efficient solutions. To achieve that numerically, we will use a grid of ε values between ideal values and the worst possible, so-called nadir values of each objective. That is $\varepsilon_m \in [\min_{\boldsymbol{w}\in\mathcal{W}} f_m(\boldsymbol{w}),\ \max_{\boldsymbol{w}\in\mathcal{W}} f_m(\boldsymbol{w})]$.

Remark. It should be noted that if $\hat{\boldsymbol{w}} \in \mathcal{W}$ is an optimal solution of (2.4) for a single $i \in \{1, \ldots, M\}$, it is not a sufficient condition for the efficiency of $\hat{\boldsymbol{w}}$ as a solution of multi-objective optimization. That is because there might exist $\tilde{\boldsymbol{w}} \in \mathcal{W}$ such that $f_i(\hat{\boldsymbol{w}}) = f_i(\tilde{\boldsymbol{w}})$ and $f_m(\tilde{\boldsymbol{w}}) < f_m(\hat{\boldsymbol{w}}) < \varepsilon_m$ for some $m \in \{1, \ldots, M\},\ m \neq i$, rendering $\hat{\boldsymbol{w}}$ inefficient.

2.2.3 Other methods

Both of the presented methods are so-called scalarization methods, as they transform the problem into a single-objective version to find efficient solutions. Several other techniques exist for solving a multi-objective optimization problem. For example, a mixed method, where we can minimize a weighted sum of convex functions while ε-constraints are used for non-convex functions. Another common option is goal-programming, which finds solutions that minimize the deviation from an ideal solution. For further details and more methods, see Ehrgott [9].

The choice of a specific method should depend on our goals, and there are many other options, often useful if we need properties of our solutions other than the efficiency as we defined it, but for our needs, the two methods described above will be sufficient.

3 Extended mean-risk models

In this chapter, we study mean-risk models that employ only CVaR (conditional value-at-risk) as a risk measure. As Rockafellar and Uryasev [4] note, "In numerical applications, the joint convexity of $f_\alpha(z, L)$, from (1.3), with respect to both z and $\boldsymbol{w}$ in Theorem 4 is even more valuable than the convexity of CVaR in $\boldsymbol{w}$." This follows from the fact that the minimization of CVaR in $\boldsymbol{w} \in \mathcal{W}$ can be reformulated as minimizing f_α jointly in $(z, \boldsymbol{w})$.

Theorem 10 (Rockafellar and Uryasev [4])**.** *Minimizing* $\mathrm{CVaR}_\alpha(L)$ *with respect to* $\boldsymbol{w} \in \mathcal{W}$ *is equivalent to minimizing* $f_\alpha(z, L)$ *from* (1.3) *over all* $(z, \boldsymbol{w}) \in (\mathbb{R} \times \mathcal{W})$, *meaning:*

$$\min_{\boldsymbol{w}\in\mathcal{W}} \mathrm{CVaR}_\alpha(L) = \min_{(z,\boldsymbol{w})\in(\mathbb{R}\times\mathcal{W})} f_\alpha(z, L),$$

and moreover:

$$\begin{aligned} &(z^*, \boldsymbol{w}^*) \in \operatorname*{arg\,min}_{(z,\boldsymbol{w})\in(\mathbb{R}\times\mathcal{W})} f_\alpha(z, -\boldsymbol{w}^T\boldsymbol{r}) \\ \iff\ & \boldsymbol{w}^* \in \operatorname*{arg\,min}_{\boldsymbol{w}\in\mathcal{W}} \mathrm{CVaR}_\alpha(-\boldsymbol{w}^T\boldsymbol{r}), z^* \in \operatorname*{arg\,min}_{z\in\mathbb{R}} f_\alpha(z, -\boldsymbol{w}^{*T}\boldsymbol{r}). \end{aligned}$$

The proof of this theorem, as can be seen in [4], derives from the expression of CVaR_α as $\min f_\alpha$ in (1.3). This theorem is useful if CVaR is in the objective function and is to be minimized. A similar property can be used when CVaR is a constraint.

Theorem 11 (Rockafellar and Uryasev [4])**.** *For probability thresholds* α_i *and loss values* l_i, $i = 1, \dots, M-1$, *the problem:*

$$\begin{aligned} &\min_{\boldsymbol{w}} g(\boldsymbol{w}, \boldsymbol{r}), \\ &\ s.t.\ \mathrm{CVaR}_{\alpha_i}(-\boldsymbol{w}^T\boldsymbol{r}) \leq l_i,\ i = 1, \dots, M-1, \end{aligned} \tag{3.1}$$

where g *is any objective function of loss, is equivalent to the problem:*

$$\begin{aligned} &\min_{z_1,\dots,z_l,\boldsymbol{w}} g(\boldsymbol{w}, \boldsymbol{r}), \\ &\quad s.t.\quad f_{\alpha_i}(z_i, -\boldsymbol{w}^T\boldsymbol{r}) \leq l_i,\ i = 1, \dots, M-1. \end{aligned} \tag{3.2}$$

That is, $(z_1^*, \dots, z_l^*, \boldsymbol{w}^*)$ *solves* (3.2) *if and only if* $\boldsymbol{w}^*$ *solves* (3.1) *and the corresponding constraints are equivalently active.*

Proof. Similarly to the poof of Theorem 10, from (1.3) we have:

$$\mathrm{CVaR}_{\alpha_i}(-\boldsymbol{w}^{*T}\boldsymbol{r}) = \min_{z_i\in\mathbb{R}} f_{\alpha_i}(z_i, -\boldsymbol{w}^{*T}\boldsymbol{r}),$$

and therefore $\mathrm{CVaR}_{\alpha_i}(-\boldsymbol{w}^{*T}\boldsymbol{r}) \leq l_i$ if and only if there exists $z_i \in \mathbb{R}$ such that $f_{\alpha_i}(z_i, -\boldsymbol{w}^{*T}\boldsymbol{r}) \leq l_i$. □

These two theorems present different approaches to mean-CVaR models, either the inclusion of CVaR as an objective function, or as a constraint, or possibly even both, since the objective function in Theorem 11 can also include CVaR.

With a restriction to long-only, fully invested portfolios with budget constraint, and from CVaR convexity shown in Theorem 4, we conclude that all resulting formulations are convex programs, regardless of the use of CVaR as an additional constraint or in the objective function.

Although Rockafellar and Uryasev's original application [4] presented these tools and showed the convexity of the problems, it focused on index replication with one CVaR constraint at confidence level $\alpha = 0.9$. In contrast, this work focuses on return maximization within mean-CVaR framework. We further extend the model by considering CVaR at multiple confidence levels or in combination with other risk measures. We utilize the mean-CVaR model from [4], modified to maximize the expected return, as a baseline against which to compare the performance and set efficient solutions of these extended formulations.

3.1 CVaR constraints approach

We begin by examining the multi-CVaR constraint model derived from Theorem 11, which maximizes the expected return subject to CVaR limits at several confidence levels:

$$\begin{aligned} \min_{\boldsymbol{w}\in\mathbb{R}^K} \quad & -\mu(\boldsymbol{w}), \\ \text{s.t.} \quad & \mathrm{CVaR}_{\alpha_i}(-\boldsymbol{w}^T\boldsymbol{r}) \leq l_i,\ i = 1,\dots,M-1, \\ & w_k \geq 0,\ k = 1,\dots,K, \\ & \sum_{k=1}^{K} w_k = 1, \end{aligned} \tag{3.3}$$

where $\mu(\boldsymbol{w}) = \mathsf{E}[\boldsymbol{w}^T\boldsymbol{r}]$ denotes the expected return of portfolio produced by weights $\boldsymbol{w}$. The simulation in the final chapter will use percentage returns and so l_i must be set accordingly as percentage.

From a theoretical and computational standpoint, this approach is very simple. First, the investor must choose probability thresholds α_i, loss tolerances l_i and set the budget. Then, from Theorem 11 and the convexity of the problem, typically a single solution can be found (in nondegenerate cases where higher risk is associated with higher return portfolios).

As a cost of its simplicity, the CVaR-constraint approach places a burden on the investor to select appropriate pairs (α_i, l_i). If any loss limit l_i is set below the minimal CVaR achievable at level α_i, no feasible portfolio exists. In addition, accurately translating an investor's risk preferences into multiple tail-loss thresholds can be demanding. Lastly, it is necessary to perform the calculation for each individual investor based on their preferences, diminishing the computational advantage of this approach.

From a theoretical point of view, the CVaR-constraint model is simply an instance of the ε-constraint method (2.4). Specifically, if (3.3) has an efficient solution, it is also an efficient solution of (2.4) for $\varepsilon_i = l_i$, $i = 1,\dots,M-1$ and ε_M corresponds to the expected loss left as the objective function. Therefore, some or all CVaR constraints can be converted into the objective function and the solution of (3.3) will lie in the set of efficient solutions of the new problem.

Despite the highlighted disadvantages, if an optimal solution is unique, it is not dominated by any other feasible solution with respect to second-order stochastic dominance (see Chapter 1), and the resulting portfolio merits consideration by any risk-averse investor.

3.2 Extreme-tail CVaR constraint

An alternative method that might be more intuitive for investors while yielding interesting results is to minimize a "baseline" tail risk while imposing a single high-confidence CVaR constraint. For example, minimize the expected loss and CVaR at 90% level, while imposing a CVaR constraint at 99.5%. This method, where the minimization is subject only to a constraint on very rare, severe losses, will be called an extreme-tail CVaR approach.

Formally, we solve:

$$\begin{aligned} \min_{\boldsymbol{w}\in\mathbb{R}^K} \ & \Big(-\mu(\boldsymbol{w}),\ \mathrm{CVaR}_{\alpha_1}(-\boldsymbol{w}^T\boldsymbol{r}) \Big), \\ \text{s.t.}\ & \mathrm{CVaR}_{\alpha_2}(-\boldsymbol{w}^T\boldsymbol{r}) \le l_2, \\ & w_k \ge 0,\ k = 1,\dots,K, \\ & \sum_{k=1}^{K} w_k = 1, \end{aligned} \tag{3.4}$$

where $\alpha_1 < \alpha_2$. Here CVaR_{α_1} enters the multi-objective pair to be minimized and CVaR_{α_2} restricts the extreme tail. By choosing one loss threshold at a very high confidence level, the investor only needs to specify the value of l_2, which can, for example, be "expected worst loss in three years", a concept that should be more approachable compared to the previous model (3.3).

As mentioned before, from convexity of CVaR, (3.4) remains a convex problem. To find the set of all efficient portfolios, we may apply the ε-constraint method, treating CVaR_{α_1} as a constraint $\mathrm{CVaR}_{\alpha_1} \le \varepsilon_1$ and varying ε_1 between its minimum and maximum feasible values. The upper bound $\varepsilon_1^{(u)}$ is equal to CVaR_{α_1} of a single-asset portfolio with the largest CVaR_{α_1}, from subadditivity. Rockafellar and Uryasev [3] show that thanks to the equivalence in Theorem 10, an approximation of the lower bound $\varepsilon_1^{(l)}$ can be found by solving the linear programming problem:

$$\min_{\boldsymbol{w},\boldsymbol{u},z} \left(z + \frac{1}{N(1-\alpha)} \sum_{i=1}^{N} u_i \right), \tag{3.5}$$

subject to the linear constraints on the weights from (3.4), $u_i \ge -\boldsymbol{w}^T\boldsymbol{r}_i - z$ and $u_i \ge 0$ for $i = 1,\dots,N$, where $\boldsymbol{w}$ are weights of a portfolio, $N \in \mathbb{N}$ is the number of observations and $\boldsymbol{r}_1,\dots,\boldsymbol{r}_N$ is the observed sample set of returns.

Maximization of the return for all constraint values $\varepsilon_1 \in (\varepsilon_1^{(l)}, \varepsilon_1^{(u)})$, followed by elimination of non-efficient results yields the set of efficient solutions (as the asset with the largest CVaR may not have the highest return). The resulting set will be closed and connected, as shown in Section 2.1. While this guarantees some stability of the solution, the efficient set may not necessarily be convex. An example of a non-convex efficient set can be constructed using three simple bonds.

Example 3. Let X, Y and Z be bonds with nominal value 1, where X does not show any loss or return (equivalent to not investing the money), Y has a final value 6 or 0, both with probability 0.5 and Z has a final value 3 with probability 0.9, and 0.5 with probability 0.1, Y and Z are independent. We will minimize the expected loss and CVaR_{95} with the constraint $\text{CVaR}_{99}(L) \leq 1$.

Clearly, any combination of the three bonds is feasible. Furthermore, both a single bond portfolio consisting of only X is efficient as any other feasible solution has $\text{CVaR}_{95}(L) > 0$ and a single bond portfolio consisting of only Y is efficient as it is the only feasible portfolio with expected loss less than or equal to -2.

If the efficient set were convex, then the portfolio $\frac{1}{2}X + \frac{1}{2}Y$ with the expected loss -1 and $\text{CVaR}_{95}(L_{\frac{1}{2}X+\frac{1}{2}Y}) = 0.5$ would also be efficient. However, the bond Z has an expected loss $\mathsf{E}(L_Z) = -1.75$ and $\text{CVaR}_{95}(L_Z) = 0.5$, and therefore $\frac{1}{2}X + \frac{1}{2}Y$ is not an efficient portfolio, as it is dominated by Z. △

Compared to the single-level mean–CVaR benchmark, with the assumption of a reasonable extreme-tail threshold, the shape of the efficient frontier under an extreme-tail constraint is similar, but with some high-risk portfolios excluded. Empirically, it will be of interest to observe how much in-sample expected return is sacrificed to satisfy the extreme-tail bound, and whether the difference persists out of sample.

3.3 Efficient mean–CVaR frontier

The most general formulation of the mean–CVaR optimization problem includes all relevant CVaR levels directly in the objective function, without requiring any risk thresholds to be specified in advance. In the case of three objectives, expected return and two CVaR levels, the following model can be used:

$$\begin{aligned}
\min_{\boldsymbol{w}\in\mathbb{R}^K} \; & \Big(-\mu(\boldsymbol{w}),\ \text{CVaR}_{\alpha_1}(-\boldsymbol{w}^T\boldsymbol{r}),\ \text{CVaR}_{\alpha_2}(-\boldsymbol{w}^T\boldsymbol{r})\Big), \\
\text{s.t. } & w_k \geq 0,\ k = 1, \ldots, K, \\
& \sum_{k=1}^{K} w_k = 1.
\end{aligned} \tag{3.6}$$

This formulation treats all risk measures as objectives to be minimized alongside the (negative) expected return. It generalizes the single-level mean–CVaR model by expanding the notion of efficiency across multiple confidence levels.

The addition of CVaR levels to the objective expands the set of efficient portfolios. As more CVaR terms are introduced, the efficient set approximates the set of all non-dominated portfolios under second-order stochastic dominance (SSD). In the limit, this set is equivalent to SSD-efficient set if we consider all CVaR levels.

As in earlier formulations, the set of efficient portfolios $\mathcal{W}_E$ is closed and connected, but not necessarily convex. This non-convexity can be seen by extending Example 3, if a new CVaR level is added, it may not affect the efficiency of the mentioned portfolios, preserving the same non-convex structure. However, from the properties of $\mathcal{W}$ and the mapping $\boldsymbol{f}(\cdot)$, the image $\boldsymbol{f}(\mathcal{W})$ is a closed set, and the image of the efficient set $\boldsymbol{f}(\mathcal{W}_E)$ lies on its boundary.

More importantly, from Section 2.1, an image of any efficient point must also lie on the boundary of $\boldsymbol{f}(\mathcal{W}) + \mathbb{R}^K_{\geq}$, a convex set. Therefore, $\boldsymbol{f}(\mathcal{W}_E)$ is a connected, closed subset of a boundary of a convex set.

In most cases, $\boldsymbol{f}(\mathcal{W}_E)$ forms a "two-dimensional surface" of the set of feasible solutions $\boldsymbol{f}(\mathcal{W})$. However, it may partially or fully collapse into a curve if the same portfolio simultaneously minimizes multiple CVaR levels for a given return. This can occur when CVaR_{α_1} and CVaR_{α_2} are minimized by the same allocation, which is plausible, because if $\alpha_1 < \alpha_2$, then $\mathrm{CVaR}_{\alpha_1} \leq \mathrm{CVaR}_{\alpha_2}$ and the minimization of one could lead to the minimization of the other. When such alignment arises, it may be considered as a desirable property, rather than a limitation, assuming that α_1 and α_2 were set appropriately. For a given return, we obtain a portfolio that is optimal with respect to multiple risk measures.

This collapse of $\boldsymbol{f}(\mathcal{W}_E)$ has a natural analogue in the single-level CVaR level, where the efficient set may degenerate to a single point. It occurs only if it is impossible to create a portfolio that has a lower CVaR than a single asset with the highest return, reducing the efficient set to only one solution, but that scenario is significantly more arbitrary.

3.4 Spectral risk measure reformulation

The model (3.6), which treats multiple CVaR levels as separate objectives, can be reformulated by aggregation of CVaR, forming a new risk measure, called the spectral risk measure. Spectral risk measures are based on quantile functions of portfolio returns, this section will consider the return random variable $X \in \mathcal{X}$, rather than loss L_X. We modify the risk metrics for continuous returns:

$$\begin{aligned}\mathrm{VaR}_\alpha^{ret}(X) &= -\min(x \in \mathbb{R}\colon\ \mathsf{P}[X \leq x] > 1-\alpha)\\ &= -\mathrm{VaR}_{1-\alpha}(X) = \mathrm{VaR}_\alpha(-X) = \mathrm{VaR}_\alpha(L_X),\end{aligned}$$

and similarly, from the alternative expression of CVaR (1.4), which is equivalent under continuity:

$$\begin{aligned}\mathrm{CVaR}_\alpha^{ret}(X) &= -\frac{1}{1-\alpha}\int_0^{1-\alpha} \mathrm{VaR}_t(X)dt \qquad \Big|_{s=1-t,\ ds=-dt}\\ &= -\frac{1}{1-\alpha}\left(-\int_\alpha^1 \mathrm{VaR}_s(L_X)ds\right) = \mathrm{CVaR}_\alpha(L_X).\end{aligned}$$

A spectral risk measure requires risk spectra, weight functions on $[0,1]$. We denote the set of all integrable functions on $[0,1]$:

$$\mathcal{L}^1[0,1] = \left\{f\colon [0,1] \to \mathbb{R}\colon\ \int_0^1 |f(x)|dx < \infty\right\}.$$

Definition 16 (Kim [15]). *An element $\phi \in \mathcal{L}^1[0,1]$ is called an* ***admissible risk spectrum*** *if it satisfies*

1. $\phi \geq 0$,
2. *ϕ is non-increasing,*
3. $\|\phi\|_{\mathcal{L}^1} = \int_0^1 |\phi(t)|dt = 1$.

An admissible risk spectrum is also called risk aversion function, because it can be interpreted as assigning weights to quantiles, or possible scenarios, with worse scenarios having higher weights. This is very similar to the utility function of a risk-averse investor in Chapter 1. According to Acerbi [16], any rational investor can express their subjective risk aversion by ϕ, so for any concave utility function there exists a corresponding risk aversion function.

Definition 17 (Kim [15])**.** *For an admissible risk spectrum $\phi \in \mathcal{L}^1[0,1]$ and the upper quantile function $q_X^+(t) = \inf\{x \in \mathbb{R} \colon P[X \le x] > t\}$, the **spectral risk measure** M_ϕ generated by ϕ is defined as:*

$$M_\phi(X) = -\int_0^1 q_X^+(t)\phi(t)dt.$$

The upper quantile function $q_X^+(t)$ corresponds to the upper VaR from Definition 5. Additionally, for continuous distribution of returns, $q_X^+(t)$ also coincides with VaR and CVaR_α is a spectral risk measure generated by the risk spectrum $\phi_\alpha(t) = \frac{1}{1-\alpha}1_{[0,1-\alpha]}(t)$. This spectrum gives equal weight across the tail and zero elsewhere, leading to mean of the tail similarly to CVaR. For continuous returns, we have $q_X^+(t) = \mathrm{VaR}_t(X)$, and therefore:

$$\begin{aligned} M_{\phi_\alpha}(X) &= -\int_0^1 q_X^+(t)\phi_\alpha(t)dt = -\frac{1}{1-\alpha}\int_0^{1-\alpha} q_X^+(t)dt \\ &= -\frac{1}{1-\alpha}\int_0^{1-\alpha} \mathrm{VaR}_t(X)dt = \mathrm{CVaR}_\alpha^{ret}(X) = \mathrm{CVaR}_\alpha(L_X). \end{aligned}$$

The model (3.6) treats CVaR_{α_1} and CVaR_{α_2} as separate objectives, however it is a convex model and according to Theorem 8, every efficient solution can be found by minimizing the weighted sum:

$$\min_{\boldsymbol{w}\in\mathcal{W}}\Big(-\lambda_1\mu(\boldsymbol{w}) + \lambda_2\,\mathrm{CVaR}_{\alpha_1}(-\boldsymbol{w}^T\boldsymbol{r}) + \lambda_3\,\mathrm{CVaR}_{\alpha_2}(-\boldsymbol{w}^T\boldsymbol{r})\Big),$$

for some $\boldsymbol{\lambda} \in \mathbb{R}_+^3$.

Furthermore, we get an equivalent set of efficient solutions if we limit $\boldsymbol{\lambda}$ such that $\lambda_1 + \lambda_2 + \lambda_3 = 1$ and after appropriate change of notation, we get:

$$\min_{\boldsymbol{w}\in\mathcal{W}}\Big(\lambda(-\mu(\boldsymbol{w})) + (1-\lambda)(\beta_1\,\mathrm{CVaR}_{\alpha_1}^{ret}(\boldsymbol{w}^T\boldsymbol{r}) + \beta_2\,\mathrm{CVaR}_{\alpha_2}^{ret}(\boldsymbol{w}^T\boldsymbol{r}))\Big),$$

with corresponding $\lambda = \lambda_1 \in [0,1]$, $\beta_1 = \frac{\lambda_2}{1-\lambda_1}$, $\beta_2 = \frac{\lambda_3}{1-\lambda_1}$ and from $\lambda_2+\lambda_3 = 1-\lambda_1$, we have $\beta_1 + \beta_2 = \frac{\lambda_2+\lambda_3}{1-\lambda_1} = 1$.

In other words, CVaR_{α_1} and CVaR_{α_2} can be aggregated into a single coherent spectral risk measure generated by the risk spectrum:

$$\phi_{\alpha_{1,2}}(t) = \beta_1\tfrac{1}{1-\alpha_1}\mathbf{1}_{[0,1-\alpha_1]}(t) + \beta_2\tfrac{1}{1-\alpha_2}\mathbf{1}_{[0,1-\alpha_2]}(t).$$

The formula of the risk measure then is as follows:

$$\begin{aligned} M_{\phi_{\alpha_{1,2}}}(X,\beta_1,\beta_2) &= \beta_1\,\mathrm{CVaR}_{\alpha_1}^{ret}(X) + \beta_2\,\mathrm{CVaR}_{\alpha_2}^{ret}(X) \\ &= \beta_1\,\mathrm{CVaR}_{\alpha_1}(L_X) + \beta_2\,\mathrm{CVaR}_{\alpha_2}(L_X). \end{aligned}$$

Consequently, the multi-objective problem (3.6) can be reformulated as a scalar mean–spectral risk optimization:

$$\min_{\boldsymbol{w}\in\mathcal{W}} \Big(-\lambda\,\mu(\boldsymbol{w}) + (1-\lambda)\,M_{\phi_{\alpha_{1,2}}}(\boldsymbol{w}^T\boldsymbol{r},\beta_1,\beta_2)\Big), \tag{3.7}$$

where every efficient solution can be found for some $\lambda \in [0,1]$ and $\beta_1, \beta_2 \geq 0$, $\beta_1 + \beta_2 = 1$.

There are useful properties that hold for all spectral risk measures, for example, law-invariance and coherency (see Kim [15]), but these can be shown for our model without spectral measures. The main benefit of this reformulation is conceptual. It incorporates our model into the wider and more flexible class of coherent risk measures. As stated by Acerbi [16], the spectral risk measure approach shows us a more general view of coherent measures, specifically "a measure is coherent if it assigns bigger weights to worse cases". This describes a very wide class of risk measures, of which a significant part can be achieved using the spectral risk framework.

Although the original model remains highly useful, especially for investors who focus on limiting worst-case losses, it is limited compared to the possibilities of coherent risk measures. There might be investors with reasonable risk approach who are not satisfied by our model, since it does not distinguish between portfolios that differ in their distribution of high returns but not losses. Spectral risk measures present a natural extension for these investors, as they can capture differences across the whole return distribution.

4 Incorporation of additional criteria

The previous chapter examined how adding additional CVaR levels impacts the set of efficient portfolios. Another meaningful extension is to incorporate different types of criteria in an attempt to increase realized return, such as liquidity, or to introduce some cost of the transactions into the model. In this chapter, we explore the addition of transaction costs as an objective, measuring these costs using the L^1 distance from a baseline portfolio.

Definition 18. *Given a starting portfolio defined by the decision vector $\boldsymbol{w}_0 \in \mathcal{W}$ and a per-unit cost $C \in \mathbb{R},\ C > 0$, the* ***transaction cost*** *of acquiring a portfolio with the decision vector $\boldsymbol{w} \in \mathcal{W}$ is defined as $C\|\boldsymbol{w} - \boldsymbol{w}_0\|_1 = \sum_{i=1}^{K} C|w_i - w_{0,i}|$, where $\|\cdot\|_1$ denotes the L^1 norm.*

This definition assumes that the cost of a portfolio change depends only on its size, not specific assets. Since minimization does not depend on the value of C, and we do not have a specific cost, we may omit C from the definition, treating the pure L^1 distance as the transaction cost. The model must assume that the investor already holds some position in the assets, the starting portfolio, otherwise the cost would be equivalent for all feasible decisions.

In contrast to the pure mean-CVaR models, where every criterion has the same unit, the additional measure might significantly differ in scale, which could affect the scalarization method. Transaction cost might not strictly require normalization, depending on the values of expected return and CVaR. However, it is advisable to at least check its effect.

This caution is even more important for some other metrics like liquidity, which might be measured in seconds or days required to liquidate the entire portfolio. Their scale could differ entirely depending on the units used to represent them, making normalization potentially necessary.

4.1 Mean-risk model with transaction cost

Introducing transaction cost as a separate objective might not seem very intuitive. If possible, the use of expected returns that include the cost of acquiring the portfolio might seem more effective. However, that approach presumes a fixed holding period to amortize the cost, which might not be preferred by the investor. By treating transaction cost as its own objective, we make the trade-off explicit and perhaps allow the investor to decide how long to hold the portfolio based on the results.

Since the L^1 norm is convex, and given a starting portfolio $\boldsymbol{w}_0 \in \mathcal{W}$, we can formulate a convex programming problem:

$$\begin{aligned} \min_{\boldsymbol{w} \in \mathbb{R}^K} & \left(-\mu(\boldsymbol{w}),\ \mathrm{CVaR}_\alpha(-\boldsymbol{w}^T\boldsymbol{r}),\ \|\boldsymbol{w} - \boldsymbol{w}_0\|_1\right), \\ \text{s.t. } & w_k \geq 0,\ k = 1, \dots, K, \\ & \sum_{k=1}^{K} w_k = 1. \end{aligned} \tag{4.1}$$

As mentioned above, we might consider normalization of the criteria before the calculation. Specifically, we perform min-max standardization

$$\widehat{f}_i(\boldsymbol{w}) = \frac{f_i(\boldsymbol{w}) - f_i(\tilde{\boldsymbol{w}})}{\max_{\boldsymbol{w}\in\mathcal{W}} f_i(\boldsymbol{w}) - f_i(\tilde{\boldsymbol{w}})},$$

where $\tilde{\boldsymbol{w}}$ is the ideal solution from Definition 11. Since the negative expected return is linear, the highest and lowest expected returns are given by single assets. For CVaR, it was already shown how to obtain its minimum and maximum as $\varepsilon_1^{(l)}$ and $\varepsilon_1^{(u)}$ in Section 3.2.

Lastly, the transaction cost always obtains values between 0 and 2. In practice, a portfolio with no transaction cost is always obtainable (the starting portfolio). On the other hand, the maximal cost might be less than 2, but never less than $2 - \frac{2}{K}$, i.e., the minimal cost of the portfolio obtained by selling every asset except for the one where the starting position is the smallest. In practice, with sufficiently large K, we can simply divide the cost by 2 to normalize it.

Convexity and an identical set of feasible solutions guarantee that the efficient set $\mathcal{W}_E$ remains closed and connected, equivalently to Section 3.3. However, the efficient sets might differ significantly in other aspects. The collapse of $\boldsymbol{f}(\mathcal{W}_E)$ described in Section 3.3 depends on the starting portfolio. There is no apparent reason why the distance from an arbitrary portfolio and CVaR should be positively correlated, but if the starting portfolio was obtained by previous CVaR optimization, the situation can be very different.

The shape of the efficient set will be significantly affected by the specific starting portfolio. The efficient set of (4.1) can include more portfolios compared to (3.6), because it is no longer guaranteed that $\mathcal{W}_E$ is a subset of all non-dominated portfolios under second-order stochastic dominance (SSD). For example, the starting portfolio is always efficient and can be dominated with respect to SSD.

While unusual from a pure risk-return perspective, an efficient solution that is dominated with respect to SSD can realistically reflect investor preferences. A scenario can exist in which an otherwise optimal portfolio is unattractive to the investor because of the cost related to its acquisition. This phenomenon shows an imperfection of the theory when translated to the real world, not a deficiency of this approach.

Because the model (4.1) is convex, every efficient solution can be found using both the weighted sum and ε-constraint method from Chapter 2, or even their combination:

$$\begin{aligned} \min_{\boldsymbol{w}\in\mathbb{R}^K} \ & \Big(\lambda(-\mu(\boldsymbol{w})) + (1-\lambda)\,\mathrm{CVaR}_\alpha(-\boldsymbol{w}^T\boldsymbol{r})\Big), \\ \text{s.t.} \ & \|\boldsymbol{w} - \boldsymbol{w}_0\|_1 < \varepsilon, \\ & w_k \geq 0, \quad k = 1, \dots, K, \\ & \sum_{k=1}^{K} w_k = 1, \end{aligned} \tag{4.2}$$

for some $\lambda \in [0,1]$ and $\varepsilon \in [0,2]$. Theorems 8 and 9 ensure that every point on the efficient set of the original problem can be reached.

The aggregation of CVaR and mean as a single mean-risk criterion, with transaction cost as a constraint, achieves a formulation that aligns with traditional mean-risk theory. We get a mean-risk model with a limited set of feasible portfolios.

The modified model would allow an investor to set the transaction cost constraint in advance, instead of finding efficient portfolios for every transaction cost. This scenario would correspond to an investor who already holds a portfolio and wants an adjustment to acquire "the best" possible portfolio of the same value, under some budget constraint. That situation is possible, but the author believes that most investors would prefer to choose the price after the calculation, together with other parameters. Determining the cost constraint in advance could lead to overlooking a slightly more expensive portfolio with disproportionately better expected performance or significantly cheaper portfolio with expected performance similar to an optimal solution.

In general, the transaction cost of the portfolio becomes more important when portfolio adjustments are made often. The trading scenario, where an investor adjusts their portfolio regularly (e.g., every week), could be optimized better using models that focus on the transaction cost to expected return difference ratio, but that approach would require the knowledge of specific transaction costs and is beyond the scope of this thesis. More importantly, the methods presented in this chapter are not unique to transaction cost optimization and would be valid for any convex objective function.

5 Numerical study

In this chapter, we evaluate portfolios generated by the models introduced in Chapters 3 and 4 using real-world S&P 500 data. We assess practical viability of each model, explore its efficient set and corresponding objective space image, and compare these results against a benchmark model.

Subsequently, we assess out-of-sample performance of selected portfolios, analyzing their returns relative to expectations, comparing their performance with the benchmark, and identifying possible advantages or disadvantages of these models.

To find efficient portfolios, we use the methods described in Section 2.2. The CVaR minimization uses an approximation of the optimization function $f_\alpha(z, L)$, provided by Rockafellar and Uryasev [3]:

$$\widetilde{f}_\alpha(z, \boldsymbol{w}) = z + \tfrac{1}{N(1-\alpha)} \sum_{i=1}^{N} \Big(\max(-\boldsymbol{w}^T \boldsymbol{r}_i - z,\ 0) \Big).$$

where $N \in \mathbb{N}$ is the number of samples and $\boldsymbol{r}_i$ are the observed returns. According to Rockafellar and Uryasev [3], "the minimization of $\widetilde{f}_\alpha$ over $\mathbb{R} \times \mathcal{W}$, in order to get an approximate solution to the minimization of $f_\alpha(z, L)$, can be reduced to convex programming." It is equivalent to the minimization problem (3.5) presented in Section 3.2. Optimization is implemented using *CVXPY* [17], a domain-specific language for convex optimization embedded in Python.

When reporting a specific CVaR_α value of a portfolio, we use the average of observed losses exceeding the estimated quantile $\widehat{\text{VaR}}_\alpha :\ = L_{(\lceil \alpha N \rceil)}$, where $\lceil \cdot \rceil$ is the ceiling function and $L_{(1)} \leq L_{(2)} \leq \ldots \leq L_{(N)}$ are sorted loss data.

The thesis includes attached Python code that performs the presented study.

5.1 Data

We selected 10 large-cap S&P500 stocks ($K = 10$), with a maximum of two per sector. The companies are listed in Table 5.1.

Company Name	**Ticker**	**Sector**
Apple Inc.	AAPL	Information Technology
Amazon.com, Inc.	AMZN	Consumer Discretionary
Bank of America Corp.	BAC	Financials
Costco Wholesale Corp.	COST	Consumer Staples
International Business Machines Corp.	IBM	Information Technology
Johnson & Johnson	JNJ	Health Care
JPMorgan Chase & Co.	JPM	Financials
The Coca-Cola Company	KO	Consumer Staples
AT&T Inc.	T	Communication Services
UnitedHealth Group Inc.	UNH	Health Care

Table 5.1 Selected S&P 500 companies.

We utilized the *yfinance* library [18] to retrieve historical daily adjusted closing prices, which account for dividends, stock splits, and other corporate actions affecting stock prices. The data span from January 1, 2015, to December 10, 2024, containing 2501 trading days. Daily returns are calculated as follows:

$$R_t = \frac{P_t}{P_{t-1}} - 1, \; t = 2, 3, \ldots$$

We will present the results in percent units ($R_t \cdot 100\%$), both for the expected return and CVaR.

This enables us to compute daily returns for 2500 trading days. Of these, 2000 trading days (from January 1, 2015 to December 12, 2022) serve as in-sample data, and the following 500 trading days (from December 13, 2022 to December 10, 2024) are out-of-sample data. This extensive dataset is necessary for the employment of high-level CVaR, such as 99.5%.

We verify data integrity and find no missing values or anomalies. An analysis of individual stock price histories and return distributions does not reveal any irregularities, allowing us to proceed with the assumption that the dataset is complete and without any significant flaws.

Table 5.2 presents summary statistics of daily returns (mean, min, max of each stock, comparison of variances and CVaR_{90} values). While CVaR and variance correlate as expected, the portfolios produced using the mean-CVaR models will definitely differ from the traditional Markowitz portfolios. For instance, *UnitedHealth* (UNH) has lower CVaR_{90} than *IBM*, yet exhibits higher variance.

Return statistics in Table 5.2 are expressed as daily percentages. For example, *Apple* (AAPL) has a mean daily return of 0.106% (in-sample), which compounds to approximately 30.6% annually. Its worst single day loss is 12.865%, and on its best day it increased by 11.981%.

A $CVaR_{90}$ value of 3.359% for *Apple* means that, considering only the worst 10% of trading days, the stock loses on average about 3.4% of its value on those days. However, this does not imply that, in any given ten-day period, it should be expected that the worst day results in 3.4% loss. Financial time series typically exhibit volatility clustering (Cipra [19, pp. 202]), which means that days with high volatility tend to cluster together. As a result, some ten-day periods may contain multiple "bad days" with large losses, while others may have none.

Examining the performance of individual assets in Figure 5.1, all stocks delivered positive overall gains. Most stocks would have yielded a profit if sold at almost any earlier point, except for *IBM*. This suggests that long-only portfolios should not outright exclude any asset from consideration.

Among the individual assets, *Amazon* (AMZN) exhibits the highest total return, but also had the highest CVaR_{90}. Its advantage is notably smaller when considering only in-sample data. Both *Amazon* and *Apple* could be less represented in our portfolios than *UnitedHealth*, which offers a similar expected return at significantly lower CVaR_{90}, in terms of in-sample data.

Costco (COST) is another stock that is likely to be highly represented in our portfolios. It provides one of the highest expected returns, while its variance and CVaR are the third lowest, outperformed only by *Johnson & Johnson* (JNJ) and *Coca-Cola* (KO), both of which offer significantly lower returns.

Lastly, we examine the correlation between the assets considered in Figure 5.3. *Amazon* stands out again, with the lowest overall correlation values, exhibiting a higher correlation only with *Apple* and *Costco*, both companies operating in similar sector. This disconnection of *Amazon* from the rest of the market is further evident in Figure 5.2, where, during the COVID-19 pandemic, every other company experienced its highest variance period, while *Amazon* appears to be significantly less affected.

We observe the highest correlation between *JPMorgan Chase* (JPM) and *Bank of America* (BAC), two companies in the financial sector. Their rolling average variance in Figure 5.2 shows a similar pattern, suggesting that their variance is more influenced by external factors than by the companies themselves. Since *JPMorgan Chase* delivers better returns with less risk, and considering the high correlation of returns, efficient portfolios are likely to favor *JPMorgan Chase* over *Bank of America*. However, this is an isolated occurrence and does not pose a significant issue.

Ticker	**Mean**	**Min**	**Max**	**Variance**	$\mathbf{CVaR_{90}}$
			In-sample data		
AAPL	0.106	-12.865	11.981	3.551	3.359
AMZN	0.110	-14.049	14.131	4.395	3.653
BAC	0.059	-15.397	17.796	4.214	3.564
COST	0.080	-12.451	9.959	1.953	2.451
IBM	0.028	-12.851	11.301	2.415	2.809
JNJ	0.044	-10.038	7.998	1.374	2.056
JPM	0.065	-14.965	18.012	3.244	3.068
KO	0.041	-9.672	6.480	1.403	2.135
T	0.027	-9.241	10.022	1.991	2.514
UNH	0.105	-17.277	12.799	2.812	2.776
			Out-of-sample data		
AAPL	0.119	-4.817	7.265	1.936	2.409
AMZN	0.202	-8.785	8.269	3.781	3.184
BAC	0.091	-6.204	8.429	2.441	2.572
COST	0.156	-7.641	7.259	1.449	1.985
IBM	0.111	-8.251	9.487	1.613	2.099
JNJ	-0.017	-3.983	6.073	0.980	1.765
JPM	0.140	-6.468	11.545	1.931	2.353
KO	0.011	-4.833	2.885	0.681	1.473
T	0.075	-10.406	8.476	2.291	2.539
UNH	0.023	-8.112	7.241	2.217	2.685

Table 5.2 Summary statistics of daily stock returns (in percent), $CVaR_{90}$ of losses.

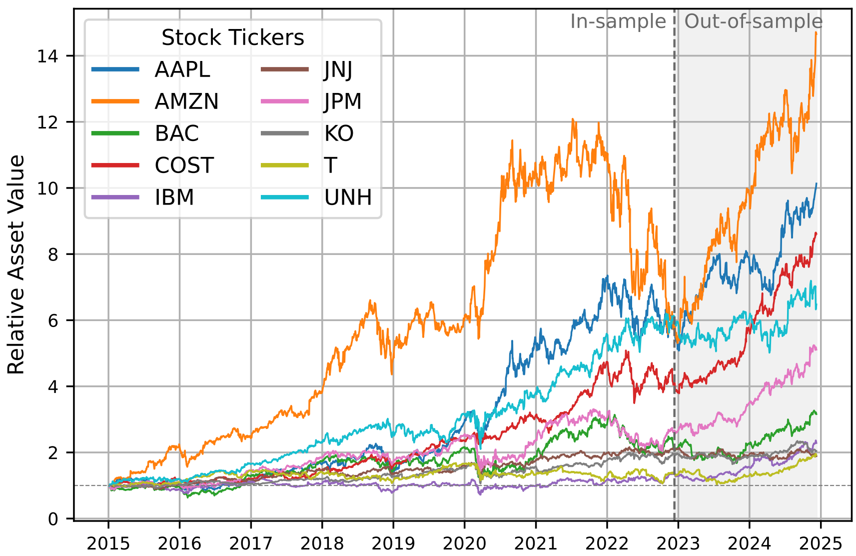


Figure 5.1 Relative price of the individual assets compared to the starting date.

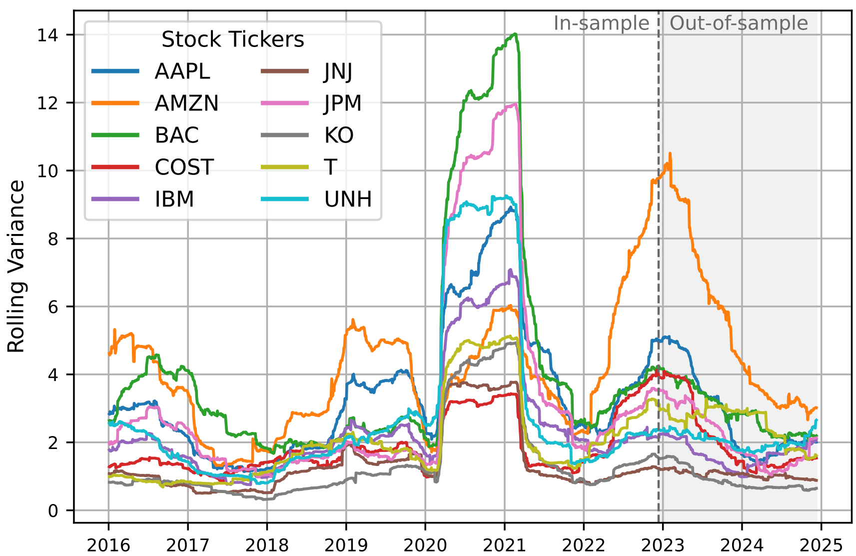


Figure 5.2 Sample variances of individual assets (rolling average over past year).

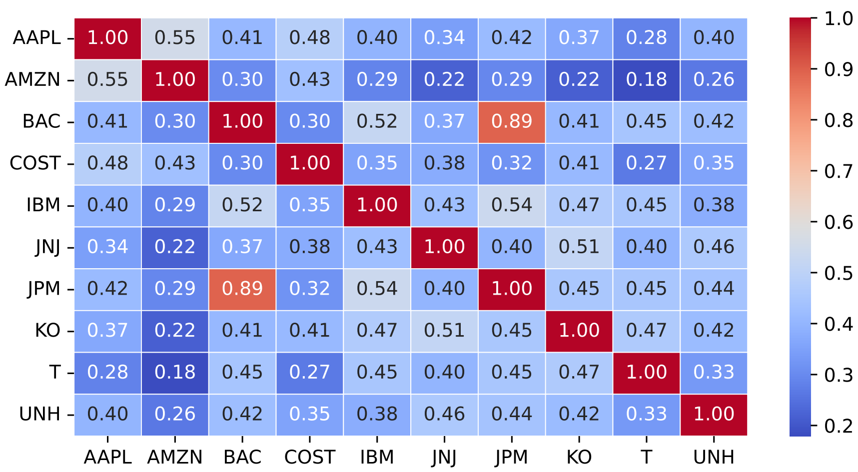


Figure 5.3 Sample correlation matrix of daily stock returns.

5.2 Benchmark model

As described in Chapter 3, we adopt the mean-CVaR model from Rockafellar and Uryasev [4], modified to maximize the expected return, as a baseline against which we compare the performance of the studied models:

$$\begin{aligned} \min_{\boldsymbol{w}\in\mathbb{R}^{10}} & \Big(-\mu(\boldsymbol{w}),\ \mathrm{CVaR}_{90}(-\boldsymbol{w}^T\boldsymbol{r})\Big), \\ \text{s.t. } & w_j \geq 0,\ j = 1, \ldots, 10, \\ & \sum_{k=1}^{10} w_k = 1. \end{aligned} \tag{M1}$$

Within the feasible set of this model, the minimum achievable expected return 0.027% corresponds to *AT&T* (T), while the maximum is from *Amazon* (0.110%).

The entire feasible set is visualized in Figure 5.4. We computed the minimum CVaR_{90} values for all feasible expected return levels using the methods described in Chapter 2, and the highest CVaR_{90} values using Theorem A2. Since CVaR is continuous and the expected return of any convex combination of two portfolios with equivalent return remains the same, any point between the two lines must also be feasible.

The resulting efficient portfolios provide the expected daily returns ranging from 0.051% to 0.110%. The portfolio weights that achieve these returns are included in Table 5.3. The weights of all efficient portfolios are shown in Figure 5.5. Notably, both *JPMorgan Chase* and *Bank of America* are excluded by the model from all efficient portfolios due to their high CVaR_{90}. Other high CVaR stocks, *Apple* and *Amazon* play roles in higher-return portfolios due to their strong return values.

Figure 5.5 further shows that as the expected return and risk increases, efficient portfolios contain fewer assets. Portfolios with expected daily returns greater than 0, 100% are composed of only four assets with the highest return. This is consistent with the notion that diversification lowers CVaR, but also limits maximum return potential.

We selected three portfolios from the efficient set, with low, medium and high risk, for further comparison with other methods. Their portfolio weights are provided in Table 5.3, along with their performance. We intentionally avoided extreme endpoints that maximize the expected return or minimize CVaR overall, as modest deviations from both ends of the efficient set provide better risk-return trade-offs, as seen in Figure 5.4.

Lastly, Table 5.3 contains the L^1 distance of each portfolio from the mean–variance (Markowitz) portfolio that offers the same expected return. Except for the highest returns, the distance ranges from 0.172 to 0.219. This shows that while both models provide qualitatively similar portfolios, substantial allocation differences remain.

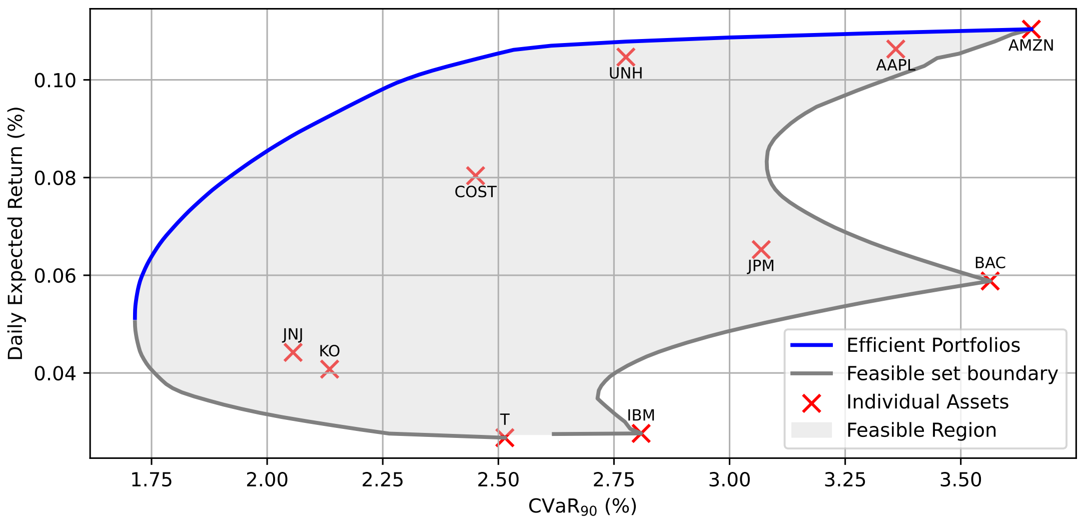


Figure 5.4 Expected daily return and CVaR_{90} of the feasible and efficient set of the benchmark model. Artifacts in feasible set boundary are present due to computational instability, but overall feasible set depicted is precise.

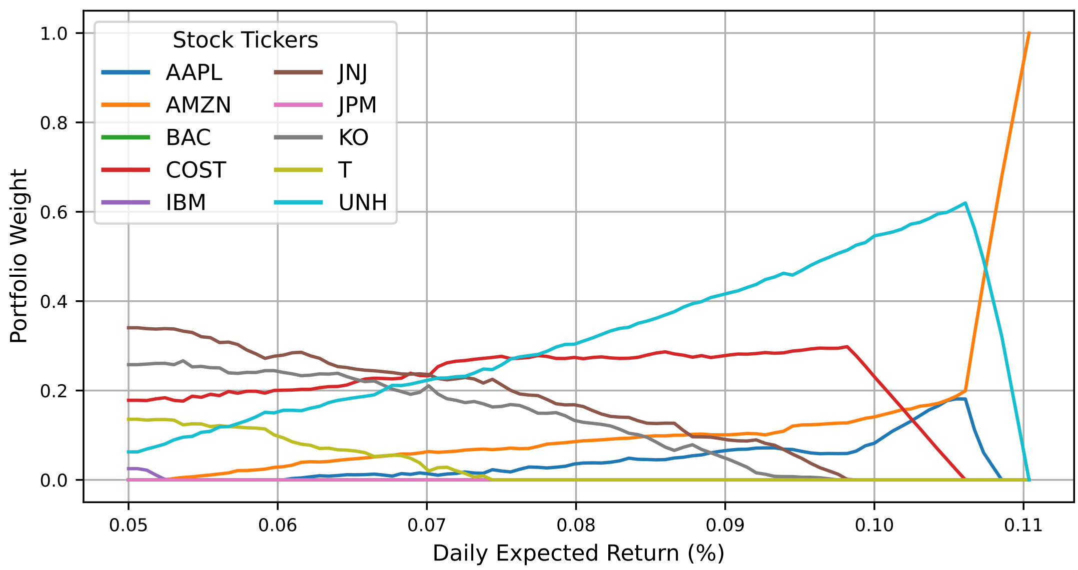


Figure 5.5 Weights of efficient portfolio with targeted expected return.

Risk tolerance	Min	Low	Medium	High	Max
$\mu(\boldsymbol{w})$ (%)	0.051	0.060	0.085	0.105	0.110
CVaR_{90} (%)	1.714	1.730	1.993	2.482	3.653
AAPL	–	–	0.044	0.179	–
AMZN	–	0.029	0.098	0.180	1.000
BAC	–	–	–	–	–
COST	0.178	0.200	0.281	0.041	–
IBM	0.025	–	–	–	–
JNJ	0.340	0.279	0.122	–	–
JPM	–	–	–	–	–
KO	0.258	0.245	0.095	–	–
T	0.136	0.096	–	–	–
UNH	0.063	0.151	0.359	0.600	–
L^1 dist	0.172	0.149	0.171	0.219	0.000

Table 5.3 Performance, weights and L^1 distance from Markowitz portfolio with equivalent expected return for selected portfolios produce by the benchmark model (M1).

5.3 CVaR constraints model

The first model introduced in the thesis was the CVaR constraints model (3.3), where the investor must pre-select the appropriate pairs of CVaR levels α_i and the corresponding loss tolerances l_i. As discussed in Section 3.1, this causes a practical issue, as selecting (α_i, l_i) pairs can be difficult.

This difficulty is evident in a two CVaR level model. Based on the results in Table 5.3, requiring CVaR_{90} to be less than 2% could be perceived as a reasonable choice. Furthermore, the investor might also want to limit CVaR_{95}. The minimum possible CVaR_{95} is 2.305%, and any lower CVaR_{95} threshold makes the model infeasible. Conversely, the portfolio that maximizes the return for $\text{CVaR}_{90} \leq 2\%$ exhibits $\text{CVaR}_{95} = 2.668\%$. Therefore, setting the loss tolerance for CVaR_{95} higher than 2.668% will return a portfolio that can already be achieved by the benchmark model (M1).

It could be suggested that widening the gap between CVaR levels would improve flexibility. However, if we use CVaR_{99} instead of CVaR_{95}, the resulting range is 4.342% to 4.826%, which is not a significant improvement.

This does not restrict the model's ability to provide a reasonable portfolio for an investor with a specific risk preference (if feasible). The issue is that any resulting portfolio can also be produced using other presented methods, which provide the investor with a more informed choice.

5.4 Extreme-tail CVaR model

An alternative approach, still based on a loss tolerance specified in advance, is the extreme-tail CVaR constraint model (3.4). This model places a constraint on CVaR_{α_2}, at a very high quantile that captures extreme losses. Consequently, it requires a long history of returns. For our choice $\alpha_2 = 99.5\%$, constructing a portfolio with fewer than 200 data points results in a $\text{CVaR}_{99.5}$ estimate that is not only imprecise (it relies on a single, highest loss value), but also systematically underestimated. This underestimation arises because, with fewer than 200 data points, the highest observed loss is less likely to underrepresent the true 99.5 percentile.

The specific model we analyze is defined as follows:

$$\begin{aligned} \min_{\boldsymbol{w}\in\mathbb{R}^{10}} \ & \Big(-\mu(\boldsymbol{w}), \text{CVaR}_{90}(-\boldsymbol{w}^T\boldsymbol{r})\Big), \\ \text{s.t.} \ & \text{CVaR}_{99.5}(-\boldsymbol{w}^T\boldsymbol{r}) \leq l, \\ & w_j \geq 0, \ j = 1, \dots, 10, \\ & \sum_{k=1}^{10} w_k = 1. \end{aligned} \tag{M2}$$

Here, l represents the investor's loss tolerance for the worst expected loss over 200 trading days. In our dataset, the lowest achievable $\text{CVaR}_{99.5}$ is 5.158% and the highest return portfolio has a $\text{CVaR}_{99.5}$ of 8.215%. Therefore, any choice of l between these boundaries is available to the investor. If l is less than 5.158%, there is no feasible solution and if it is greater than 8.215%, the new model (M2) is equivalent to the benchmark model (M1).

In Figure 5.6 compares portfolios from the benchmark model (M1) and the extreme-tail models (M2) for multiple choices of loss tolerance l. As expected, the CVaR constraint limits the maximum expected return achievable. Since the benchmark portfolio with the lowest risk has a $\text{CVaR}_{99.5}$ of 5.471%, any stricter constraint means that there will be no common portfolio between the new model and the benchmark. In that case, all new portfolios will have the constraint active. The figure also contains CVaR_{90} and the expected return of the portfolios minimizing $\text{CVaR}_{99.5}$ to show the limit of models achievable by the model (M2).

For a loss tolerance greater than 5.471%, efficient portfolios with a low expected return will be equivalent to the benchmark model, because the efficient portfolios of the benchmark model meet the constraint. However, as the target return increases, the portfolios must decouple as the constraint becomes active. This divergence is visible in Figure 5.7, where the portfolios with constraint $\text{CVaR}_{99.5} \leq 5.500\%$ exhibit this behavior.

Another observation is that, at the highest feasible expected return, portfolios tend to contain fewer assets compared to those produced by the benchmark model. This behavior is not unique to the specific loss tolerance of 5.500%. Since limited diversification typically increases risk, caution is warranted when investing in such portfolios, even though their risk measures may remain within tolerance levels.

To construct three portfolios with low, medium, and high risk, with a target return corresponding to the previous, we select appropriate loss tolerances. For a high-risk portfolio, the choice might not seem very important. Figure 5.6 suggests that any efficient portfolio will be very close to the one presented by the benchmark model. The choice is more impactful for low and medium risk, where we prefer portfolios which are different from the benchmark, in order to compare the two methods while avoiding the low-diversity portfolios close to the highest returns.

For a low-risk portfolio with expected return 0.060%, we set loss tolerance $l = 5.250\%$, and for medium-risk, we choose $l = 5.700\%$, constructing a portfolio with expected return 0.085%. For the high-risk portfolio, we use $l = 6.500\%$ and the target return 0.105%. The weights of the portfolio are recorded in Table 5.4, together with the L^1 distances from the portfolios produced by the benchmark model.

Although the assets included in all risk-level portfolios remained almost unchanged (the medium-risk portfolio being the only one to drop an asset, *Coca-Cola*), the overall distance from the benchmark is not insignificant. This is particularly surprising for the high-risk portfolio, where we expected a near match based on Figure 5.6. Furthermore, the portfolio no longer contains assets with a weight greater than 0.5, suggesting a more balanced allocation. Since CVaR_{90} values did not increase dramatically, the comparison of the out-of-sample performance could provide interesting results.

Risk tolerance	**Low**	**Medium**	**High**
$\mu(\boldsymbol{w})$ (%)	0.060	0.085	0.105
$\text{CVaR}_{99.5}$ (%)	5.250	5.700	6.500
CVaR_{90} (%)	1.811	2.022	2.514
AAPL	–	0.101	0.190
AMZN	0.146	0.138	0.276
BAC	–	–	–
COST	0.168	0.209	0.064
IBM	–	–	–
JNJ	0.448	0.257	–
JPM	–	–	–
KO	0.036	–	–
T	0.154	–	–
UNH	0.048	0.294	0.470
L^1 dist	0.689	0.464	0.259

Table 5.4 Portfolio weights of extreme-tail model (M2) for selected target returns with $\text{CVaR}_{99.5}$ constraint, distance from benchmark model (M1) portfolios with equivalent expected return.

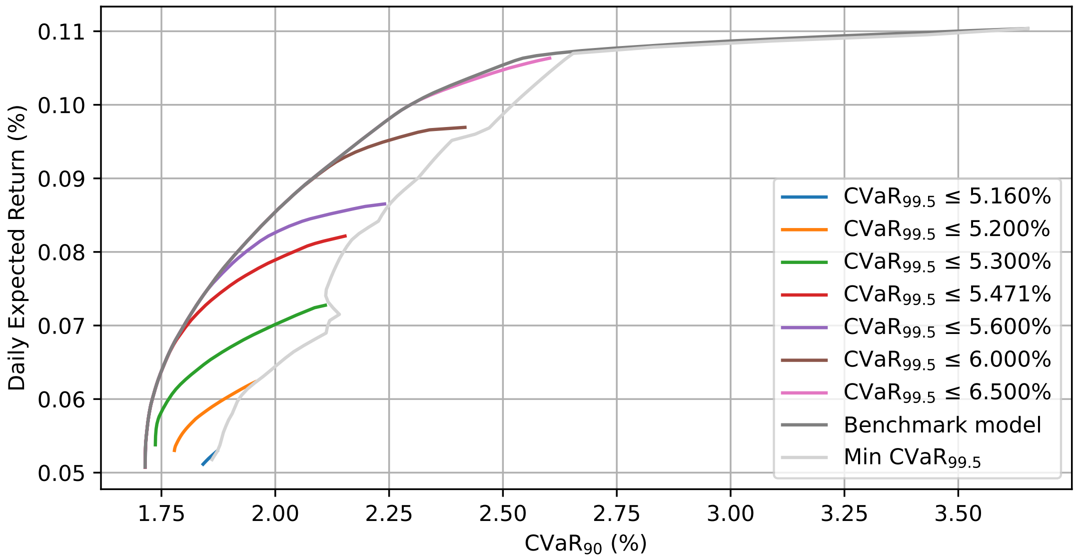


Figure 5.6 Expected daily return and $CVaR_{90}$ of efficient sets at different loss tolerance for $CVaR_{99.5}$. Benchmark model portfolios and portfolios minimizing $CVaR_{99.5}$ included.

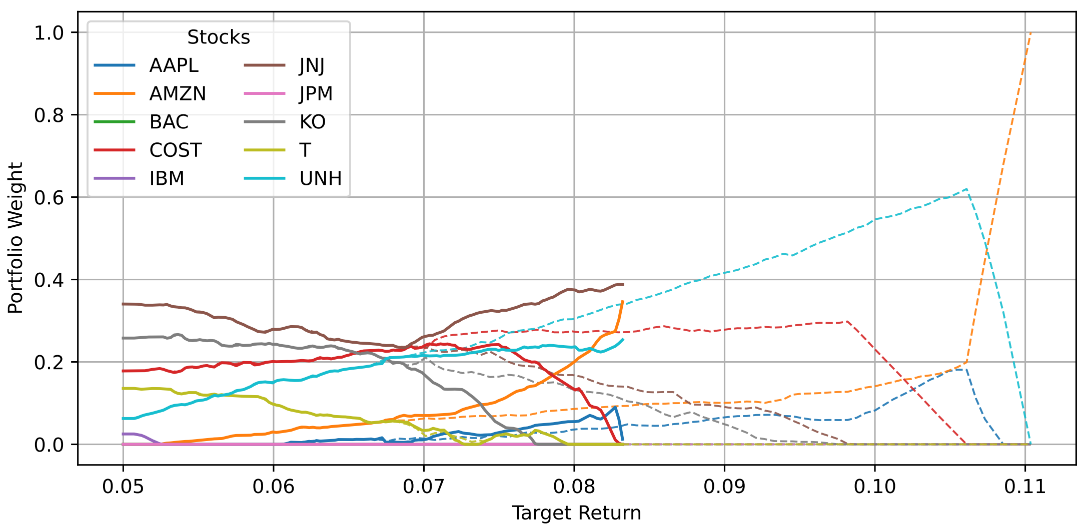


Figure 5.7 Weights of efficient portfolio with targeted expected return and $CVaR_{99.5} \leq 5.5\%$ denoted by full lines, in comparison with the benchmark model (dashed lines).

5.5 General mean-CVaR model

The most comprehensive pure mean-CVaR approach includes both risks and expected return in the objective function, as presented by the model (3.6). The first decision is the selection of CVaR levels to include. For each target return μ, efficient portfolios will have CVaR_{α_1} between two extremes, the minimum CVaR_{α_1} achievable at that return, and CVaR_{α_1} value of the portfolio that minimizes CVaR_{α_2} at μ. We will set α_1 at 0.9 to match the CVaR models.

Figure 5.8 illustrates how some choices of α_2, for example 0.8 or 0.95 lead to a very small range of values of CVaR_{90}. The efficient set $\boldsymbol{f}(\mathcal{W}_E)$ collapses, reflecting the connected, but not necessarily convex geometry discussed in Section 3.3. At the other extreme, $\text{CVaR}_{99.9}$ is also suspicious, providing a very smooth curve in Figure 5.8. This is likely due to the lack of data in that tail, leading to concentrated portfolios.

We decided to simulate two models, first with $\alpha_2 = 0.95$ to see if the small change in weights can achieve meaningful differences in portfolio performance, and second $\alpha_2 = 0.995$, creating a natural extension of the model (M2), which might provide more interesting results to study.

The two models are as follows:

$$\begin{aligned} \min_{\boldsymbol{w}\in\mathbb{R}^{10}} \ & \Big(-\mu(\boldsymbol{w}),\ \text{CVaR}_{90}(-\boldsymbol{w}^T\boldsymbol{r}),\ \text{CVaR}_{95}(-\boldsymbol{w}^T\boldsymbol{r})\Big), \\ \text{s.t. } & w_j \geq 0,\ j = 1,\dots,10, \\ & \sum_{k=1}^{10} w_k = 1, \end{aligned} \tag{M3}$$

$$\begin{aligned} \min_{\boldsymbol{w}\in\mathbb{R}^{10}} \ & \Big(-\mu(\boldsymbol{w}),\ \text{CVaR}_{90}(-\boldsymbol{w}^T\boldsymbol{r}),\ \text{CVaR}_{99.5}(-\boldsymbol{w}^T\boldsymbol{r})\Big), \\ \text{s.t. } & w_j \geq 0,\ j = 1,\dots,10, \\ & \sum_{k=1}^{10} w_k = 1. \end{aligned} \tag{M4}$$

5.5.1 Model M3

The methods necessary to find the efficient set of this model were presented in Section 2.2. Furthermore, in Section 3.3 it was mentioned that the image of the efficient set $\boldsymbol{f}(\mathcal{W}_E)$ is a connected, closed subset of a convex set boundary. Projecting onto the mean-CVaR_{90} plane, it is sufficient to trace a "boundary" of this projection via minimization of two objectives. All objective values within this "boundary" must be achievable by an efficient portfolio, which minimizes CVaR_{95}. This boundary for the model (M3) is depicted in Figure 5.9.

According to the theory presented in the previous chapters, the efficient set of this model may be non-convex. We can show that using the following portfolios. The efficient set contains efficient portfolios $X_1,\ X_2$ defined by decision vectors $\boldsymbol{w}_1, \boldsymbol{w}_2 \in \mathcal{W}_E$ such that:

- $\mu(\boldsymbol{w}_1) = 0.060\%$, $\text{CVaR}_{90}(L_{X_1}) = 1.732\%$ and $\text{CVaR}_{95}(L_{X_1}) = 2.351\%$,

- $\mu(\boldsymbol{w}_2) = 0.105\%$, $\mathrm{CVaR}_{90}(L_{X_2}) = 2.484\%$ and $\mathrm{CVaR}_{95}(L_{X_2}) = 3.244\%$.

If $\mathcal{W}_E$ was convex, we could find an efficient portfolio $\widehat{X}_3$ with expected return 0.085% as the linear combination $0.444\boldsymbol{w}_1 + 0.556\boldsymbol{w}_2$. The new portfolio has:

- $\mathrm{CVaR}_{90}(L_{\widehat{X}_3}) = 2.035\%$ and $\mathrm{CVaR}_{95}(L_{\widehat{X}_3}) = 2.727\%$.

However, there exists a portfolio X_3 with decision vector $\boldsymbol{w}_3 \in \mathcal{W}_E$, such that:

- $\mu(\boldsymbol{w}_3) = 0.085\%$, $\mathrm{CVaR}_{90}(L_{X_3}) = 2.000\%$ and $\mathrm{CVaR}_{95}(L_{X_3}) = 2.650\%$,

thus rendering $\widehat{X}_3$ inefficient and $\mathcal{W}_E$ non-convex.

Instead of choosing some small set of efficient portfolios produced by this model, we analyze the entire subset of efficient portfolios with specific expected returns. Our goal will not be to find the best portfolio in this set, but rather to study how out-of-sample properties evolve along the shifting focus from CVaR_{90} to CVaR_{95}. Figure 5.10 shows the weight distribution for any efficient CVaR_{90} value for the expected return 0.085%, which corresponds to our medium-risk portfolios from previous sections.

The values of CVaR_{95} in Figure 5.9 decrease rapidly at first, however, as CVaR_{90} increases, the advantage becomes smaller. If the goal was to select a specific portfolio, an option with CVaR_{90} of 1.995% would seem reasonable. Overall, the portfolio weights change the most towards the lowest CVaR_{90} values, suggesting that the introduction of a new objective might have an effect despite the small size of efficient set.

Analysis of individual weights reveals that *Coca-Cola* is again the only stock to be excluded at lower CVaR_{95}, suggesting that it does not perform well in terms of higher CVaR levels. Also similarly to extreme-tail model (M2), *Johnson & Johnson* sees increased representation, performing better higher CVaR levels.

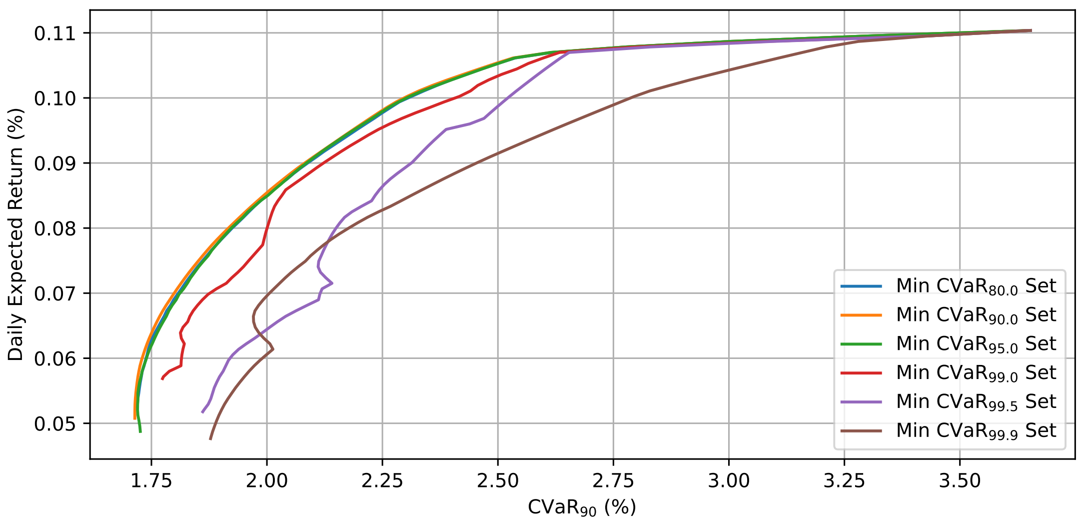


Figure 5.8 Expected daily return and CVaR_{90} of efficient sets at different levels of CVaR. CVaR_{80} and CVaR_{95} are virtually indistinguishable.

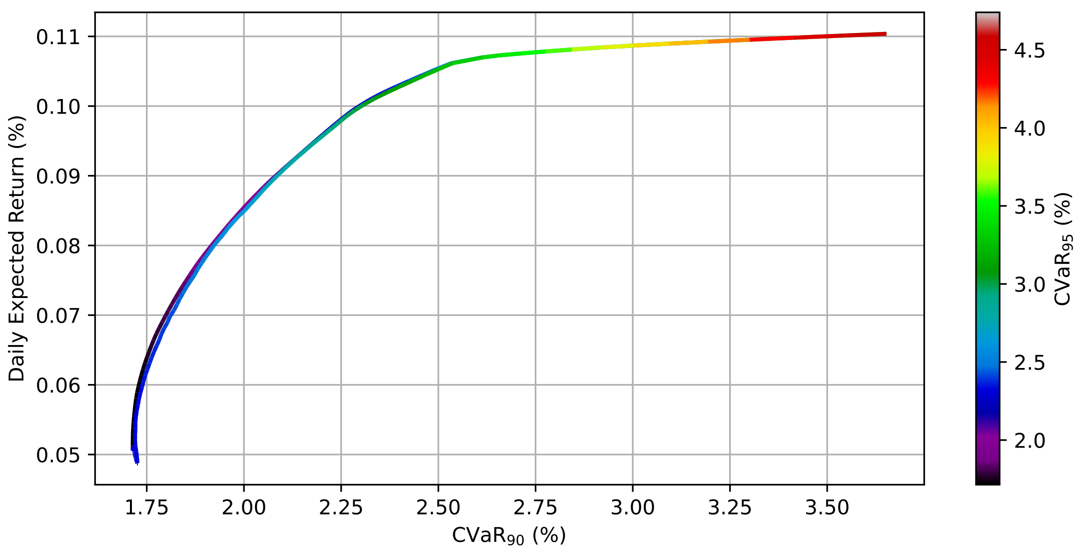


Figure 5.9 Expected daily return and CVaR_{90} for efficient set "boundary" of model (M3) with CVaR_{95} depicted using color scale.

5.5.2 Model M4

The analysis of model M4 is very similar to the previous model M3, however, the efficient set is significantly larger (see Figure 5.12). At medium-risk return (0.085%), Figure 5.11 shows more dramatic weight changes, but general trends remained. Once again, the portfolios are least stable at most extreme CVaR values. For low CVaR_{90}, there are more assets included in the portfolio, but the weights change very rapidly. A slight increase in CVaR_{90} provides a significant decrease in $\text{CVaR}_{99.5}$ and stabilization of the weights. At the other end of the spectrum, near the minimum $\text{CVaR}_{99.5}$, the weights do not behave as erratically. However, portfolios with low $\text{CVaR}_{99.5}$ are the least diversified, again suggesting that options in between might be advantageous.

Figure 5.12 also shows the weights and CVaR values of the benchmark and extreme-tail model portfolios. The benchmark model portfolio has a relatively high $\text{CVaR}_{99.5}$ value, as it does not optimize for it at all. The portfolio produced by the extreme-tail model (M2) improves on that, but if we were to choose a single portfolio for medium risk, we might select an even higher CVaR_{90} value that provides a better CVaR trade-off.

The entire efficient set is depicted in Figure 5.12. As mentioned above, the return-CVaR_{90} image of the set can be obtained from the "boundary", but to analyze $\text{CVaR}_{99.5}$, we also simulated values inside. The development of $\text{CVaR}_{99.5}$ can be analyzed more precisely using the contour lines presented in Figure 5.13, which clearly shows that the significant decrease in $\text{CVaR}_{99.5}$ close to the lowest CVaR_{90} portfolios is not limited to the expected return of 0.085%. It is present at all levels except for very high returns, where the efficient set once again collapses, with portfolios that minimize both CVaR levels very close to each other.

To show that the efficient set of this model is also not convex, we will use a similar example. Let us consider efficient portfolios X_1, X_2 and X_3 with weights $\boldsymbol{w}_1, \boldsymbol{w}_3, \boldsymbol{w}_3 \in \mathcal{W}_E$ and following properties:

- $\mu(\boldsymbol{w}_1) = 0.060\%$, $\text{CVaR}_{90}(L_{X_1}) = 1.800\%$ and $\text{CVaR}_{99.5}(L_{X_1}) = 5.260\%$,
- $\mu(\boldsymbol{w}_2) = 0.105\%$, $\text{CVaR}_{90}(L_{X_2}) = 2.500\%$ and $\text{CVaR}_{99.5}(L_{X_2}) = 6.567\%$,
- $\mu(\boldsymbol{w}_3) = 0.085\%$, $\text{CVaR}_{90}(L_{X_3}) = 2.050\%$ and $\text{CVaR}_{99.5}(L_{X_3}) = 5.651\%$.

The linear combination $0.444\boldsymbol{w}_1 + 0.556\boldsymbol{w}_2$ also produces portfolio $\widehat{X}_3$ with expected return 0.085%, with the corresponding CVaR values:

- $\text{CVaR}_{90}(L_{\widehat{X}_3}) = 2.080\%$ and $\text{CVaR}_{99.5}(L_{\widehat{X}_3}) = 5.730\%$.

The portfolio X_3 therefore dominates $\widehat{X}_3$, the convex combination of two efficient portfolios is not efficient, and the efficient set is not convex.

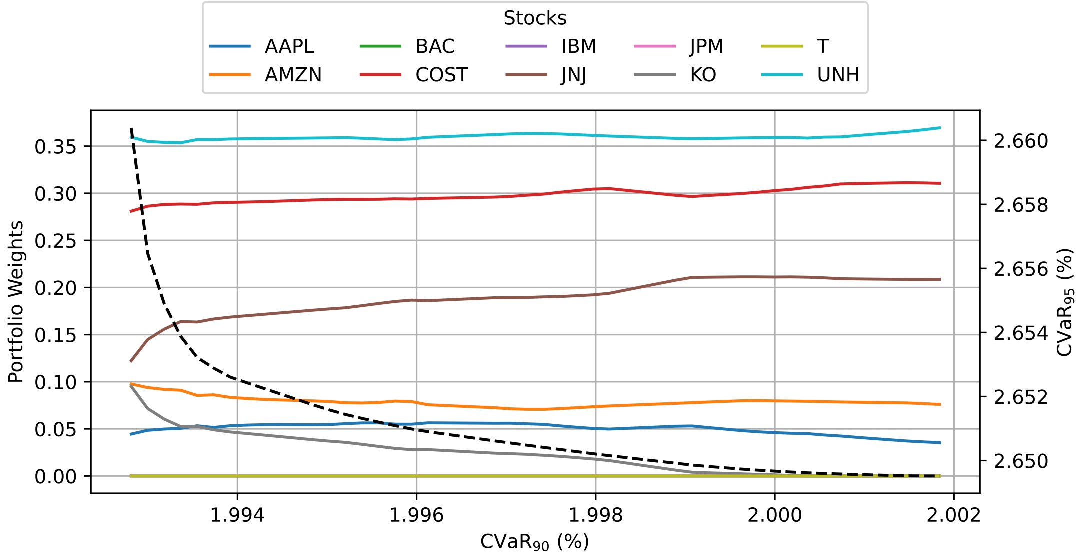


Figure 5.10 Weights of efficient portfolios with expected return 0.085% from model (M3). Dashed line shows CVaR_{95} value of the portfolios.

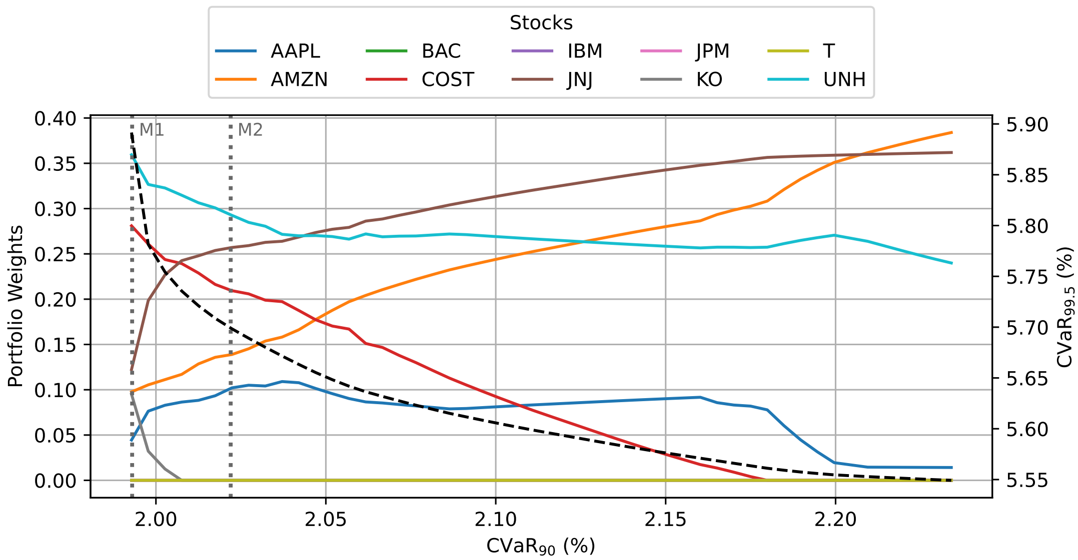


Figure 5.11 Weights of efficient portfolios with expected return 0.085% from model (M4). Dashed line shows CVaR_{95} value of the portfolios. Also shows positions of portfolios produced by benchmark model (M1) and extreme-tail model (M2).

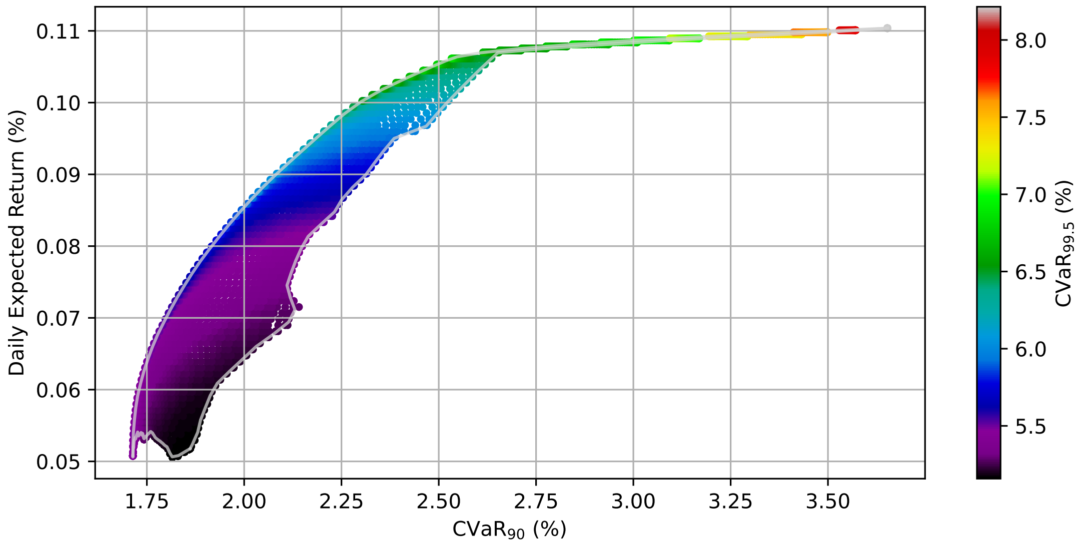


Figure 5.12 Expected daily return and CVaR_{90} for efficient set of model (M4) with $\mathrm{CVaR}_{99.5}$ depicted using color scale.

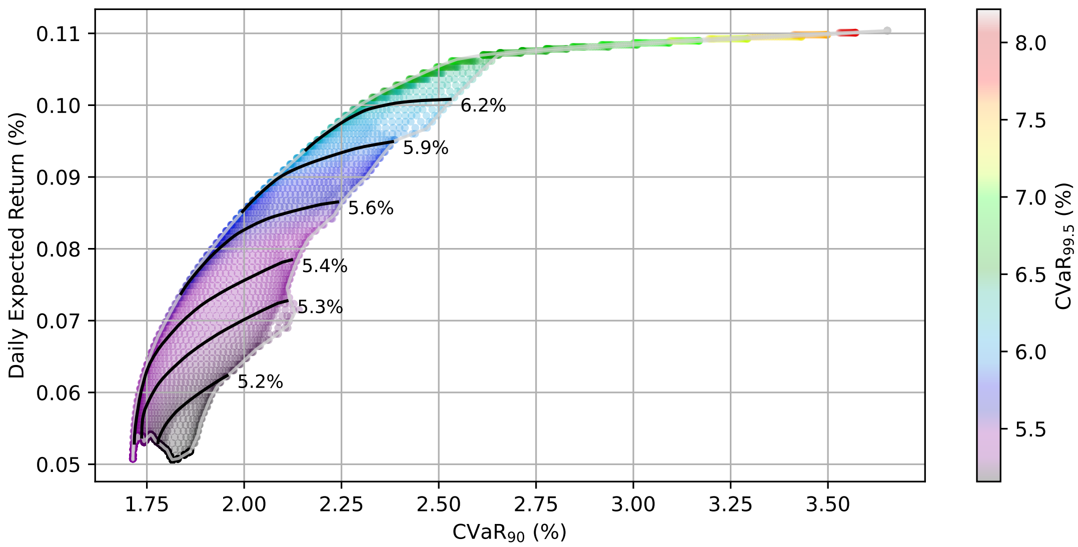


Figure 5.13 Efficient set of model (M4) with $\mathrm{CVaR}_{99.5}$ emphasized using contour lines.

5.6 Model with transaction cost

The last presented model, based on Chapter 4 and specifically (4.1) is formulated as follows:

$$\begin{aligned} \min_{\boldsymbol{w}\in\mathbb{R}^{10}} \ & \Big(-\mu(\boldsymbol{w}),\ \mathrm{CVaR}_{90}(-\boldsymbol{w}^T\boldsymbol{r}),\ \|\boldsymbol{w}-\boldsymbol{w}_0\|_1\Big), \\ \text{s.t. } & w_j \geq 0,\ j=1,\dots,10, \\ & \sum_{k=1}^{10} w_k = 1, \end{aligned} \tag{M5}$$

where $\boldsymbol{w}_0$ are weights of some starting portfolio. We will use a starting portfolio with all weights equal, set to $\frac{1}{10}$.

The goal of this model is to demonstrate how another criterion can be implemented in the mean-CVaR framework. The portfolios constructed are not guaranteed to be non-dominated with respect to second-order stochastic dominance, and we do not expect to find portfolios that would outperform the previous models.

We can once again find the "boundary" of the image of the efficient set projected into the mean-CVaR space. Any efficient portfolio with unique objective values from the benchmark model (M1) must also be efficient for the model (M5), as these are the only portfolios that maximize the return for a specific CVaR_{90} value. In addition, any convex combination of the highest-return asset and the starting portfolio lies on this efficient "boundary" as it maximizes expected return for its distance from the starting portfolio. Lastly, we need to find portfolios that minimize CVaR_{90} and distance.

An image of the efficient set comparable to previous models is presented in Figure 5.14, where we can observe all mean-CVaR trade-off options provided by model (M5). It might seem very similar to the efficient set of a previous model (M4) in Figure 5.12, however, this similarity is purely coincidental, caused by the fact that our starting model performs very similarly in terms of expected return and CVaR_{90} to some portfolios that minimize $\mathrm{CVaR}_{99.5}$. The portfolio weights and $\mathrm{CVaR}_{99.5}$ of the portfolios are not similar, and the efficient sets differ significantly.

Perhaps a more useful representation of $\boldsymbol{f}(\mathcal{W}_E)$ is presented in Figure 5.15, which shows the expected return to the L^1 distance trade-off, with CVaR_{90} represented by the contour lines. The information presented in this figure can allow for some reasonable choices with respect to transaction cost. We can notice that for any level of return, CVaR_{90} will initially decrease rapidly with increasing distance, but at higher distances CVaR_{90} almost plateaus, suggesting that if transaction cost is a concern, the investor might prefer a portfolio that is cheaper to acquire and performs very similarly.

Furthermore, the distance of mean-CVaR efficient portfolios differs significantly with relatively small changes in expected return. This allows the investor to choose a return level with lower acquisition cost of a mean-CVaR optimal portfolio. Combining both of these strategies for the selection of portfolios comparable to the previous models, we chose portfolios with expected returns 0.060%, 0.085% and 0.105%, and L^1 distances 0.85, 0.75 and 1.25 for a low-, medium- and high-risk portfolio, respectively. We will be interested in whether the performance out-of-sample is significantly impacted in comparison to the other options.

Since all convex combinations of the starting portfolio and highest-return asset are efficient, we could ask whether the new objective caused the entire efficient set to be convex. However, a convex combination of the low- and high-risk portfolio with expected return 0.085% has both CVaR_{90} and distance from the starting portfolio larger than the medium-risk portfolio, rendering the efficient set non-convex once again.

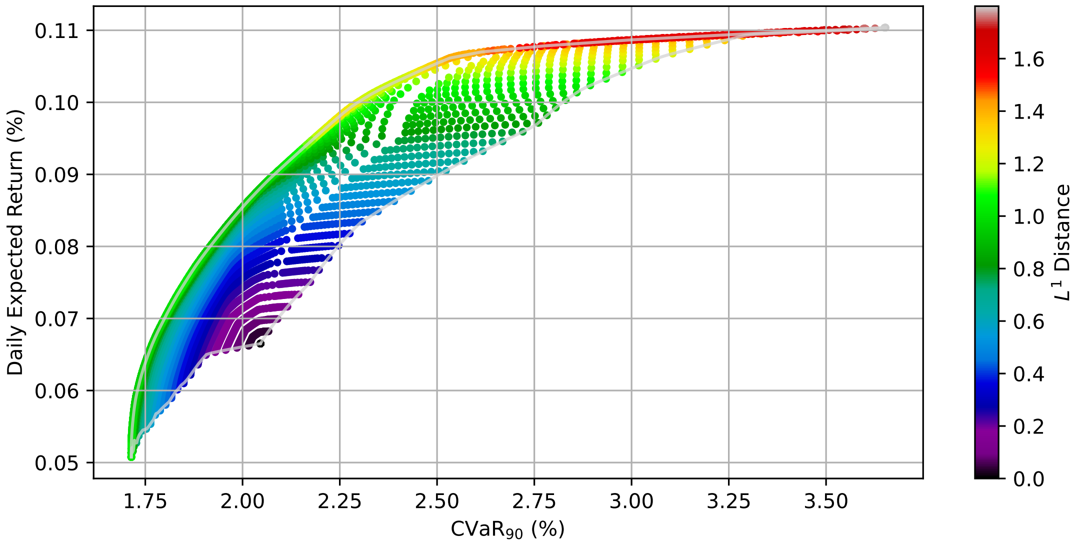


Figure 5.14 Expected daily return and CVaR_{90} for efficient set of model (M5) with L^1 distance depicted using color scale.

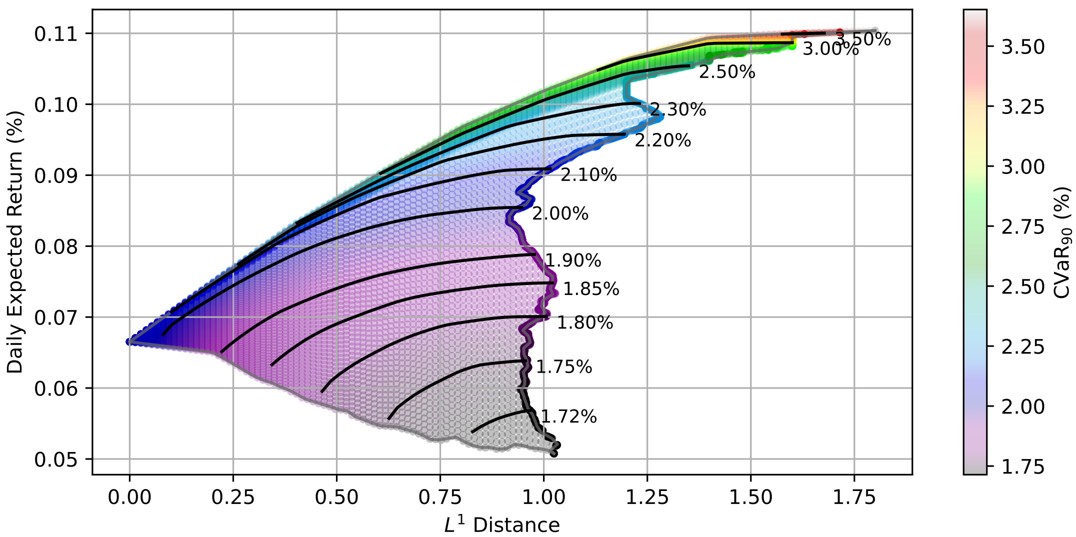


Figure 5.15 Expected daily return and L^1 distance form equal weights starting portfolio for the efficient set of model (M5) with CVaR_{90} emphasized using contour lines.

5.7 Out-of-sample performance

To evaluate how the portfolios might fare in practice, we examine their performance on data not used for their construction. Specifically, we measure the return and CVaR_{90} in the out-of-sample data, daily returns for 500 trading days that directly follow the period used to construct the portfolios. To better encapsulate the risks associated with a specific portfolio, we also look at maximum one-day loss. Although a medium- to long-term investor might not be particularly interested in a one-day loss, a suspiciously high value could suggest a flawed portfolio with insufficient diversity or unusually highly correlated assets.

Furthermore, we analyze the relative maximum drawdown, i.e. the largest relative decrease of the price during the period. This risk measure shows the losses inquired in a scenario where an investor bought and sold the portfolio at the worst possible time in the observed sample. This measure is very practical, as it shows a potential investor the absolute worst scenario that could have happened.

Lastly, we look at the L^1 distance from the starting equal weight portfolio. This is not a very useful measure to choose a correct investment, but it serves well to show the possibility of a decrease in acquisition costs that is not detrimental to the performance of the portfolio.

The results of specific portfolios we mentioned in the previous sections, produced by the corresponding models (M1, M2, M5) are presented in Table 5.5.

The three portfolios produced by the benchmark model (M1) underperformed in terms of return. However, this was not accompanied by an increase in CVaR_{90}. This suggests that the underperformance is due to generally lower volatility in the out-of-sample period, rather than increased losses. Note that direct comparisons of maximum loss and drawdown are not possible between the samples, as they differ in size.

The extreme-tail constraint model (M2) either improved or matched (within 0.001 p.p.) the performance of the benchmark model in all metrics except the maximum one-day loss. The low-risk portfolio delivered 0.061% daily out-of-sample returns (16.611% annualized), matching its target. More details on how this performance was achieved will be produced in the analysis of the model (M4), which can also produce these portfolios.

The transaction cost constrained portfolios produced by the model (M5) also performed quite well, matching or exceeding the benchmark across all metrics except the maximum one-day loss. This is not to suggest that limiting the L^1 distance from an arbitrary portfolio should systematically improve performance, but highlights the fact that, if transaction cost was a focus, similar models based on the principles presented in this thesis would be worth further study.

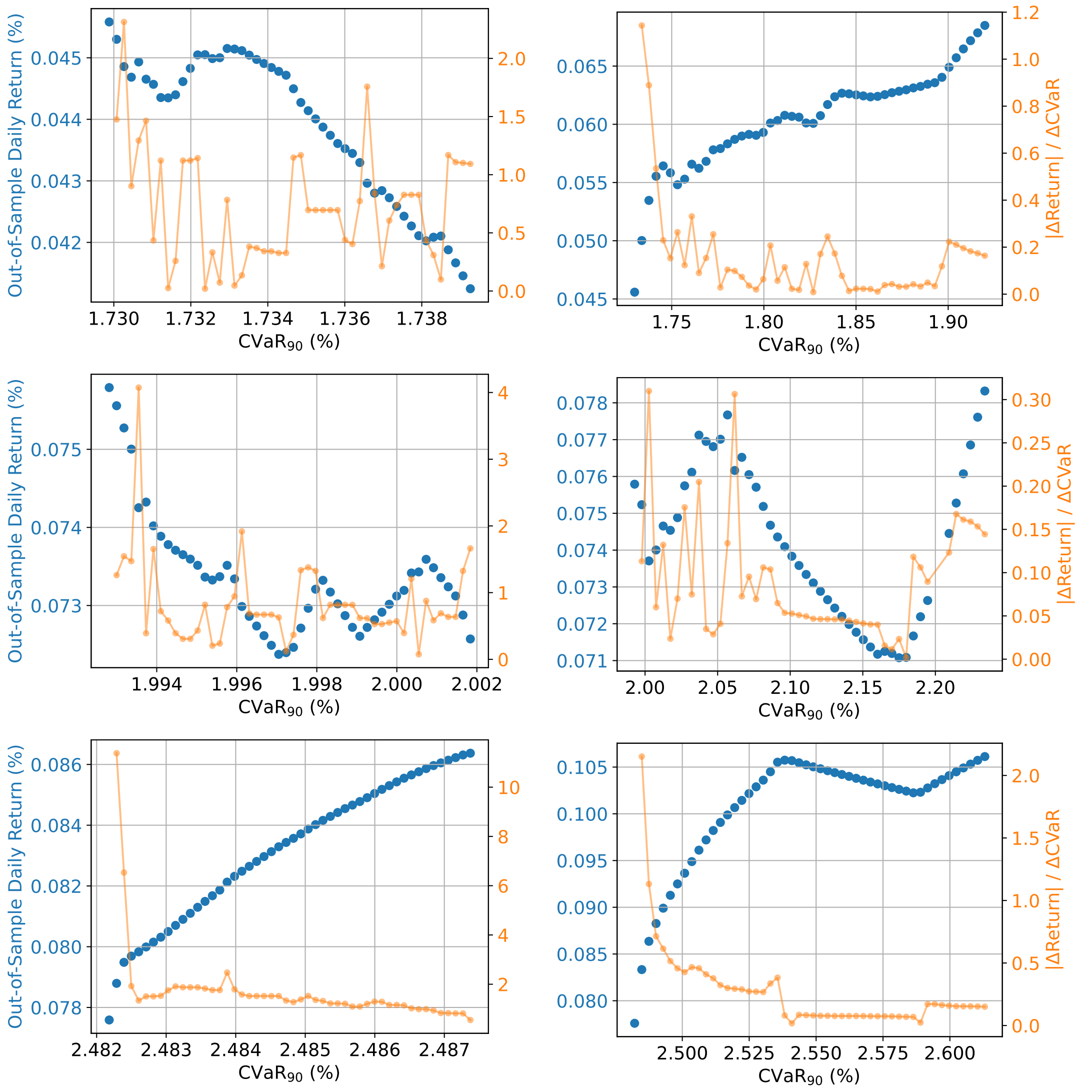
Lastly, we analyze the sets of portfolios presented by the models (M3) and (M4). Figure 5.16a shows the daily return with respect to the in-sample CVaR_{90} value at a specific expected return level. The portfolios produced by the model (M3), which has a very small range of efficient CVaR_{90} values and, therefore, the range of all returns is smaller than 0.009 percentage points. Furthermore, the returns behave very differently for the three target return levels. The returns of low-risk portfolios slowly decrease with increasing CVaR_{90}, the medium return has a very small decrease at the beginning and then remains nearly constant, while high-risk portfolios see an increase in returns with increasing target CVaR_{90}.

The returns of the model (M4) in Figure 5.16b are perhaps more interesting, due to the wider range of possible CVaR_{90}. The low- and high-risk portfolios both see an increase in returns at higher CVaR_{90} and therefore lower $\text{CVaR}_{99.5}$. In comparison, changes in medium-risk portfolio returns are less significant, with no clear trend. This is certainly not sufficient evidence that the minimization of $\text{CVaR}_{99.5}$ is better than CVaR_{90}, rather a suggestion that a larger simulation with more assets and multiple testing time periods might provide interesting results.

A trend that does seem to persist in all subplots in Figure 5.16a and 5.16b, is that the stability of returns with respect to CVaR changes, depicted in orange, is the lowest in portfolios that significantly prioritize one CVaR value over the other, especially if the prioritized value is CVaR_{90}. This suggests that while benefits in terms of return are unclear, multicriteria optimization at two CVaR levels could be a more robust method compared to a single CVaR level.

Model	**M1**		**M2**		**M5**	
Risk Approach	**Low** (0.060% Daily return)					
Sample	in	out	in	out	in	out
Daily Return (%)	0.060	0.046	0.060	0.061	0.060	0.049
CVaR (%)	1.730	1.083	1.811	1.226	1.735	1.076
Max Loss (%)	7.877	2.275	6.565	2.392	7.759	2.333
Max Drawdown (%)	27.156	8.921	22.743	9.601	27.045	8.826
L^1 **Distance**	0.949		1.033		0.850	
Risk Approach	**Medium** (0.085% Daily return)					
Sample	in	out	in	out	in	out
Daily Return (%)	0.085	0.076	0.085	0.075	0.085	0.077
CVaR (%)	1.993	1.372	2.022	1.310	2.007	1.322
Max Loss (%)	10.364	2.902	9.822	3.013	10.507	3.020
Max Drawdown (%)	25.567	8.452	24.333	8.524	26.338	7.664
L^1 **Distance**	0.926		1.000		0.750	
Risk Approach	**High** (0.105% Daily return)					
Sample	in	out	in	out	in	out
Daily Return (%)	0.105	0.078	0.105	0.099	0.105	0.105
CVaR (%)	2.482	1.814	2.514	1.713	2.529	1.781
Max Loss (%)	13.892	4.630	12.450	3.787	11.858	3.700
Max Drawdown (%)	29.909	9.990	27.112	10.120	25.661	9.728
L^1 **Distance**	1.318		1.272		1.250	

Table 5.5 Return and risk measures of low-, medium- and high-risk portfolios produced by the benchmark model (M1), the extreme-tail model (M2) and the transaction cost model (M5). Last lines contains the L^1 distance from the equal weight starting portfolio.



(a) M3 model.

(b) M4 model.

Figure 5.16 Daily returns with respect to the target $CVaR_{90}$ value of portfolios with target returns 0.060%, 0.085%, 0.085%, in that order. Also shown is the ratio of absolute change in return and change in the target $CVaR_{90}$ compared to the previous portfolio, to illustrate stability of solutions.

Conclusion

This thesis addresses multi-criteria portfolio optimization using mean-CVaR models and investigates the effects of additional optimization criteria, which allows the investor to set a risk profile that improves alignment with their subjective risk preferences.

In the first chapter, we introduced two risk measures, VaR and CVaR, highlighting their connection to first- and second-order stochastic dominance. This ties the thesis to a theoretic framework and clarifies why incorporating multiple CVaR levels can shape more appropriate portfolios. We also defined a key property of risk measures, coherence, which shows why CVaR is generally preferred over VaR.

The second chapter focused on portfolio efficiency. We formalized portfolio dominance and defined an efficient set of a model. We also proved that, while efficient sets of the considered models are closed and connected, they need not be convex. Lastly, we showed two optimization techniques, weighted sums and ε-constraints, used to calculate the efficient frontier.

The following two chapters presented the specific models. The third chapter focused on pure mean-CVaR models. We systematically compared approaches, treating CVaR as an objective, a constraint, or both, and demonstrated how these models can be approached using a spectral risk measure reformulation. That reformulation broadens the flexibility of mean-CVaR methods, allowing for even more specific risk preferences.

Chapter four introduced transaction cost as a non-CVaR criterion, modeled using the $\mathcal{L}^1$ norm, measured in relation to some starting portfolio. We show how this alters the efficient set and discuss possibly necessary normalization of data prior to the portfolio calculations. We also considered an alternative, constraints-based approach, which yielded a traditional mean-CVaR model with limited feasible set.

The last chapter presented results numerical study which applied all models to real-world data. We confirmed the presented theoretical properties empirically and compared out-of-sample performance. Although none of the extensions consistently improved the returns in our limited dataset, the multi-CVaR formulations appeared to improve stability, suggesting that this approach could be more robust compared to the single CVaR level.

We presented an extension to the traditional mean-CVaR approach, studied the effects on the efficient set, and verified our results using a numerical study. Our findings offer insights for investors using multi-criteria risk frameworks with CVaR risk measure. A natural extension of this thesis could focus on providing a more comprehensive and realistic approach to portfolio optimization with transaction cost, introducing an explicit cost, and measuring the performance of regularly updated portfolios that attempt to retain their performance while reducing transaction cost in a similar way to what we presented in the last chapter.

Additionally, a significantly larger numerical study could be performed with multiple time periods and larger pools of assets, no longer focused on confirming the theoretical properties but looking for statistically significant differences in performance of the presented models. It would also be interesting to analyze whether the observed robustness benefits persist in that study.

Bibliography

1. Markowitz, H. Portfolio Selection. *The Journal of Finance*. 1952, vol. 7, no. 1, pp. 77–91. Available from DOI: 10.2307/2975974.
2. Artzner, P.; Delbaen, F.; Jean-Marc, E.; Heath, D. Coherent Measures of Risk. *Mathematical Finance*. 1999, vol. 9, pp. 203–228. Available from DOI: 10.1111/1467-9965.00068.
3. Rockafellar, R. T.; Uryasev, S. Optimization of conditional value-at-risk. *Journal of Risk*. 2000, vol. 3, pp. 21–41. Available from DOI: 10.21314/JOR.2000.038.
4. Rockafellar, R. T.; Uryasev, S. Conditional value-at-risk for general loss distributions. *Journal of Banking & Finance*. 2002, vol. 26, no. 7, pp. 1443–1471. Available from DOI: 10.1016/S0378-4266(02)00271-6.
5. Levy, H. *Stochastic Dominance: Investment Decision Making under Uncertainty*. 3rd ed. Cham, Switzerland: Springer, 2015. ISBN 978-3-319-21707-9. Available from DOI: 10.1007/978-3-319-21708-6.
6. Hanoch, G.; Levy, H. The Efficiency Analysis of Choices Involving Risk. *The Review of Economic Studies*. 1969, vol. 36, no. 3, pp. 335–346. Available from DOI: 10.2307/2296431.
7. Acerbi, C.; Tasche, D. On the Coherence of Expected Shortfall. *Journal of Banking & Finance*. 2002, vol. 26, pp. 1487–1503. Available from DOI: 10.1016/S0378-4266(02)00283-2.
8. Rockafellar, R. T. *Convex Analysis*. Princeton: Princeton University Press, 1970. ISBN 978-1-400-87317-3. Available from DOI: 10.1515/9781400873173.
9. Ehrgott, M. *Multicriteria Optimization*. 2nd ed. Berlin, Heidelberg: Springer, 2005. ISBN 978-3-540-21398-7. Available from DOI: 10.1007/3-540-27659-9.
10. Mishra, S.; Wang, S.-Y.; Lai, K. K. Generalized Convexity and Vector Optimization. *Nonconvex Optimization and its Applications*. 2009, vol. 90. Available from DOI: 10.1007/978-3-540-85671-9.
11. Miettinen, K. *Nonlinear Multiobjective Optimization*. New York, NY: Springer, 1998. ISBN 978-0-7923-8278-2. Available from DOI: 10.1007/978-1-4615-5563-6.
12. Hartley, R. On Cone-Efficiency, Cone-Convexity and Cone-Compactness. *SIAM Journal on Applied Mathematics*. 1978, vol. 34, no. 2, pp. 211–222. Available from DOI: 10.1137/0134018.
13. Dupačová, J.; Hurt, J.; Štěpán, J. *Stochastic Modeling in Economics and Finance*. New York, NY: Springer, 2002. ISBN 978-1-4020-0840-5. Available from DOI: 10.1007/b101992.
14. Haimes, Y.; Leon, S. L.; Dang, D. On a bicriterion formation of the problems of integrated system identification and system optimization. *IEEE Transactions on Systems, Man, and Cybernetics*. 1971, pp. 296–297. Available from DOI: 10.1109/TSMC.1971.4308298.

15. Kim, J. H. On relation among coherent, distortion and spectral risk measures. *Theoretical Mathematics and Pedagogical Mathematics*. 2009, vol. 16, no. 1, pp. 121–131.

16. Acerbi, C. Spectral measures of risk: A coherent representation of subjective risk aversion. *Journal of Banking & Finance*. 2002, vol. 26, no. 7, pp. 1505–1518. Available from doi: 10.1016/S0378-4266(02)00281-9.

17. Diamond, S.; Eric Chu, E.; Boyd, S. *CVXPY (version 0.2)*. 2014. Available also from: http://cvxpy.org/. A Python-Embedded Modeling Language for Convex Optimization.

18. Aroussi, R. *yfinance (version 0.2.65)*. 2025. Available also from: https://pypi.org/project/yfinance/. Python package for downloading market data from Yahoo Finance.

19. Cipra, T. *Time Series in Economics and Finance*. 1st ed. Cham, Switzerland: Springer, 2020. isbn 978-3-030-46346-5. Available from doi: 10.1007/978-3-030-46347-2.

A Additional theorems

A.1 Highest CVaR portfolios

Lemma A1. *Any feasible portfolio constructed using three assets X_1, X_2, X_3 with returns $\boldsymbol{r}^T = (r_1, r_2, r_3)$, defined by weights $\boldsymbol{w}^T \in \mathbb{R}^3_+$, $w_1 + w_2 + w_3 = 1$ and with expected return $\mu^*(\boldsymbol{w}) = \mathsf{E}(w_1 r_1 + w_2 r_2 + w_3 r_3)$ can be expressed as a convex combination of feasible portfolios with equivalent expected return $\mu(\boldsymbol{w})$ and at most two positive weights.*

Proof. Let $\Delta = \{\boldsymbol{w} \in \mathbb{R}^3_+ \colon w_1 + w_2 + w_3 = 1\}$ be the set of all feasible weights and $H = \{\boldsymbol{w} \in \mathbb{R}^3_+ \colon \mathsf{E}(w_1 r_1 + w_2 r_2 + w_3 r_3) = \mu^*(\boldsymbol{w})\}$ set of all weights (regardless of feasibility), such that the expected return is $\mu^*(\boldsymbol{w})$.

Δ is a triangle in $\mathbb{R}^3$ and H is a hyperplane, all feasible portfolios with the expected return $\mu^*(\boldsymbol{w})$ lie in their intersection $\mathcal{W}_{\mu^*} = \Delta \cap H$. We show that the statement holds for any shape of $\mathcal{W}_{\mu^*}$.

1. $\mathcal{W}_{\mu^*}$ is an empty set, no feasible portfolio with desired return exists.
2. $\mathcal{W}_{\mu^*}$ is a single point, the only feasible portfolio is a vertex of Δ, where two of the weights are 0.
3. $\mathcal{W}_{\mu^*}$ is a line segment, where the end points lie on the edges or vertices of Δ, where at least one of the weights is 0, and any other point is a convex combination of the end points.
4. $\mathcal{W}_{\mu^*}$ is the entire Δ, which is convex and every point on its boundary has at least one weight equal to 0.

□

Theorem A2. *Let $\mathcal{W}_\mu$ denote the set of all feasible portfolios with expected return μ. If $\mathcal{W}_\mu$ is not empty, the maximum CVaR of any portfolio in $\mathcal{W}_\mu$, at any level α, can always be achieved by a portfolio consisting of at most two assets.*

Proof. Let $X_1, \ldots, X_K$ be a set of assets with random returns $r_1, \ldots, r_K$. Assume that $\mathsf{E}(\lambda_1 r_1 + (1 - \lambda_1) r_2) = \mu$ for some $\lambda_1 \in [0, 1]$ and $\mathrm{CVaR}_\alpha(\lambda_1 L_{X_1} + \lambda_2 L_{X_2})$ is the highest among all two-asset portfolios with expected return μ (if no such portfolio existed, $\mathcal{W}_\mu$ would be empty). We will show that the addition of an asset cannot decrease CVaR of a two-asset portfolio while the expected return remains constant, therefore $\lambda_1 X_1 + \lambda_2 X_2$ maximizes CVaR. Since $\mathsf{E}(\lambda_1 r_1 + (1 - \lambda_1) r_2) = \mu$, either $r_1 > \mu$ and $r_2 < \mu$ (or the opposite), or at least one of r_1, r_2 is equal to μ. We analyze each option individually.

- $r_1 > \mu$ and $r_2 < \mu$:
  - For any $k = 3, \ldots, K$ such that $r_k = \mu$, based on the assumption, $\mathrm{CVaR}_\alpha(L_{X_k}) > \mathrm{CVaR}_\alpha(\lambda_1 L_{X_1} + (1 - \lambda_1) L_{X_2})$ must hold and X_k cannot increase CVaR.

– If there exists $k \in \{3, \dots, K\}$ such that $r_k \neq \mu$, then for some $\lambda_2 \in (0,1)$ either $\mathsf{E}(\lambda_2 r_1 + (1-\lambda_2) r_k) = \mu$ or $\mathsf{E}(\lambda_2 r_2 + (1-\lambda_2) r_k) = \mu$. Then

$$\begin{aligned}
&\mathsf{E}\Big(\beta_1(\lambda_1 r_1 + (1-\lambda_1) r_2) \\
&\quad + \mathbf{1}_{[\mathsf{E}(\lambda_2 r_1 + (1-\lambda_2) r_k) = \mu]}(1-\beta_1)(\lambda_2 r_1 + (1-\lambda_2) r_k) \\
&\quad + \mathbf{1}_{[\mathsf{E}(\lambda_2 r_2 + (1-\lambda_2) r_k) = \mu]}(1-\beta_1)(\lambda_2 r_2 + (1-\lambda_2) r_k)\Big) = \mu
\end{aligned}$$

for any $\beta_1 \in [0,1]$ and from the assumption of maximum CVaR pair,

$$\begin{aligned}
&(1-\beta_1)\,\mathrm{CVaR}_\alpha\Big(\lambda_1 L_{X_1} + (1-\lambda_1) L_{X_2}\Big) \\
&\qquad \geq \mathbf{1}_{[\mathsf{E}(\lambda_2 r_1 + (1-\lambda_2) r_k) = \mu]}(1-\beta_1)\,\mathrm{CVaR}_\alpha\Big(\lambda_2 L_{X_1} + (1-\lambda_2) L_{X_k}\Big) \\
&\qquad + \mathbf{1}_{[\mathsf{E}(\lambda_2 r_2 + (1-\lambda_2) r_k) = \mu]}(1-\beta_1)\,\mathrm{CVaR}_\alpha\Big(\lambda_2 L_{X_2} + (1-\lambda_2) L_{X_k}\Big).
\end{aligned}$$

If we add $\beta_1\,\mathrm{CVaR}_\alpha\Big(\lambda_1 L_{X_1} + (1-\lambda_1) L_{X_2}\Big)$ to both sides of the inequality, since β_1 was chosen arbitrarily, using Lemma A1 we get that any portfolio consisting of X_1, X_2, X_k has higher or equal CVaR than the original

- $r_1 = \mu, r_2 \neq \mu$: CVaR of any two-asset portfolio from $\mathcal{W}_\mu$ is higher or equal to CVaR of L_{X_1} and any portfolio from $\mathcal{W}_\mu$ gained by addition of an asset X_k to X_1 and X_2 is a convex combination of X_1 and a portfolio from $\mathcal{W}_\mu$ consisting of X_2 and X_k, which has higher CVaR than X_1 from the assumption.
- $r_1 = r_2 = \mu$: Any assets that can be added so that the new portfolio belongs to $\mathcal{W}_\mu$ must have also return μ and higher CVaR.

□

List of Figures

5.1 Relative price of the individual assets compared to the starting date. 29
5.2 Sample variances of individual assets (rolling average over past year). 29
5.3 Sample correlation matrix of daily stock returns. 29
5.4 Expected daily return and CVaR_{90} of the feasible and efficient set of the benchmark model. Artifacts in feasible set boundary are present due to computational instability, but overall feasible set depicted is precise. 31
5.5 Weights of efficient portfolio with targeted expected return. 31
5.6 Expected daily return and CVaR_{90} of efficient sets at different loss tolerance for $\mathrm{CVaR}_{99.5}$. Benchmark model portfolios and portfolios minimizing $\mathrm{CVaR}_{99.5}$ included. 35
5.7 Weights of efficient portfolio with targeted expected return and $\mathrm{CVaR}_{99.5} \leq 5.5\%$ denoted by full lines, in comparison with the benchmark model (dashed lines). 35
5.8 Expected daily return and CVaR_{90} of efficient sets at different levels of CVaR. CVaR_{80} and CVaR_{95} are virtually indistinguishable. 38
5.9 Expected daily return and CVaR_{90} for efficient set "boundary" of model (M3) with CVaR_{95} depicted using color scale. 38
5.10 Weights of efficient portfolios with expected return 0.085% from model (M3). Dashed line shows CVaR_{95} value of the portfolios. . 40
5.11 Weights of efficient portfolios with expected return 0.085% from model (M4). Dashed line shows CVaR_{95} value of the portfolios. Also shows positions of portfolios produced by benchmark model (M1) and extreme-tail model (M2). 40
5.12 Expected daily return and CVaR_{90} for efficient set of model (M4) with $\mathrm{CVaR}_{99.5}$ depicted using color scale. 41
5.13 Efficient set of model (M4) with $\mathrm{CVaR}_{99.5}$ emphasized using contour lines. 41
5.14 Expected daily return and CVaR_{90} for efficient set of model (M5) with L^1 distance depicted using color scale. 43
5.15 Expected daily return and L^1 distance form equal weights starting portfolio for the efficient set of model (M5) with CVaR_{90} emphasized using contour lines. 43
5.16 Daily returns with respect to the target CVaR_{90} value of portfolios with target returns 0.060%, 0.085%, 0.085%, in that order. Also shown is the ratio of absolute change in return and change in the target CVaR_{90} compared to the previous portfolio, to illustrate stability of solutions. 46

List of Tables

5.1 Selected S&P 500 companies. 26
5.2 Summary statistics of daily stock returns (in percent), CVaR_{90} of losses. 28
5.3 Performance, weights and L^1 distance from Markowitz portfolio with equivalent expected return for selected portfolios produce by the benchmark model (M1). 32
5.4 Portfolio weights of extreme-tail model (M2) for selected target returns with $\mathrm{CVaR}_{99.5}$ constraint, distance from benchmark model (M1) portfolios with equivalent expected return. 34
5.5 Return and risk measures of low-, medium- and high-risk portfolios produced by the benchmark model (M1), the extreme-tail model (M2) and the transaction cost model (M5). Last lines contains the L^1 distance from the equal weight starting portfolio. . . 45

List of Abbreviations and Notation

u	Utility function
U_1	Set of all utility functions
U_2	Set of all concave utility functions
$\boldsymbol{W}$	Level of wealth
$\mathsf{E}_X\, u(x)$	Expected utility of investment X
K	Number of assets
$\boldsymbol{r}$	Random vector of returns
$\boldsymbol{r}_i$	Observed realizations of $\boldsymbol{r}$
$\boldsymbol{R}$	Set of all random returns
$\boldsymbol{w}$	Decision vector, vector of portfolio weights
$\mathcal{W}$	Set off all feasible portfolio weights
$\mathcal{X}$	Set off feasible portfolio returns
L_X	Random loss associated with portfolio X
$(\Omega, \mathcal{F}, \mathsf{P})$	Probability space
FSD	First-order stochastic dominance
SSD	Second-order stochastic dominance
F_X	Cumulative distribution function of investment X
$F_X^{(-1)}$	quantile function of the distribution of investment X
$F_X^{(-2)}$	integrated quantile of the distribution of investment X
F_α	α-tail distribution function of investment X
VaR_α	Value at Risk at level α
VaR_α^+	Upper Value at Risk at level α
CVaR_α	Conditional Value at Risk at level α
ρ	Risk measure
$\tilde{w}$	Ideal solution
$\mathcal{W}_E$	Efficient set
$\mathrm{cl}\, A$	Closure of set A
$\mathbb{R}^p_\geq$	$\{\boldsymbol{r} \in \mathbb{R}^p : \boldsymbol{r} \geq 0\}$
$\mathbb{R}^p_+$	$\{\boldsymbol{r} \in \mathbb{R} : \boldsymbol{r} \geq 0, \boldsymbol{r} \neq 0\}$
$\mathrm{VaR}_\alpha^{ret}$	Value at Risk at level α for returns
$\mathrm{CVaR}_\alpha^{ret}$	Upper Value at Risk at level α for returns
$\mathcal{L}^1$	Set of all integrable functions
ϕ	Admissible risk spectrum
q_X^+	Upper quantile function of of the distribution of investment X
M_ϕ	Spectral risk measure
$\|\cdot\|_1$	L^1 norm