University of Milano - Bicocca

DEPARTMENT OF STATISTICS AND QUANTITATIVE METHODS

PhD in Statistics and Mathematical Finance

PORTFOLIO OPTIMIZATION WITH EXPECTILES

Advisor: Professor Fabio Bellini

PhD Thesis of:
Christian Colombo
Badge N: 721421

Academic Year 2016/2017

Contents

Chapter 1

Introduction

Portfolio selection is a classical topic in mathematical finance since the seminal work of Markowitz (1952), that introduced the classical mean/variance framework, later generalized to the so called mean/risk models. To this aim, several risk measures have been studied and compared from different point of views: their theoretical properties, their robustness, the complexity of the corresponding portfolio problem, their empirical performance on real data. There is a general agreement that a "good" risk measure should satisfy the axioms of coherence, introduced by Artzner et al. (1999). Acerbi and Tasche (2002a) even stated:

> "To avoid confusion, if a measure is not coherent we just choose not to call it a risk measure at all."

Computational issues are also extremely relevant from a practical point of view, since rebalancing a portfolio is an operation that in some cases is done even several times a day. For this reason, portfolio selection is preferable to be a Linear Programming Problem or at least a convex optimization problem, in order to assure the convergence of the algorithms used by the numerical solvers. Indeed if the problem is not convex, not only his computational cost may be extremely high, but also the solutions provided by the solver algorithms could be local minima and not global. For a recent review of the main LP solvable models for portfolio optimization we refer to Mansini et al. (2003a), Mansini et al. (2003b) Mansini et al. (2014).

A risk measure that is coherent and gives rise to an LP minimization problem when applied on discrete scenarios is the CVaR, as it has been shown by Rockafellar and Uryasev (2000). CVaR became extremely popular, especially since when

Basel Committee on Banking Supervision (2012) changed the standard regulatory risk measure from Value at Risk to CVaR.

Another family of risk measures that has become quite popular in the recent literature is the expectiles. Expectiles have been introduced in the statistical literature by Newey and Powell (1987) as an asymmetric generalization of the mean, and it has been recently shown by several authors that they are also coherent risk measures. Moreover, it has been proved (see Ziegel (2016), Delbaen et al. (2015), Bellini and Bignozzi (2015)) that they are the only coherent risk measure that satisfy the so-called elicitability property, that gives a natural scoring function for backtesting purposes.

This thesis explores the use of expectiles and of related risk measures in portfolio optimization. More precisely, we introduce the concept of Expectile Value at Risk (EVaR), and consider mean/risk models of the mean/EVaR type. We show that they also can be recast as LP problems, and we provide numerical examples both on real and simulated data, following the pattern that Bertsimas et al. (2004) used for an analogous study with CVaR.

Further, we consider portfolio optimization with the interexpectile difference as an objective, a novel variability measure introduced in Bellini et al. (2017) for measuring implied volatility from option prices, and compare it with the most well known interquantile difference.

A robust version of the problem of EVaR minimization is also provided, following the pattern of Zhu and Fukushima (2009).

Finally, starting from the work of Maillard et al. (2010) and all the following works on the equally-weighted risk contributions (see e.g. Cesarone and Colucci (2017) and Cesarone and Tardella (2017)), we investigate empirically equally-weighted risk contributions portfolios using the EVaR as risk measure.

The thesis is structured as follows.

In Chapter 2 we review the main theoretical concepts needed to introduce expectiles and EVaR. We start from the definition of monetary risk measure that can be found e.g. in Föllmer and Schied (2010) and we recall the definitions of convex and coherent risk measures from the axiomatic point of view or from their acceptance sets. We remind their most important properties and their Dual Representation. Moreover we talk about differentiability of risk measures due to its importance in the definition of equally-weighted risk contributions portfolios. Finally we recall the notion of Deviation Risk Measure.

In Chapter 3 we introduce expectiles as defined by Newey and Powell (1987) and we report the main results related to them (see e.g. Delbaen (2013), Bellini et al. (2014), Bellini and Bignozzi (2015)).

In Chapter 4 we recall the most important mean-risk models and we present the classical mean-risk problem using EVaR as risk measure to minimize. Starting from the dual representation of EVaR we show that this is an LP solvable problem. We also provide numerical examples using simulated and historical data.

In Chapter 5 we discuss about Interexpectile Difference as the analogous of the interquantile difference. We show that the relative mean-risk model optimization this is again an LP problem: we still perform numerical examples using simulated and historical data.

In Chapter 6 following the approach of Zhu and Fukushima (2009) we define a Robust version of EVaR, supposing that the random variable which describes the possible outcome of the portfolio belongs to a certain class of probability density functions. Then we perform tests using this risk measure.

Finally in Chapter 7 we recall the notion of equally-weighted risk contributions portfolio as presented in Maillard et al. (2010) and we extend this method to the case of EVaR. In the end we compare numerically equally-weighted risk contributions portfolios using different risk measures.

Chapter 2

Risk measures and deviation measures

⟨RM⟩

2.1 Monetary, Convex and Coherent Risk Measures

We recall the definition of monetary risk measures that can be found e.g. in Föllmer and Schied (2010) and Föllmer and Weber (2015). As usual, let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, and let $\mathcal{X}$ be a linear space of functions $X \colon \Omega \to \mathbb{R}$ containing the constants; in finance, for instance, Ω represents possible "states of the world" and X is a random variable which could be the value of an asset or a portfolio.

Definition 1. *A* monetary risk measure *is a map* $\rho \colon \mathcal{X} \to \mathbb{R} \cup \{+\infty\}$ *that satisfies the following properties:*

Normalization $\rho(0) = 0$

Translation invariance *if* $a \in \mathbb{R}$ *and* $X \in \mathcal{X}$, *then* $\rho(X + a) = \rho(X) - a$

Monotonicity *if* $X_1, X_2 \in \mathcal{X}$ *and* $X_1 \leqslant X_2$, *then* $\rho(X_2) \leqslant \rho(X_1)$.

A monetary risk measure ρ *is said to be a* convex risk measure *if it satisfies also:*

Convexity *for each* $X_1, X_2 \in \mathcal{X}$ *and* $\lambda \in [0, 1]$, *it holds that*

$$\rho(\lambda X_1 + (1 - \lambda) X_2) \leqslant \lambda \rho(X_1) + (1 - \lambda) \rho(X_2).$$

These properties have natural financial interpretations. The property of Normalization means that the risk of holding nothing is equal to 0. The property of Translation invariance, also called cash additivity, means that if we add an amount of cash a to a position X, then the risk decreases by a; in this case $\rho(X)$ may be interpreted as a capital requirement, i.e. the amount of cash to be added to a position X in order to make it acceptable from the point of view of a regulator. The Monotonicity property says that that if a position X_1 has less value than X_2, then the risk of holding it is higher. Finally Convexity represents the fact that the risk of a diversified portfolio should be less than the corresponding average of the risks of the single assets.

The first axiomatization of risk measures has been given in Artzner et al. (1999), that introduced the concept of *coherent* risk measure defining a set of desirable properties that a "good" risk measure should have. Coherent risk measures are monetary risk measures that satisfy also the subadditivity and positive homogeneity properties:

⟨coherentRM⟩ **Definition 2.** *A monetary risk measure is said to be* coherent *if it satisfies also the following properties:*

Subadditivity *for each $X_1, X_2 \in \mathcal{X}$, then $\rho(X_1+X_2) \leqslant \rho(X_1)+\rho(X_2)$*

Positive homogeneity *if $a \geqslant 0$ and $X \in \mathcal{X}$, then $\rho(aX) = a\rho(X)$.*

A coherent risk measure is always convex; more precisely, the link between subadditivity, convexity and positive homogeneity is given by the following (see e.g. Föllmer and Schied (2010)).

Proposition 3. *Let ρ be a risk measure, $\rho\colon \mathcal{X} \to \mathbb{R} \cup \{+\infty\}$, with $\rho(0) = 0$. Then:*

a) Subadditivity + Positive homogeneity $\Rightarrow$ Convexity;

b) Positive homogeneity + Convexity $\Rightarrow$ subadditivity;

c) Subadditivity + Convexity $\Rightarrow$ Positive homogeneity.

Proof. a) Subadditivity + Positive homogeneity $\Rightarrow$ Convexity

$$\rho(\lambda X + (1-\lambda)Y) \leqslant \rho(\lambda X) + \rho((1-\lambda)Y) = \lambda\rho(X) + (1-\lambda)\rho(Y)$$

The first inequality is true due to subadditivity, the second inequality is true from positive homogeneity.

b) Positive homogeneity + Convexity $\Rightarrow$ Subadditivity

$$\begin{aligned}\rho(X+Y) &= \rho\left(2\left(\frac{1}{2}X+\frac{1}{2}Y\right)\right) = 2\rho\left(\left(\frac{1}{2}X+\frac{1}{2}Y\right)\right)\\ &\leqslant 2\left(\frac{1}{2}\rho(X)+\frac{1}{2}\rho(Y)\right) = \rho(X)+\rho(Y)\end{aligned}$$

The second equality is true due to positive homogeneity and the inequality is true due to convexity.

c) Subadditivity + Convexity $\Rightarrow$ Positive homogeneity
Notice first that for $\alpha \leqslant 1$, from convexity and the fact that $\rho(0) = 0$, it follows that

$$\rho(\alpha X) = \rho(\alpha X + (1-\alpha)0) \leqslant \alpha\rho(X) + (1-\alpha)\rho(0) = \alpha\rho(X). \tag{2.1}$$

eqnsopra

Let now $\alpha \geqslant 1$ and let $\alpha = \lfloor\alpha\rfloor + (\alpha - \lfloor\alpha\rfloor)$, where $\lfloor\cdot\rfloor$ is the floor function. We have

$$\begin{aligned}\rho(\alpha X) &= \rho((\lfloor\alpha\rfloor + (\alpha-\lfloor\alpha\rfloor))X) \leqslant \rho(\lfloor\alpha\rfloor X) + \rho((\alpha-\lfloor\alpha\rfloor)X)\\ &\leqslant \lfloor\alpha\rfloor\rho(X) + (\alpha-\lfloor\alpha\rfloor)\rho(X) = \alpha\rho(X),\end{aligned} \tag{2.2}$$

eqnrhosub2

where the first inequality follows from subadditivity, the second inequality follows from the fact that by subadditivity $\rho(nX) \leqslant n\rho(X)$ and by (2.1), since $\alpha - \lfloor\alpha\rfloor \leqslant 1$. It follows that $\rho(\alpha X) \leqslant \alpha\rho(X)$, for each $\alpha \geqslant 0$.

In order to prove the opposite inequality notice that for each $\alpha \geqslant 0$

$$\rho(X) = \rho\left(\alpha\frac{1}{\alpha}X\right) \leqslant \frac{1}{\alpha}\rho(\alpha X) \quad\Rightarrow\quad \rho(\alpha X) \geqslant \alpha\rho(X) \tag{2.3}$$

eqnrhosub

Equations (2.2) and (2.3) together imply that $\rho(\alpha X) = \alpha\rho(X)$ for each $\alpha \geqslant 0$ that is the positive homogeneity property.

□

2.1.1 Acceptance sets

Given a monetary risk measure $\rho\colon \mathcal{X} \to \mathbb{R}$, its *acceptance set* $\mathcal{A}_\rho$ is given by:

$$\mathcal{A}_\rho := \{X \in \mathcal{X} | \rho(X) \leqslant 0\}.$$

The acceptance set of a risk measure represents the set of positions that have non-positive risk, and hence are termed acceptable. Similarly, given an acceptance set $\mathcal{A}$, we can associate to it a monetary risk measure $\rho_{\mathcal{A}}$ as follows:

$$\rho_{\mathcal{A}}(X) := \inf\{m \in \mathbb{R} | X + m \in \mathcal{A}\}.$$

The properties of ρ and $\mathcal{A}_\rho$ are strictly linked. Indeed:

⟨propArho⟩ **Proposition 4.** *Let ρ be a risk measure and let $\mathcal{A}_\rho$ be its acceptance set. Then:*

a) ρ is translation invariant if and only if $\rho = \rho_{\mathcal{A}_\rho}$.

b) If ρ is monotone then $\mathcal{A}_\rho$ is a monotone set, in the sense that

$$X \geqslant Y, \quad Y \in \mathcal{A} \Rightarrow X \in \mathcal{A}.$$

c) If ρ is positive homogeneous then $\mathcal{A}_\rho$ is a cone.

d) If ρ is convex then $\mathcal{A}_\rho$ is convex.

Proof. a) ρ is translation invariant if and only if $\rho = \rho_{\mathcal{A}_\rho}$.
By definition of $\mathcal{A}_\rho$, we have that:

$$\rho_{\mathcal{A}_\rho}(X) = \inf\{m \in \mathbb{R} | X + m \in \mathcal{A}_\rho\} = \inf\{m \in \mathbb{R} | \rho(X + m) \leq 0\}$$

By the property of translation invariance

$$\inf\{m \in \mathbb{R} | \rho(X + m) \leq 0\} = \inf\{m \in \mathbb{R} | \rho(X) \leqslant m\}$$

and clearly

$$\inf\{m \in \mathbb{R} | \rho(X) \leqslant m\} = \rho(X).$$

b) If ρ is monotone then $\mathcal{A}_\rho$ is a monotone set.
Let $X \geqslant Y$ and $Y \in \mathcal{A}_\rho$. By the monotonicity of ρ we have that $\rho(X) \leqslant \rho(Y)$. Since $\rho(Y) \leqslant 0$ because it belongs to $\mathcal{A}_\rho$, it follows that $\rho(X) \leqslant 0$ and $X \in \mathcal{A}_\rho$

c) If ρ is positive homogeneous then $\mathcal{A}_\rho$ is a cone.
We must prove that if $X \in \mathcal{A}_\rho$ and $\lambda \geqslant 0$ then $\lambda X \in \mathcal{A}_\rho$. By positive homogeneity we have that

$$\rho(\lambda X) = \lambda \rho(X) \leq 0$$

so $\lambda X \in \mathcal{A}_\rho$ and consequently $\mathcal{A}_\rho$ is a cone.

d) If ρ is a convex map then $\mathcal{A}_\rho$ is convex.
We must prove that if $X, Y \in \mathcal{A}_\rho$ and $\lambda \in [0,1]$ then $\lambda X + (1-\lambda)Y \in \mathcal{A}_\rho$. By the convexity of ρ we can write

$$\rho(\lambda X + (1-\lambda)Y) \leqslant \lambda\rho(X) + (1-\lambda)\rho(Y).$$

Since $X, Y \in \mathcal{A}_\rho$ and $\lambda \in [0,1]$ we have that

$$\lambda\rho(X) + (1-\lambda)\rho(Y) \leq 0$$

which ends the proof.

□

Similarly there is a relationship between the properties of the acceptance set $\mathcal{A}$ and the properties of the risk measure $\rho_{\mathcal{A}}$.

Proposition 5. *Let $\mathcal{A}$ be an acceptance set and let $\rho_{\mathcal{A}}$ be the corresponding risk measure, then*

a) $\rho_{\mathcal{A}}$ is translation invariant.

b) If $\mathcal{A}$ is a monotone set then $\rho_{\mathcal{A}}$ is a monotone.

c) If $\mathcal{A}$ is a cone then $\rho_{\mathcal{A}}$ is positive homogeneous .

d) If $\mathcal{A}$ is convex then $\rho_{\mathcal{A}}$ is a convex.

Proof. See e.g. Föllmer and Schied (2004). □

Proposition 4 says that if ρ is translation invariant, then $\rho = \rho_{\mathcal{A}_\rho}$. There is an analogous result for $\mathcal{A}_{\rho_{\mathcal{A}}}$. Recall that a set $\mathcal{A} \in \mathcal{X}$ is said to be *closed from above* if for each sequence $X_n \downarrow X$ with $X_n \in \mathcal{A}$, it holds that $X \in \mathcal{A}$. We have the following:

Proposition 6. *Let $\mathcal{A}$ be an acceptance set monotone and closed from above. Then $\mathcal{A} = \mathcal{A}_{\rho_{\mathcal{A}}}$.*

2.1.2 Dual Representations of Risk Measures

A fundamental result in the axiomatic theory of risk measures is the so called dual representation. Roughly, it states that every coherent risk measure is equal to a worst case expectation over a set of probability measures called *generalized scenarios*. We give a precise statement for risk measures on L^∞. In order to formulate it properly, we recall two very important continuity properties known as the *Fatou property* and the *Lebesgue property*. The Lebesgue property is a stronger continuity requirement than the Fatou property.

Definition 7. *A risk measure $\rho\colon L^\infty \to \mathbb{R}$ is said to have the Fatou property if*

$$\|X_n\|_\infty \leqslant k, \quad X_n \to X \ a.s. \Rightarrow \rho(X) \leqslant \liminf \rho(X_n).$$

It is said to have the Lebesgue property if

$$\|X_n\|_\infty \leqslant k, \quad X_n \to X \ a.s. \Rightarrow \lim_{n\to+\infty} \rho(X_n) = \rho(X).$$

Theorem 8. *Let $\rho\colon L^\infty \to \mathbb{R}$ be a coherent risk measure satisfying Fatou property. Then there exists a set of probability measures $\mathcal{M}$, with*

$$\mathcal{M} \subset \mathcal{Q} := \{\mathbb{Q} \ \textit{probability measures on} \ (\Omega, \mathcal{F}) | \mathbb{Q} \ll \mathbb{P}\},$$

such that

$$\rho(X) = \sup_{\mathbb{Q}\in\mathcal{M}} \{\mathbb{E}_\mathbb{Q}[-X]\}.$$

If moreover ρ satisfies the Lebesgue property, then

$$\rho(X) = \max_{\mathbb{Q}\in\mathcal{M}} \{\mathbb{E}_\mathbb{Q}[-X]\}.$$

Recall that $\mathbb{Q} \ll \mathbb{P}$ means that $\mathbb{Q}$ is absolutely continuous with respect to $\mathbb{P}$, i.e. that for every measurable set A, $\mathbb{P}(A) = 0$ implies $\mathbb{Q}(A) = 0$.

2.2 Value at Risk and Conditional Value at Risk

The most studied and discussed risk measures in the financial literature are the *Value at Risk* and the *Conditional Value at Risk*, briefly *VaR* and *CVaR*. There are several papers that compare these two risk measures from different points of views and explain in details their properties. Among them, we recall Linsmeier and Pearson (2000) who analyzed in detail the VaR and the methodologies to evaluate it, Rockafellar and Uryasev (2000) who showed how the problem of minimizing the CVaR of a portfolio on discrete scenarios can be recast as a Linear Programming problem, and Pflug (2000) who summarized the properties of these two risk measures and studied in details the structure of the portfolio optimization problem. In the following we recall their definitions.

Definition 9. *Let X be a random variable on $(\Omega, \mathcal{F}, \mathbb{P})$ with distribution function*

$$F_X(t) := \mathbb{P}(X \leqslant t).$$

For $\alpha \in (0,1)$, the left and right α-quantiles of X are defined respectively by:

$$q_\alpha^-(X) := \inf\{x \in \mathbb{R} \colon F_X(x) \geqslant \alpha\},$$
$$q_\alpha^+(X) := \sup\{x \in \mathbb{R} \colon F_X(x) \leqslant \alpha\}.$$

For continuous distribution functions, that are the most relevant in financial applications, it holds that $q_\alpha^-(X) = q_\alpha^+(X)$. Since the random variable X represent the value of a position, the loss tail corresponds to low values of α, typically $\alpha = 0.05$ or $\alpha = 0.01$. Value at Risk is usually defined as the negative of the right quantile:

Definition 10. *Let X be a random variable and let $\alpha \in (0,1)$. Then*

$$VaR_\alpha(X) := -q_\alpha^+(X).$$

It is easy to see that the acceptance set of VaR_α is the following:

$$\mathcal{A}_{VaR_\alpha} = \{X \in L^0(\Omega, \mathcal{F}, \mathbb{P}) | \mathbb{P}(X < 0) \leqslant \alpha\}.$$

A position is thus acceptable for VaR_α if the probability of a loss is less than a prespecified (low) level α. In other words, $VaR_\alpha(X)$ can be defined as the minimum amount of capital that has to be added to a position X in order to make the probability of a loss less than α. Indeed, we can check:

$$\begin{aligned}
&\inf\{m \in \mathbb{R} | X + m \in \mathcal{A}_{VaR_\alpha}\} = \inf\{m \in \mathbb{R} | \mathbb{P}(X + m < 0) \leqslant \alpha)\} = \\
&\inf\{m \in \mathbb{R} | \mathbb{P}(X < -m) \leqslant \alpha)\} = -\sup\{m \in \mathbb{R} | \mathbb{P}(X < m) \leqslant \alpha)\} = \\
&-\sup\{m \in \mathbb{R} | \mathbb{P}(X \leqslant m) \leqslant \alpha) = -q_\alpha^+(X).
\end{aligned}$$

From the properties of quantiles, it follows that VaR_α satisfies the aforementioned properties of Normalization, Translation invariance, Monotonicity and Positive homogeneity. As it is well known, the main drawback of VaR_α is that it does not satisfy the Subadditivity property and the Convexity property. That is one of the reasons why VaR is criticized in finance: diversification does not decrease its value, and may increase it.

The Conditional Value at Risk (CVaR), also known as Expected Shortfall or Tail Mean, intuitively represents a mean of the α worst case scenarios, with changed sign. It can be defined as follows:

Definition 11. *Let X be a L^1 random variable and let $\alpha \in (0,1]$. Then*

$$CVaR_\alpha(X) := -\frac{1}{\alpha}\int_0^\alpha q_\beta^-(X)\,d\beta = -\frac{1}{\alpha}\int_0^\alpha q_\beta^+(X)\,d\beta = \frac{1}{\alpha}\int_0^\alpha VaR_\gamma(X)\,d\gamma.$$

It is well known that CVaR satisfies the properties of Normalization, Translation invariance, Monotonicity, Positive homogeneity and Subadditivity, and hence is a coherent risk measure (see e.g. Pflug (2000) and Acerbi and Tasche (2002b)). Rockafellar and Uryasev (2000) showed that the CVaR can be computed as the optimal value of a simple minimization problem:

$$CVaR_\alpha(X) := \inf\left\{\alpha + \frac{1}{1-\alpha}\mathbb{E}[X-a]^+, \quad a \in \mathbb{R}\right\},$$

where $[x]^+ = \max(x,0)$.

2.3 Continuity and Differentiability of Convex Risk Measures

In this section we focus our attention on the continuity and differentiability properties of risk measures. We refers to Ruszczynski and Shapiro (2006), where they studied various notions of differentiability for convex risk functions.
Recall that the *domain* of ρ is defined as $\mathrm{dom}(\rho) := \{X \in \mathcal{X} | \rho(X) \le +\infty\}$ and that a risk measure ρ is said to be *proper* if $\rho(X) > -\infty$ for all $X \in \mathcal{X}$ and $\mathrm{dom}\,\rho \neq \emptyset$.

Proposition 12. *Let ρ be a proper and convex risk measure and* $\mathrm{int}(\mathrm{dom}\,\rho)$ *be the interior of the domain of ρ. If ρ is bounded from above on a neighborhood of some point $\overline{X} \in \mathcal{X}$, then ρ is continuous on* $\mathrm{int}(\mathrm{dom}\,\rho)$.

Definition 13. *A linear functional* $l : \mathcal{X} \to \mathbb{R}$ *is called an* algebraic subgradient *of* ρ *at* $\overline{X} \in \operatorname{dom} \rho$ *if*

$$\rho(X) \geqslant \rho(\overline{X}) + l(X - \overline{X}), \quad \forall X \in \mathcal{X}. \tag{2.4}$$ defnsub

The set of all subgradients l *which satisfy 2.4 is called the* subdifferential *of* ρ *at* $\overline{X}$*, and is denoted by* $\partial\rho(\overline{X})$.

Definition 14. *The* directional derivative function $\delta(\cdot) := \rho'(\overline{X}, \cdot)$ *is*

$$\rho'(\overline{X}, X) := \lim_{t \downarrow 0} \frac{\rho(\overline{X} + tX) - \rho(\overline{X})}{t}.$$

Proposition 15. *Suppose that* $\mathcal{X}$ *is a Banach lattice and* $\rho \colon \mathcal{X} \to \overline{\mathbb{R}}$ *is a proper convex risk measure, then* $\rho(\cdot)$ *is continuous and subdifferentiable on* $\operatorname{int}(\operatorname{dom} \rho)$.

If we consider the Banach space $\mathcal{X} := L^p(\Omega, \mathcal{F}, \mathbb{P})$, with $p \in [1, +\infty)$ and dual $\mathcal{Y} := L^q(\Omega, \mathcal{F}, \mathbb{P})$, with $\frac{1}{p} + \frac{1}{q} = 1$, we have that $\mathcal{X}$ is also a Banach lattice, so any proper convex risk measure $\rho \colon \mathcal{X} \to \overline{\mathbb{R}}$ is continuous and subdifferentiable on the interior of its domain. Moreover, if we consider a point $\overline{X} \in \operatorname{dom}(\rho)$, by the definition of subgradient we have that the probability measure $\mathbb{P}$ belongs to the subdifferential if and only if $\rho^\star(\mathbb{P}) = \langle \mathbb{P}, \overline{X} \rangle - \rho(\overline{X})$, where $\rho^\star \colon \mathcal{Y} \to \overline{\mathbb{R}}$ is the convex conjugate of ρ

$$\rho^\star(Y) = \sup_{X \in \mathcal{X}} \left\{ \langle Y, X \rangle - \rho(X) \right\}.$$

Since ρ satisfies the hypothesis of the *Fenchel-Moreau* theorem (see e.g. Borwein and Lewis (2006)), we have that $\rho = \rho^{\star\star}$ and

$$\partial\rho^{\star\star}(\overline{X}) = \arg \max_{\mathbb{P} \in \operatorname{dom}(\rho^\star)} \left\{ \langle \mathbb{P}, \overline{X} \rangle - \rho(\overline{X}) \right\}.$$

Hence, if ρ is a lower semicontinuous convex risk measure, then

$$\partial\rho(\overline{X}) = \partial\rho^{\star\star}(\overline{X}) = \arg \max_{\mathbb{P} \in \operatorname{dom}(\rho^\star)} \langle \mathbb{P}, \overline{X} \rangle.$$

Moreover, if ρ is subdifferentiable at $\overline{X}$ and $\rho'(\overline{X}, \cdot)$ is lower semicontinuous at 0, then

$$\rho'(\overline{X}, X) = \sup_{\mathbb{P} \in \partial\rho(\overline{X})} \langle \mathbb{P}, X \rangle, \quad X \in \mathcal{X}.$$

Recall the definition of *Hadamard directional differentiability* and the definition of *Gâteaux directional differentiability*:

Definition 16. *A function $\rho\colon L^p \to L^q$ is said to be* Hadamard directionally differentiable *at $\overline{X} \in L^p$ on the direction $X' \in L^P$ if there exist a function $\rho'(\overline{X}, \cdot)\colon L^p \to L^q$ such that*

$$\rho'(\overline{X}, X) = \lim_{t \downarrow 0,\ X' \to X} \frac{\rho(\overline{X} + tX') - \rho(\overline{X})}{t}.$$

for all sequences $X' \to X$.

Definition 17. *A function ρ is said to be* Gâteaux differentiable *at $\overline{X}$ if ρ is directionally differentiable at $\overline{X}$ and $\rho'(\overline{X}, \cdot)$ is linear continuous, i.e., there exists an element $\nabla\rho(\overline{X}) \in \mathcal{Y}$ such that*

$$\rho'(\overline{X}, \cdot) = \langle \nabla\rho(\overline{X}), X \rangle, \quad \forall X \in \mathcal{X}.$$

These two definitions of differentiability are related, in fact if the Hadamard directional derivative exists, then the Gâteaux derivative exists and the two derivatives coincide (see e.g. Shapiro (1991)). Hence, Hadamard directional differentiability induces continuity of the directional derivative function $\rho'(\overline{X}, \cdot)$.
It follows that if $\mathcal{X}$ is a Banach space and ρ is continuous at $\overline{X}$, then ρ is Gâteaux differentiable at $\overline{X}$ if and only if $\partial\rho(\overline{X})$ is a singleton, and in this case we have that $\partial\rho(\overline{X}) = \nabla\rho(\overline{X})$.
For further clarification we refer to Ruszczynski and Shapiro (2006) and Rockafellar (1974).

Following Ruszczynski and Shapiro (2006), we now demonstrate that the Conditional Value at Risk is Hadamard differentiable.
Consider the space $\mathcal{X} = L_1(\Omega, \mathcal{F}, \mathbb{P})$ and its dual space $\mathcal{Y} = L_\infty(\Omega, \mathcal{F}, \mathbb{P})$, the dual representation of $CVaR_\alpha$ (see e.g. Föllmer and Schied (2004)) is

$$CVaR_\alpha(X) = \sup_{Y \in \mathcal{U}_\alpha} \langle Y, X \rangle \tag{2.5}$$ eqCVAR

where

$$\mathcal{U}_\alpha := \left\{ Y \in \mathcal{Y} \text{ s.t. } \mathbb{E}[Y] = 1,\ Y \in [0, \alpha^{-1}],\ \mathbb{P}\text{-a.e.} \right\}.$$

Then the subdifferential is defined as

$$\begin{aligned} \partial CVaR_\alpha(X) &= \operatorname*{argmax}_{Y \in \mathcal{U}_\alpha} \langle Y, X \rangle \\ &= \operatorname*{argmax}_{Y \in \mathcal{Y}} \left\{ \langle Y, X \rangle \text{ s.t. } \mathbb{E}[Y] = 1,\ Y \in [0, \alpha^{-1}],\ \text{a.s.} \right\}. \end{aligned}$$

Relaxing the constraint $\mathbb{E}[Y] = 1$, the Lagrangean of the problem (2.5) is:

$$\begin{aligned} L(Y, \lambda; X) &= \langle Y, X \rangle + \lambda(1 - \mathbb{E}[Y]) \\ &= \langle Y, X \rangle + \lambda - \langle \lambda, Y \rangle \\ &= \langle Y, X - \lambda \rangle + \lambda \end{aligned}$$

Introducing the dual function of the Lagrangean $d(\lambda)$ we can write

$$\begin{aligned} d(\lambda) &= \sup_{Y \in [0, \alpha^{-1}]} L(Y, \lambda) \\ &= \sup_{Y \in [0, \alpha^{-1}]} \langle Y, X - \lambda \rangle + \lambda \\ &= \sup_{Y \in [0, \alpha^{-1}]} \int_\Omega Y(X - \lambda) d\mathbb{P} + \lambda, \end{aligned}$$

Clearly the supremum is attained for $Y = \alpha^{-1}\mathbb{1}_{[X - \lambda \geqslant 0]}$, and in this case

$$d(\lambda) = \alpha^{-1}\mathbb{E}[X - \lambda]_+ + \lambda.$$

The dual problem of 2.5 is found by minimizing the dual function, hence

$$\inf_{\lambda \in \mathbb{R}} \alpha^{-1}\mathbb{E}[X - \lambda]_+ + \lambda.$$

Since the set of minimizers is a bounded set, we have strong duality, so

$$CVaR_\alpha = \inf_{\lambda \in \mathbb{R}} \alpha^{-1}\mathbb{E}[X - \lambda]_+ + \lambda.$$

If we define $\overline{\lambda}$ as

$$\overline{\lambda} = \inf\{\lambda \text{ s.t. } F_X(\lambda) \geq 1 - \alpha\},$$

where F_X is the cumulative density function of the random variable X, we have that $Y \in \partial CVaR_\alpha$ if and only if

$$\begin{aligned} \mathbb{E}[Y] &= 1 \\ X > \overline{\lambda} &\implies Y = \alpha^{-1} \\ X < \overline{\lambda} &\implies Y = 0 \\ X = \overline{\lambda} &\implies Y \in [0, \alpha^{-1}]. \end{aligned} \tag{2.6}$$

system

It follows that $\partial CVaR_\alpha$ is a singleton, which implies that $CVaR_\alpha$ is Hadamard differentiable, if and only if (2.6) has a unique solution.

So we have that $Y = \partial CVaR_\alpha$ when one of the following statement is true:

$$\begin{aligned} \mathbb{P}(X = \overline{\lambda}) &= 0 \\ \mathbb{P}(X < \overline{\lambda}) &= \alpha \\ \mathbb{P}(X > \overline{\lambda}) &= 1 - \alpha, \end{aligned}$$

and in this case

$$CVaR'_\alpha(\overline{X}) = \langle \partial CVaR_\alpha(X), \overline{X} \rangle.$$

2.4 Deviation Risk Measures

As we have seen in the previous sections, coherent risk measures may be interpreted as minimum capital requirements, i.e. they are the minimum amount of capital that has to be added to a position in order to make it acceptable. The ordinary variability measures of the statistical literature such as the variance and the standard deviation do not belong to this class, because they do not satisfy the Translation invariance property. Rockafellar et al. (2006) proposed an axiomatic definition of *deviation risk measures* by replacing the axiom of Translation invariance with a natural axiom that they called Shift invariance.

Definition 18. *Let $\mathcal{X} \subseteq L^0(\Omega, \mathcal{F}, \mathbb{P})$. A map $D\colon \mathcal{X} \to [0, +\infty]$ is a* deviation risk measure *if satisfies the following properties:*

Shift-invariance *If $X \in \mathcal{X}$ and $c \in \mathbb{R}$, then $D(X + c) = D(X)$*

Positive homogeneity *If $X \in \mathcal{X}$ and $\lambda > 0$, then $D(\lambda X) = \lambda D(X)$*

Subadditivity *If $X_1, X_2 \in \mathcal{X}$, then $D(X_1 + X_2) \leqslant D(X_1) + D(X_2)$*

Positivity *If $X \in \mathcal{X}$, then $D(X) \geqslant 0$ and $D(X) = 0$ if and only if $X = c$ a.s.*

Standard deviation is an example of a deviation risk measure, while the interquantile differences is not because it does not satisfy the subadditivity. Other important examples of deviation risk measure are:

- Mean absolute deviation, defined by $MAD(X) := \mathbb{E}[|X - \mathbb{E}[X]|]$

- Range-based deviations, given by

$$D_-(X) := \mathbb{E}[X] - \inf(X) \text{ or } D_+(X) := \sup(X) + \mathbb{E}[X]$$

- CVaR deviation, defined as $CVaR_\alpha^\Delta := CVaR_\alpha(X - \mathbb{E}[X])$.

More generally, Rockafellar et al. (2006) noticed that if ρ is a coherent risk measure, then the quantity

$$D(X) := \rho(X) + \mathbb{E}[X] = \rho(X - \mathbb{E}[X])$$

is a deviation risk measure.

Chapter 3

Expectiles and EVaR

⟨EVaR⟩ Expectiles have been first introduced in the statistical literature by Newey and Powell (1987), as an asymmetric version of the least squares regression. For $\tau \in (0,1)$, they introduced the loss function (see Fig. 3.1)

$$\rho_\tau(x) = \tau x_+^2 + (1-\tau)x_-^2,$$

where $x_+ = \max(x,0)$ and $x_- = (-x)_+$. For $X \in L^2$, they defined the expectiles $e_\tau(X)$ as follows:

$$e_\tau(X) = \operatorname*{argmin}_{x\in\mathbb{R}} \mathbb{E}[\rho_\tau(X-x)]$$

that can be explicitly written as

$$e_\tau(X) = \operatorname*{argmin}_{x\in\mathbb{R}} \left\{\tau\mathbb{E}\left[(X-x)_+^2\right] + (1-\tau)\mathbb{E}\left[(X-x)_-^2\right]\right\}. \tag{3.1}$$ defexpectile

Expectiles are thus a straightforward one-parameter asymmetric generalization of the mean, that arises when $\tau = 1/2$, since obviously

$$\begin{aligned} e_{1/2}(X) &= \operatorname*{argmin}_{x\in\mathbb{R}} \left\{1/2\,\mathbb{E}\left[(X-x)_+^2\right] + 1/2\,\mathbb{E}\left[(X-x)_-^2\right]\right\} \\ &= \operatorname*{argmin}_{x\in\mathbb{R}} \mathbb{E}\left[(X-x)^2\right] = \mathbb{E}[X]. \end{aligned}$$

If $\tau > 1/2$ the lose function gives more weights to the right tail of the distribution and consequently $e_\tau(X) > \mathbb{E}[X]$, the opposit situation arises if $\tau < 1/2$.
The present definition has the drawback of being well-posed only for square-integrable random variables; in order to define expectiles for each $X \in L^1$, Newey and Powell (1987) adopted the following slightly modified definition:

$$e_\tau(X) = \operatorname*{argmin}_{x\in\mathbb{R}} \mathbb{E}[\rho_\tau(X-x) - \rho_\tau(X)].$$

Notice that the additional term $\rho_\tau(X)$ does not contain x and has the only effect of making the objective function in the definition finite for each $X \in L^1$.

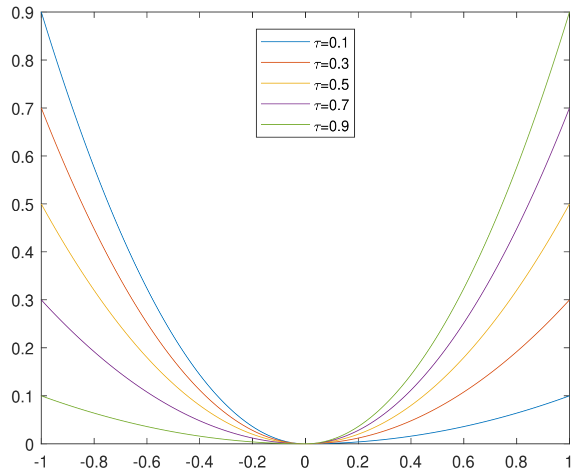


Figure 3.1: Expectile loss functions ρ_τ for different values of τ.

unctionExpectile⟩

3.1 Properties of Expectiles

Newey and Powell (1987) proved the following properties:

Theorem 19. *Let be $X \in L^1$ and $\tau \in (0,1)$. Then:*

a) $e_\tau(X)$ is the solution of the equation

$$x - \mathbb{E}[X] = \frac{2\tau - 1}{1-\tau} \int_x^\infty (y-x) dF_X(y); \qquad (3.2)$$ foc0

a) as a function of τ, $e_\tau \colon (0,1) \to \mathbb{R}$ is strictly increasing;

a) if $E = (essinf\, X, esssup\, X)$ then $e_\tau(X) \colon (0,1) \to E$;

a) if $\widetilde{X} = sX + t$, with $s > 0$, then $e_\tau(\widetilde{X}) = se_\tau(X) + t$;

a) if F_X is continuously differentiable, then e_τ is continuously differentiable, and for $y \neq \mathbb{E}[X]$ and $\overline{\tau}$ such that $y = e_{\overline{\tau}(X)}$,

$$F_X(y) = \lim_{y \to \mathbb{E}[X]} -\frac{y - \mathbb{E}[X] + \overline{\tau}(1-2\overline{\tau}) \cdot \left.\frac{de_\tau(X)}{d\tau}\right|_{\tau=\overline{\tau}}}{(1-2\overline{\tau})^2 \cdot \left.\frac{de_\tau(X)}{d\tau}\right|_{\tau=\overline{\tau}}}$$

From property d) we have that expectiles are location and scale invariant. Item a) shows that expectiles may equivalently be defined by means of the first order condition (3.2), that adimits several equivalent formulations. For the sake of completeness we provide a direct proof of a).

Let $X \in L^1$ and let

$$f(X,x) := \tau\mathbb{E}\left[(X-x)_+^2\right] + (1-\tau)\mathbb{E}\left[(X-x)_-^2\right].$$

The function f is convex in x, continuous and differentiable, as a consequence of the dominated convergence theorem. By definition, the expectile is given by

$$e_\tau(X) = \underset{x\in\mathbb{R}}{\operatorname{argmin}}\left\{\tau\mathbb{E}\left[(X-x)_+^2\right] + (1-\tau)\mathbb{E}\left[(X-x)_-^2\right]\right\}$$

and it must satisfy

$$\left.\frac{df(X,x)}{dx}\right|_{x=e_\tau(X)} = 0.$$

Hence

$$\left.\frac{d}{dx}\left(\tau\mathbb{E}\left[(X-x)_+^2\right] + (1-\tau)\mathbb{E}\left[(X-x)_-^2\right]\right)\right|_{x=e_\tau(X)} = 0,$$

and interchanging the derivative with the expectation by the dominated convergence theorem we get

$$\left.\frac{d}{dx}\left(\tau\mathbb{E}\left[(X-x)_+^2\right]\right)\right|_{x=e_\tau(X)} + \left.\frac{d}{dx}\left((1-\tau)\mathbb{E}\left[(X-x)_-^2\right]\right)\right|_{x=e_\tau(X)} = 0,$$

$$\tau\mathbb{E}\left[\left.\frac{d}{dx}\left((X-x)_+^2\right)\right|_{x=e_\tau(X)}\right] + (1-\tau)\mathbb{E}\left[\left.\frac{d}{dx}\left((X-x)_-^2\right)\right|_{x=e_\tau(X)}\right] = 0,$$

that gives the first order condition in the simplest form

$$\tau\mathbb{E}\left[(X-e_\tau(X))_+\right] = (1-\tau)\mathbb{E}\left[(X-e_\tau(X))_-\right]. \tag{3.3}$$ foc

Notice that Equation (3.3) has always exactly one solution; it is actually an alternative definition of e_τ. For $\tau = 1/2$, we recover the definition of the mean as the value for which the expectation of the right excess given by $\mathbb{E}[(X-\mathbb{E}[X])_+]$ is equal to the expectation of the left excess, given by $\mathbb{E}[(X-\mathbb{E}[X])_-]$. Rewriting the first order condition (3.3) as follows

$$\frac{\mathbb{E}\left[(X-e_\tau(X))_+\right]}{\mathbb{E}\left[(X-e_\tau(X))_-\right]} = \frac{1-\tau}{\tau},$$

we see that the expectile e_τ is that quantity that sets the *gain-loss ratio* in the left hand side equal to the prespecified constant $(1-\tau)/\tau$.

Another interesting interpretation of expectiles can be seen by considering the following piecewise linear utility function:

$$u_\tau(x) = \begin{cases} \tau x & \text{if } \ x \geq 0 \\ (1-\tau)x & \text{if } \ x < 0 \end{cases}.$$

Notice that u_τ is strictly increasing, concave for $\tau < 1/2$ and convex for $\tau > 1/2$. It is easy to check that e_τ satisfies the equation

$$\mathbb{E}[u_\tau(X - e_\tau(X))] = 0.$$

Indeed,

$$\begin{aligned} \mathbb{E}[u_\tau(X - e_\tau(X))] &= \mathbb{E}[\tau(X - e_\tau(X))_+ - (1-\tau)(X - e_\tau(X))_-] = \\ &= \tau\mathbb{E}\left[(X - e_\tau(X))_+\right] - (1-\tau)\mathbb{E}\left[(X - e_\tau(X))_-\right] = \\ &= 0 \end{aligned}$$

by the first order condition. So $e_\tau(X)$ can be seen as a *zero utility premium* or a *shortfall risk measure* with respect to the utility function u_τ; in literature the theory of zero utility premia have been studied for instance in Rolski et al. (2008)).

Expectiles are thus strictly related to the piecewise linear expected utility. However, these objects should not be confused. The following example shows a situation in which the two criteria give opposite orderings.

Example 20. *Let consider the following lotteries*

$$A = \begin{cases} 1 & \textit{with probability} \quad 0.5 \\ 3 & \textit{with probability} \quad 0.5 \end{cases} \quad \textit{and} \quad B = \begin{cases} 0 & \textit{with probability} \quad 0.05 \\ 2 & \textit{with probability} \quad 0.95 \end{cases}$$

Then considering the first order condition

$$\begin{aligned} 0.25 \cdot 0.5 \cdot (3 - e_{0.25}(A)) &= 0.75 \cdot 0.5 \cdot (e_{0.25}(A) - 1) \quad &\Rightarrow \quad e_{0.25}(A) = 1.5 \\ 0.25 \cdot 0.95 \cdot (2 - e_{0.25}(B)) &= 0.75 \cdot 0.05 \cdot (e_{0.25}(B)) \quad &\Rightarrow \quad e_{0.25}(B) = 1.72 \end{aligned}$$

So the expected utility is

$$\begin{aligned} U_\tau(A) &= \frac{0.5 \cdot 0.25 \cdot 3 + 0.5 \cdot 0.75 \cdot 1}{2} = 0.38 \\ U_\tau(B) &= \frac{0.95 \cdot 0.25 \cdot 2 + 0.05 \cdot 0.75 \cdot 0}{2} = 0.23 \end{aligned}$$

So the expectile is higher for the lottery B, but the expected utility of A is higher: optimizing expectiles and expected utility provides different results. The same conclusion can be achieved by considering the gain-loss ratio: take for instance the following lotteries

$$C = \begin{cases} -1 & \textit{with probability} \quad 0.3 \\ 1 & \textit{with probability} \quad 0.7 \end{cases} \quad \textit{and} \quad D = \begin{cases} -1 & \textit{with probability} \quad 0.5 \\ 2 & \textit{with probability} \quad 0.5 \end{cases}$$

again the lottery chosen by minimizing the gain loss ratio is different from the lottery chosen by maximizing the expectile.

It is worth noting the analogy between expectiles and the usual quantiles, that may be equivalently defined by means of the following optimization problem:

$$[q_\alpha^-(X), q_\alpha^+(X)] = \underset{x \in \mathbb{R}}{\operatorname{argmin}} \left\{ \alpha \mathbb{E}\left[(X-x)_+\right] + (1-\alpha)\mathbb{E}\left[(X-x)_-\right] \right\}.$$

Both expectiles and the usual quantiles are special cases of the so called generalized quantiles or M-quantiles (see e.g. Breckling and Chambers (1988)), defined as follows:

Definition 21. *Let Φ_1 and Φ_2 be convex loss functions with $\Phi_i(0) = 0$ and let $X \in L^\infty$. The* generalized quantile $\widetilde{q}_\tau$ *is*

$$\widetilde{q}_\tau(X) = \underset{x \in \mathbb{R}}{\operatorname{argmin}} \left\{ \tau \mathbb{E}\left[\Phi_1((X-x)_+)\right] + (1-\tau)\mathbb{E}\left[\Phi_2((X-x)_-)\right] \right\}.$$

The expectile is therefore the *generalized quantile* with $\Phi_1(x) = \Phi_2(x) = x^2$.

In Bellini and Di Bernardino (2017) there is a comparison between expectiles and quantiles curves for the most common distribution: in Fig. 3.2 we report the cases of the most common distributions. In general the expectiles are closer to the center of the distribution than the corresponding quantiles and the two curves typically intersect in a single point.

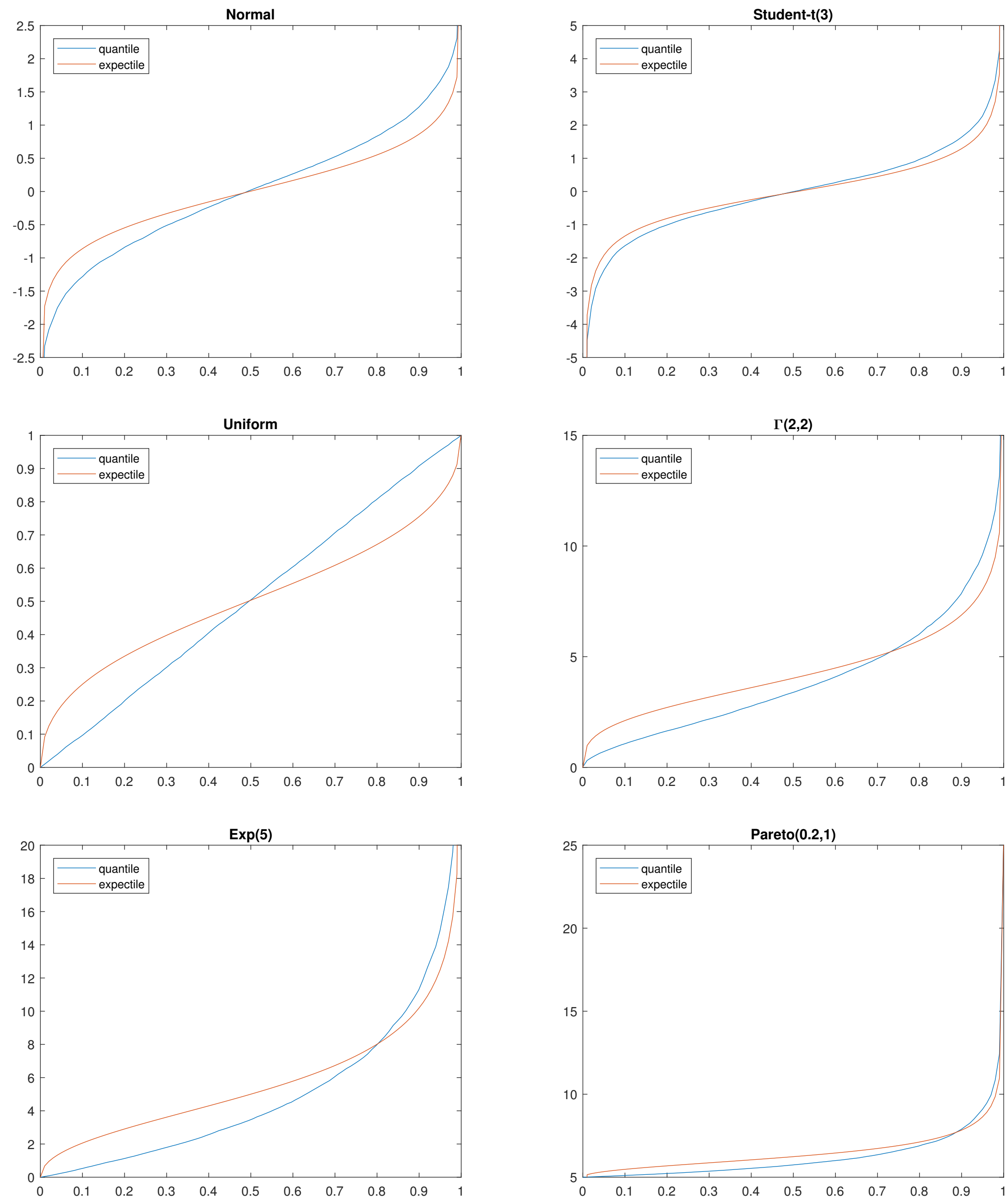


Figure 3.2: Comparison of the quantiles and expectiles curves for different distributions.

⟨exp_quant⟩

Further properties of the expectiles (see e.g. Bellini et al. (2014)) are recalled in the next Proposition.

Proposition 22. *Let X, Y be random variables in L^1 and let $\tau \in (0,1)$*

a) If $X \geqslant_{FSD} Y$ then $e_\tau(X) \geqslant e_\tau(Y)$; if moreover $\mathbb{P}(X > Y) > 0$, then $e_\tau(X) > e_\tau(Y)$

a) *If* $\tau \leqslant \frac{1}{2}$, *then* $e_\tau(X+Y) \geqslant e_\tau(X) + e_\tau(Y)$; *if* $\tau \geqslant \frac{1}{2}$, *then* $e_\tau(X+Y) \leqslant e_\tau(X) + e_\tau(Y)$

a) $e_\tau(X) = -e_{1-\tau}(-X)$.

Property a) shows that the expectiles are strictly monotonic with respect to the first order stochastic dominance $\leqslant_{FSD}$; property b) is extremely relevant from a financial point of view, since it shows that for $\tau \geqslant 1/2$ expectiles have the fundamental subadditivity property; finally, property c) is a symmetry property. Notice that neither of the three properties is satisfied by usual quantiles.

Moreover, as pointed out by many authors, the expectile is the unique risk measure that is coherent and elicitable (see e.g. Gneiting (2011), Ziegel (2016), Bellini and Bignozzi (2015), Delbaen et al. (2015)).
For the sake of completeness we report the definition of scoring function and elicitable functional.

Definition 23. *A function* $S : \mathbb{R}^2 \to [0, +\infty)$ *is said to be* scoring function *if for every* $x, y \in \mathbb{R}$*:*

a) $S(x, y) \geq 0$ *and* $S(x, y) = 0$ *if and only if* $x = y$

a) $S(x, y)$ *is increasing in* x *for* $x > y$ *and decreasing for* $x < y$

a) $S(x, y)$ *is continuous in* x

Definition 24. *Let* $T : \mathcal{M} \to 2^{\mathbb{R}}$ *be a possibly set-valued functional, where* $2^{\mathbb{R}}$ *is the power set of* $\mathbb{R}$. T *is* elicitable *relatively to the class* $\mathcal{M}$ *if there exist a scoring function* $S : \mathbb{R}^2 \to [0, +\infty)$ *such that*

$$\int S(x, y) dF(y) < +\infty \quad \forall x \in \mathbb{R}, \ \forall F \in \mathcal{M}$$
$$T(F) = \operatorname{argmin}_x \int S(x, y) dF(y), \quad \forall F \in \mathcal{M}.$$

In this case we say that the scoring function S *is strictly consistent with* T.

The elicitablity of expectiles derives directly from their definition, with the consistent scoring function

$$S(x, y) = \rho_\tau(y - x) = \tau(y - x)_+^2 + (1 - \tau)(y - x)_-^2.$$

See Bellini et al. (2017) for examples and applications of scoring functions for backtesting.

3.2 Expectile Value at Risk

Recall from the previous sections that VaR_α is defined by

$$VaR_\alpha(X) = -q_\alpha^+(X).$$

By paralleling this definition, we introduce the Expectile Value at Risk (EVaR).

Definition 25. *Let X be a random variable in L^1 and let $\tau \in (0, 1/2]$. We define*

$$EVaR_\tau(X) = -e_\tau(X).$$

The requirement $\tau \leqslant 1/2$ is motivated by the fact that in this case e_τ is superadditive and hence, as a consequence, $EVaR_\tau(X)$ is subadditive. Notice also that from the symmetry property of expectiles described in the previous section it holds that

$$EVaR_\tau(X) = -e_\tau(X) = e_{1-\tau}(-X),$$

so EVaR may be equivalently defined as an high expectile of the losses.

From the properties of expectiles it follows immediately that, for $\tau \leqslant 1/2$, $EVaR_\tau(X)$ is a coherent risk measure.

Proposition 26. *Let X and Y be random variables in L^1, and let $\tau \in (0, 1/2]$. Then*

Translation Invariance $EVaR_\tau(X + a) = EVaR_\tau(X) - a, \ \forall a \in \mathbb{R}$

Positive Homogeneity $EVaR_\tau(\lambda X) = \lambda EVaR_\tau(X), \ \forall \lambda > 0$

Monotonicity *if $X \leqslant Y$ almost surely, then $EVaR_\tau(X) \geqslant EVaR_\tau(Y)$*

Subadditivity $EVaR_\tau(X + Y) \leqslant EVaR_\tau(X) + EVaR_\tau(Y)$.

In order to better understand the financial meaning of $EVaR_\tau$, we study its acceptance set, defined as usual by

$$\mathcal{A}_{EVaR_\tau} = \{X \in L^1 | EVaR_\tau(X) \leqslant 0\} = \{X \in L^1 | e_\tau(X) \geqslant 0\}.$$

From the first order condition it follows immediately that

$$e_\tau(X) \geqslant 0 \iff \tau\mathbb{E}[X_+] - (1 - \tau)\mathbb{E}[X_-] \geqslant 0,$$

that implies that

$$X \in \mathcal{A}_{EVaR_\tau} \iff \frac{\mathbb{E}[X_+]}{\mathbb{E}[X_-]} \geqslant \frac{1-\tau}{\tau}.$$

A position is thus acceptable for $EVaR_\tau$ if and only if its gain loss ratio is bigger than a prespecified value $(1-\tau)/\tau$, that depends on the level τ. Notice that since $\tau \leqslant 1/2$ it follows that $(1-\tau)/\tau \geqslant 1$.

3.2.1 Dual Representation of EVaR and CVaR

The dual representation of $EVaR_\tau$ has been derived in Bellini et al. (2014).

Proposition 27. *Let $X \in L^1$, $\tau \in (0, 1/2]$ and $EVaR_\tau(X)$. Then*

$$EVaR_\tau(X) = \max_{\varphi \in \mathcal{M}_\beta} E[-X\varphi], \tag{3.4}$$ DualEVaR

where $\beta = (1-\tau)/\tau$ and

$$\mathcal{M}_\beta = \left\{ \varphi \in L^\infty(\Omega, F, P),\ \varphi \geqslant 0,\ \mathbb{E}[\varphi] = 1,\ \frac{\operatorname{ess\,sup}(\varphi)}{\operatorname{ess\,inf}(\varphi)} \leqslant \beta \right\}.$$

It is interesting to compare the set of probability measures $\mathcal{M}_\beta$ with the corresponding sets of $CVaR_\alpha$, in fact his dual representation is

$$CVaR_\alpha(X) = \max_{\varphi \in \mathcal{U}_\alpha} E[-X\varphi],$$

where

$$\mathcal{U}_\alpha := \left\{ \varphi \in L^\infty \text{ s.t. } \mathbb{E}[\varphi] = 1,\ \varphi \in [0, \alpha^{-1}],\ \mathbb{P}\text{-a.e.} \right\}.$$

While the densities $\varphi \in \mathcal{U}_\alpha$ are bounded by the constraint $\varphi \in [0, \alpha^{-1}]$, in the case of $EVaR_\tau$ we have a constraint that depends on the ratio between the essential supremum and the essential infimum of φ.
Moreover, the constraint $\mathbb{E}[\varphi] = 1$ implies that

$$\frac{1}{\beta} \leqslant \operatorname{ess\,inf}(\varphi) \leqslant 1 \leqslant \operatorname{ess\,sup}(\varphi) \leqslant \beta.$$

Another interesting comparison can be done with the *Bilateral Conditional Value at Risk* of levels γ, δ introduced by Pflug and Ruszczynski (2004).

Definition 28. *Let X be a L^1 random variable and let $0 \leqslant \gamma \leqslant 1 \leqslant \delta$. Then*

$$BCVaR_{\gamma,\delta}(X) := \max_{\varphi \in \mathcal{M}_{\gamma,\delta}} \mathbb{E}[-X\varphi]$$

with

$$\mathcal{M}_{\gamma,\delta} := \{\varphi \in L^\infty \ s.t.\ \mathbb{E}[\varphi] = 1,\ \varphi \in [\gamma, \delta],\ \mathbb{P}\text{-}a.e.\}.$$

In Fig. 3.3 are represented the constraints of the sets $\mathcal{U}_\alpha$, $\mathcal{M}_{\gamma,\delta}$ and $\mathcal{M}_\beta$: clearly we have that

$$\mathcal{M}_\beta = \bigcup_{\gamma \in [1/\beta, 1]} \mathcal{M}_{\gamma,\beta\gamma},$$

thus

$$EVaR_\tau(X) = \max_{\varphi \in \bigcup\limits_{\gamma \in [1/\beta,1]} \mathcal{M}_{\gamma,\beta\gamma}} \mathbb{E}[-X\varphi] = \max_{\gamma \in [1/\beta,1]} BCVaR_{\gamma,\beta\gamma}.$$

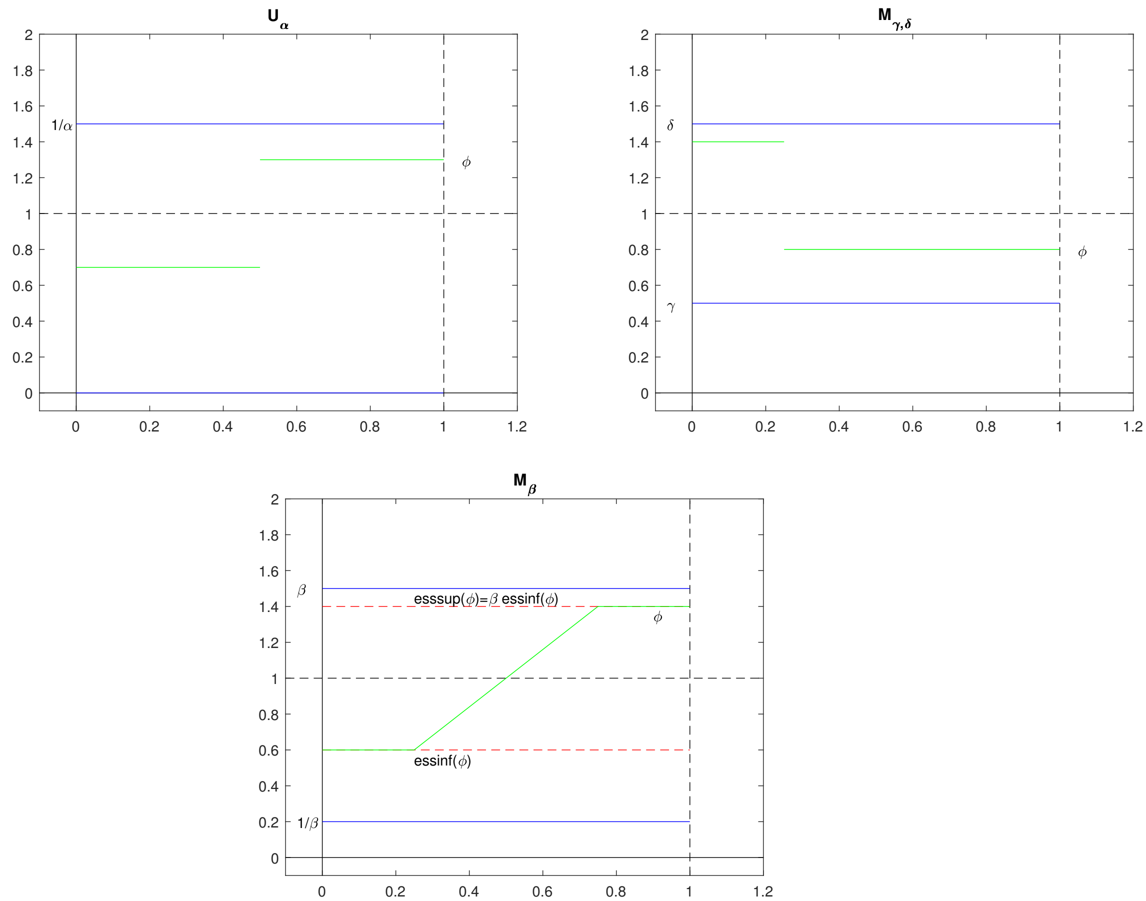


Figure 3.3: Representation of the domain of the sets of probability measures of the CVaR, BCVaR and EVaR

⟨setsprob⟩

In their work, Pflug and Ruszczynski (2004) showed also that

$$BCVaR_{\gamma,\delta}(X) = (1-\gamma)CVaR_{\frac{1-\gamma}{\delta-\gamma}}(X) + \gamma\mathbb{E}[-X],$$

and consequently,

$$BCVaR_{\gamma,\beta\gamma}(X) = (1-\gamma)CVaR_{\frac{1-\gamma}{\beta\gamma-\gamma}}(X) + \gamma\mathbb{E}[-X],$$

finally, substituting $z = \frac{1-\gamma}{\gamma(\beta-1)}$, we have that

$$EVaR_\tau(X) = \max_{z\in[0,1]} \left\{ \frac{z(\beta-1)}{1+z(\beta-1)} CVaR_z(X) + \frac{1}{1+z(\beta-1)} \mathbb{E}[-X] \right\}.$$

In conclusion we think that EVaR deserves to be discussed in detail because of its extremely interesting theoretical properties: indeed EVaR is a coherent and elicitable risk measure which take into account the whole distribution and not only the left tail. Comparing EVaR with the most common risk measures we notice that: CVaR is a coherent risk measure but lacks of elicitability and considers only the left tail of the distribution; VaR is an elicitable risk measure but lacks of coherency and considers only a percentile of the distribution; variance considers the whole distribution but lacks of coherency and elicitability.

Chapter 4

Mean-EVAR Optimal Portfolios

⟨capME⟩

4.1 Mean-risk models

⟨capME1⟩

In the work of Markowitz (1952) the aim is to build a portfolio which minimizes the standard deviation and maximizes the expected return: mathematically, let $J = \{1, \ldots, m\}$ be the investment universe and r_j the random return of asset $j \in J$. The expected return is given by $\mu_j = \mathbb{E}[r_j]$ and the related covariance matrix is given by $C = [C_{i,j}]_{i,j=1,\ldots,m} = [\mathbb{E}[(r_i - \mu_i)(r_j - \mu_j)]]_{i,j=1,\ldots,m}$.
In this thesis, instead of working with returns, we prefer to adopt logreturns: from a computational point of view, the difference between using logreturns instead of returns is negligible when using daily or weekly data. From a statistical point of view logreturns are preferable to returns because they satisfy the property of additivity. Here and thereafter we adopt the notation r for logreturns.
Moreover consider $w = (w_1, \ldots, w_m) \in \Pi$ with $e^T w = 1$ as the portfolio weights of the investment, with Π defining the feasible set of possible portfolios, excluding short-selling: more precisely Π is supposed to be a set of feasible linear inequalities with non-negative variables. Hence the portfolio logreturn is given by $\mu(w) = \sum_{j=1}^{m} \mu_j w_j$.
The portfolio optimization problem can be modeled as the following parametric optimization problem:

$$\begin{aligned} &\min_{w\in\Pi} wCw^T \\ &\mu(w) \geqslant \mu^0 \end{aligned} \tag{4.1}$$

markQP

or equivalently

$$\begin{aligned} &\max \mu(w) \\ &wCw^T \leqslant var^0 \\ &w \in \Pi \end{aligned}$$

with parameters μ^0 and var^0.
This is a quadratic programming problem which nowadays is easy to solve: varying the parameter μ^0 from $\min_j \mu_j$ to $\max_j \mu_j$ (or equivalently varying var^0 from $\min_j var(r_j)$ to $\max_j var(r_j)$) and solving (4.1) we obtain the *optimal portfolios*.
Despite the Markowitz portfolio has a relatively easy formulation, there are some

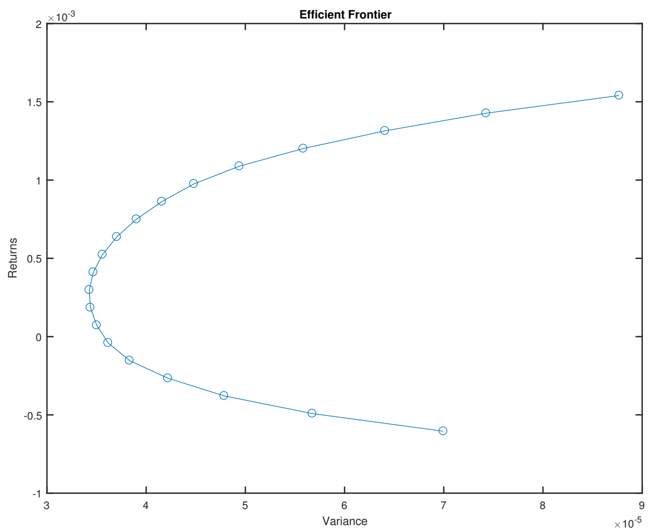


Figure 4.1: Example of optimal portfolios

⟨effron⟩

limitations. First of all, if we extend the set of the possible portfolios Π adding for example integrality or cardinality constraints (see e.g. Chang et al. (2000) and Cesarone et al. (2013)), the complexity of the problem increases and a quadratic objective function performs much worse than a linear objective function, in terms of computational costs and precision of the solution. Moreover the Markowitz model is not in general consistent with the Second Degree Stochastic Dominance, see e.g. Ogryczak and Ruszczynski (1999), because the solution of the problem (4.1) may not be consistent with the risk aversion axioms (Artzner et al. (1999)):

Definition 29. *A risk measure ρ is consistent with the Second Degree Stochastic*

Dominance if given two distribution X and Y with $X \geq_{SSD} Y$, then

$$\mu(X) - \rho(X) \geqslant \mu(Y) - \rho(Y)$$

.

This can be illustrated by means of the following simple example proposed by Mansini et al. (2014).

⟨example2⟩ **Example 30.** *Let us consider two portfolios A and B with the following logreturn distributions*

$$r_A = \begin{cases} 1 \text{ with probability } 1 \\ 0 \text{ otherwise} \end{cases} \qquad r_B = \begin{cases} 3 \text{ with probability } 0.5 \\ 5 \text{ with probability } 0.5 \end{cases}$$

Clearly the portfolio A is dominated by the portfolio B but they are both optimal portfolios because $\mu(A) - var(A) = 1$ and $\mu(B) - var(B) = 1$. In general the Second Degree Stochastic Dominance is not consistent with anyone of the mean-risk models that use a dispersion risk measure.

These are some of the reasons why many authors started looking for LP models with alternative risk measures which satisfy consistency with respect to Second Degree Stochastic Dominance (see e.g. Konno and Yamazaki (1991), Yitzhaki (1982), Rockafellar and Uryasev (2000), Ogryczak and Śliwiński (2011)).
Generalizing the problem in (4.1), a *mean-risk problem* for a risk measure $\rho \colon \Pi \to \mathbb{R} \cup \{+\infty\}$ is a bicriteria optimization problem given by:

$$\max_{w \in \Pi} [\mu(w), -\rho(w)] \tag{4.2}$$ bicriteria

A feasible portfolio $w^0 \in \Pi$ is called an *efficient solution* of problem (4.2) if $\nexists w \in \Pi$ such that $\mu(w) \geqslant \mu(x^0)$ and $\rho(w) \leqslant \rho(w^0)$ with at least one inequalities strict.
In general it is possible to find the set of efficient solutions specifying a lower bound for $\mu(w)$:

$$\begin{aligned} &\min_{w \in \Pi} \rho(w) \\ &\mu(w) \geqslant \mu^0 \end{aligned}$$

or equivalently specifying an upper bound for $\rho(w)$:

$$\begin{aligned} &\max_{w \in \Pi} \mu(w) \\ &\rho(w) \leqslant \rho^0. \end{aligned}$$

Considering μ^0 (or ρ^0) as a parameter we obtain the set of solutions which solve the problem (4.2). Moreover, if $\rho(w)$ is a convex function, then the solutions of the bicriteria optimization problem (4.2) lays on a convex line called *optimal portfolios frontier* (Fig. 4.1).

From a practical point of view, the lower part of the optimal portfolio frontier it is not taken into consideration because it is dominated by the upper part also known as *efficient frontier*: generally for every portfolio that lays the lower part, there exists a portfolio with the same risk but higher expected logreturn that lays on the higher part. In this work we consider the entire frontier because from a mathematical point of view it is also interesting to investigate the shape of the lower part of the optimal portfolio frontier.

4.2 LP portfolio optimization problems

In order to solve the bicriteria optimization problem in (4.2)

$$\max_{w\in\Pi}\left[\mu(w), -\rho(w)\right],$$

it is necessary to know the logreturns' distributions and the related functions $\mu(\cdot)$ and $\rho(\cdot)$, which in general are unknown quantities.

A common approach is to rely on historical data in order to provide an estimation of μ and ρ. For instance, considering the Markowitz portfolio problem in (4.1)

$$\begin{aligned} &\min_{w\in\Pi} wCw^T \\ &\mu(w) \geqslant \mu^0 \end{aligned} \tag{4.3}$$

markQ2

we evaluate $\mu(w)$ and C as the sample mean of the portfolio logreturn and the sample covariance matrix of the historical logreturns. Thus, supposing that the historical logreturn of the asset j is described by a discrete random variable r_j with $\mathbb{P}(r_j = r_{jt}) = p_t$, we can write

$$\hat{\mu}_j = \sum_{t=1}^{T} r_{jt}p_t \quad \text{and} \quad \hat{\mu}(w) = \frac{1}{J}\sum_{j=1}^{J} w_j\hat{\mu}_j.$$

Similarly the sample covariance matrix $\hat{C} = [c_{ij}]$ is given by

$$c_{ij} = \frac{1}{T-1}\sum_{t=1}^{T}(r_{jt} - \hat{\mu}_j)(r_{it} - \hat{\mu}_i),$$

and the Markowitz portfolio problem in (4.3) become

$$\min_{w\in\Pi} w\hat{C}w^T$$
$$\hat{\mu}(w) \geqslant \mu^0.$$

In this work we assume that the i.i.d assumption on the logreturns holds, for alternative methods we refer to Kempf and Memmel (2006), Lai et al. (2011) and Ledoit and Wolf (2003).

We now present some of the risk measures whose relative mean-risk problem is a LP problem. We refer to Mansini et al. (2014) for a recent detailed review on LP portfolio problems.

As above, let us suppose that the random variables r_j are discrete with $\mathbb{P}(r_j = r_{jt}) = p_t$.

Yitzhaki (1982) in his work introduced a mean-risk model using the *Gini's mean difference*

$$\Gamma(r_j) = \frac{1}{2}\sum_{t_1=1}^{T}\sum_{t_2=1}^{T} |r_{jt_1} - r_{jt_2}|\, p_{t_1}p_{t_2}.$$

Hence the Gini's mean difference for a portfolio composed of the assets is

$$\Gamma(w) = \sum_{j=1}^{J} w_j\Gamma(r_j) = \frac{1}{2}\sum_{j=1}^{J}\sum_{t_1=1}^{T}\sum_{t_2=1}^{T} |r_{jt_1} - r_{jt_2}|\, p_{t_1}p_{t_2}w_j$$

Then the mean-risk problem associated to the Gini's mean difference is

$$\begin{aligned} \min_{w\in\Pi}\Gamma(w) \quad s.t. \qquad &= \qquad \min_{w\in\Pi}\frac{1}{2}\sum_{j=1}^{J}\sum_{t_1=1}^{T}\sum_{t_2=1}^{T} |r_{jt_1} - r_{jt_2}|\, p_{t_1}p_{t_2}w_j \quad s.t. \\ \mu(w) \geqslant \mu^0 \qquad & \qquad \sum_{j=1}^{J}\sum_{t=1}^{T} w_j r_{jt} p_t \geqslant \mu^0 \end{aligned} \tag{4.4}$$ GiniLP

As pointed out by Yitzhaki (1982), the Gini's mean difference is consistent with the Second Degree Stochastic Dominance and moreover the problem (4.4) is LP solvable.

Another mean-risk model which is consistent with the Second Degree Stochastic Dominance and is LP solvable is based on the *mean absolute deviation*

$$\delta(x) = \mathbb{E}[|x - \mathbb{E}[x]|].$$

Konno and Yamazaki (1991) wrote the corresponding mean-risk model as

$$\min_{w\in\Pi}\sum_{t=1}^{T}\left|\sum_{j=1}^{J}(r_{jt}-\mu(r_j))w_j\right|p_t \quad s.t.$$
$$\sum_{j=1}^{J}\sum_{t=1}^{T}w_j r_{jt}p_t \geqslant \mu^0,$$

that can be written as the following LP problem:

$$\begin{aligned}
&\min_{w\in\Pi}\sum_{t=1}^{T}y_t p_t \quad s.t.\\
&\quad y_t+\sum_{j=1}^{J}(r_{jt}-\mu(r_j))w_j \geqslant 0, \quad t=1,\ldots,T\\
&\quad y_t-\sum_{j=1}^{J}(r_{jt}-\mu(r_j))w_j \geqslant 0, \quad t=1,\ldots,T\\
&\quad \sum_{j=1}^{J}\sum_{t=1}^{T}w_j r_{jt}p_t \geqslant \mu^0.
\end{aligned} \tag{4.5}$$

madLP

According to Yitzhaki (1982) and Konno and Yamazaki (1991), the mean-risk models in (4.4) and (4.5) are comparable with the mean-variance model, and they can be used as an alternative to the classic Markowitz's model with the benefit of dealing with an LP instead of a QP-problem.

In recent years, after the Basel II agreements (which in 2015 evolved into Basel III), the work of Rockafellar and Uryasev (2000) has become extremely important because they found that the mean-risk problem minimizing the *CVaR* is also an LP problem:

$$\begin{aligned}
&\min_{\gamma\in\mathbb{R},\ z_t\geq 0,\ w\in\Pi}\gamma+\frac{1}{(1-\alpha)T}\sum_{t=1}^{T}z_t \quad s.t.\\
&\quad \sum_{j=1}^{J}\mathbb{E}[r_{jt}]w_j \geqslant \mu^0\\
&\quad z_t \geqslant \sum_{j=1}^{J}w_j r_{jt}-\gamma, \quad t=1,\ldots,T,
\end{aligned}$$

where α is the level of the $CVaR$, γ is the value of the $CVaR_\alpha$, μ^0 is the threshold on the expected portfolio logreturn and z_t are artificial variables.

4.3 LP formulation of mean-EVaR portfolio problems

In this section we study the mean-EVaR portfolio problems and we write it as a LP problem. We can find analogous studies in recent papers of Jakobsons (2016a) and Jakobsons (2016b).

Let X be a random variable. Starting from the dual representation of the EVaR in (3.4) we want to solve:

$$EVaR_\tau(X) := \max_{\varphi \in \mathcal{M}_\beta} \mathbb{E}_p[-X\psi] \tag{4.6}$$ EVaRproblem

where

$$\mathcal{M}_\beta = \left\{ \varphi \in L^\infty(\Omega, \mathcal{F}, \mathbb{P}),\ \varphi \geq 0,\ \mathbb{E}[\varphi] = 1, \frac{\operatorname{ess\,sup}(\varphi)}{\operatorname{ess\,inf}(\varphi)} \leqslant \beta \right\}.$$

If X is a discrete random variable with $\mathbb{P}(X = x_i) = p_i, \forall i = 1, \dots, n$, we have that:

$$M_\beta = \left\{ f \in \mathbb{R}^n,\ f_i \geqslant 0,\ \sum_{i=1}^n f_i p_i = 1,\ f_i \leqslant \beta f_j,\ \forall i \neq j \right\}.$$

So the problem in (4.6) becomes:

$$EVaR_\alpha(X) = \begin{cases} \sup\limits_{f_i} \sum\limits_{i=1}^n -x_i f_i p_i \text{ s.t.} \\ \sum\limits_{i=1}^n f_i p_i = 1 \\ f_i - \beta f_j \leqslant 0,\ \forall i \neq j \\ f_i \geqslant 0,\ \forall i = 1, \dots, n. \end{cases} \tag{4.7}$$ dual_discrete

We observe that this is a LP problem with n variables and $n(n-1)+1$ constraints: that is problematic for random variables with a large number of possible outcomes, since the number of constraints grows quadratically in n.
Considering the quantity $m = \min_i f_i$, the constraint $f_i - \beta f_j \leqslant 0$ holds $\forall i \neq j$ and in particular for $f_j = m$, so we have that

$$\{f_i - \beta f_j \leqslant 0,\ \forall i \neq j\} \equiv \{m \leqslant f_i \leqslant \beta m,\ \forall i = 1, \dots, n\},$$

hence the problem (4.7) is equivalent to the following:

$$EVaR_\alpha(X) = \begin{cases} \sup\limits_{f_i} \sum\limits_{i=1}^{n} -x_i f_i p_i \text{ s.t.} \\ \sum\limits_{i=1}^{n} f_i p_i = 1 \\ m \leqslant f_i \leqslant \beta m, \ \forall i, \forall i = 1, \dots, n \\ m \geqslant 0, f_i \geqslant 0, \forall i = 1, \dots, n. \end{cases} \tag{4.8}$$ dual_discrete2

This LP problem has only $n+1$ variables and $2n+1$ constraints, so the number of constraints grow only linearly with respect to n. Moreover we have that the dual of the LP problem in (4.8) is

$$EVaR_\tau(X) = \begin{cases} \min\limits_{u_i, v_i} \ y \quad \text{s.t.} \\ p_i y - u_i + v_i \geqslant -p_i x_i \ \forall i \\ \sum\limits_{i=1}^{n} u_i - \beta \sum\limits_{i=1}^{n} v_i \geqslant 0 \\ u_i, v_i, \geqslant 0, \forall i = 1, \dots, n \end{cases} \tag{4.9}$$ LPdual

Suppose now that the outcomes of the random variable X are portfolio logreturns, so $X = Rw$, where R is the matrix of the historical logreturns and $w = (w_1, \dots, w_d)$ is the vector of portfolio weights. Finding the portfolio which minimize the $EVaR_\tau$ requires finding w such that

$$\begin{aligned} w \quad = \quad & \operatorname*{argmin}_{w} EVaR_\tau(Rw) \quad s.t. \\ & \sum w = 1, \end{aligned}$$

with possible additional portfolio constraints such as cardinality contstraints, integrality constraints, ... Considering the dual formulation in (4.9), we have that:

$$\begin{array}{lcl} w = & \operatorname*{argmin}\limits_{w_j} & \min\limits_{u_i, v_i} \ y \quad \text{s.t.} \\ & \sum_{j=1}^{d} w_j = 1 & p_i y - u_i + v_i \geqslant -p_i (Rw)_i \ \forall i \\ & \text{portfolio constraints} & \sum\limits_{i=1}^{n} u_i - \beta \sum\limits_{i=1}^{n} v_i \geqslant 0 \\ & & u_i, v_i, \geqslant 0, \forall i = 1, \dots, n \end{array}$$

Grouping the two minimization problem, we finally obtain

$$\begin{aligned}
&\min_{u_i,v_i,w_j} y \quad \text{s.t.}\\
&p_i y - u_i + v_i \geqslant p_i(Rw)_i \ \forall i\\
&\sum_{i=1}^{n} u_i - \beta \sum_{i=1}^{n} v_i \geqslant 0\\
&u_i, v_i \geqslant 0,\ \forall i = 1,\dots,n\\
&\sum_{j=1}^{m} w_j = 1 \qquad\qquad (4.10)\ \boxed{\texttt{EVARLP}}\\
&\left.\begin{array}{ll}
\text{No Shortselling} & w_j \geqslant 0\ \forall j = 1,\dots,m\\
\text{Optimal Portfolios} & \sum_{i=1}^{m} w_i \mathbb{E}[r_i] = \mu_p\\
\dots &
\end{array}\right\} \Longrightarrow \text{Additional Portfolio Constraints}
\end{aligned}$$

4.4 Numerical examples

In the first part of this section we consider optimal portfolios on simulated data, while in the second part we consider real data: our purpose is to compare the portfolio problem using the $EVaR_\tau$ as risk measure with the usual risk measures used in literature.
All the experiments are performed using MatLab and the optimization problems are solved with the functions `linprog`, `quadprog` and `fmincon`. The solutions to Linear and Quadratic Programming problems are checked with GAMS (General Algebraic Modeling System).

4.4.1 Simulated data

⟨numEVAR⟩
We first consider three assets A, B and C described by standard normal distributions without correlation. In this case clearly the distribution of the three assets is exchangeable, and for symmetry reasons the optimal portfolio under any coherent risk measure coincides with the equally weighted portfolio $(w_A, w_B, w_C) = (1/3, 1/3, 1/3)$. In order to assess the variability of the optimal portfolio, we simulate samples of length $T = 20, 50, 100$ from the aforementioned normal distribution, and compute the distribution of minimal $EVaR_\tau$ portfolios for τ ranging from 1% to 20%. In this experiment our aim is to observe the variability of the weight of optimal portfolios considering different values of τ and sample sizes.
In order to have convexity of the EVaR, τ must be in $(0, 1/2]$: we focus on the interval $(0, 0.2]$ because for value of τ close to $1/2$, the $EVaR_\tau$ loose its meaning as risk measure by the fact that $EVaR_{1/2}(\cdot) = -e_{1/2}(\cdot) = -\mathbb{E}[\cdot]$, and the related mean-risk optimization problem becomes

$$\max_{w \in \Pi} [\mu(w), -\rho(w)]$$

with $-\rho(w) \approx \mu(w)$.

Moreover, considering the sample size, we expect that increasing the possible scenarios T, the variability should decrease. We then focus on the case $T = 100$ that we believe might be a realistic value by practical portfolio optimization.

The results are reported in Fig. 4.3; the three lines in each subpanel correspond to the 10$^{\text{th}}$ percentile, the median and the 90$^{\text{th}}$ percentile of the portfolio weights. For the sake of completeness we compare with the corresponding percentiles of the Markowitz portfolios.

In Fig. 4.4 we compare the distribution of the optimal portfolios originated by the minimization of the variance, of $CVaR_{0.05}$, and of $EVaR_{0.05}$ and $EVaR_{0.20}$, with a sample size $T = 100$.

In Fig. 4.5 and 4.6 we perform the same experiments of Fig. 4.3 and 4.4 but considering the cases of correlated standard normal marginals. We use the following correlation matrices:

$$\Sigma_+ = \begin{pmatrix} 1 & 0.3 & 0.5 \\ 0.3 & 1 & 0.7 \\ 0.5 & 0.7 & 1 \end{pmatrix}, \qquad \Sigma_- = \begin{pmatrix} 1 & -0.3 & 0.5 \\ -0.3 & 1 & -0.7 \\ 0.5 & -0.7 & 1 \end{pmatrix} \tag{4.11}$$ corrmat

From theory, considering symmetric distributions with or without correlation, we do not expect substantial differences between minimizing EVaR and CVaR: for this reason we now consider a portfolio composed by two normal distributions and a reflected gamma distribution with probability density function

$$f(x; k, \theta, \mu, \sigma) = -\frac{x^{k-1}e^{-\frac{x}{\theta}}}{\theta^k \Gamma(k)} \cdot \sigma + \mu,$$

where μ and σ are the mean and the standard deviation of the distribution, and

$$\Gamma(k) = \int_0^\infty y^{k-1} e^{-t} dt.$$

In order to highlight the differences between CVaR minimization and EVaR minimization, we set the parameters of the distributions to ensure that the Values at Risk at level $\alpha = 0.05$ coincides and the left tails are similar.
In this way we expect that

$$CVaR_{0.05}(\mathcal{N}(\mu_\mathcal{N}, \sigma_\mathcal{N})) \approx CVaR_{0.05}(\Gamma(k, \theta, \mu_\Gamma, \sigma_\Gamma)),$$

hence theoretically the CVaR minimization should give the same weight to the three asset while the EVaR minimization should prefer the assets with a fatter right tail.
In particular we set

$$\begin{aligned} &\text{Asset A} \sim \mathcal{N}(\mu_\mathcal{N}, \sigma_\mathcal{N}), && \text{with } \ \mu_\mathcal{N} = 0, \ \sigma_\mathcal{N} = 2.25 \\ &\text{Asset B} \sim \mathcal{N}(\mu_\mathcal{N}, \sigma_\mathcal{N}), && \text{with } \ \mu_\mathcal{N} = 0, \ \sigma_\mathcal{N} = 2.25 \\ &\text{Asset C} \sim -\Gamma(k, \theta, \mu_\Gamma, \sigma_\Gamma), && \text{with } \ k = 5, \ \theta = 1, \ \mu_\Gamma = -0.912, \sigma_\Gamma = 1.5 \end{aligned} \tag{4.12}$$ parameters

In Fig. 4.2 we report the normal distribution and the reflected gamma distribution with the parameters in (4.12), and consequently we have

$$
\begin{aligned}
&CVaR_{0.05}(\mathcal{N}(\mu_{\mathcal{N}}, \sigma_{\mathcal{N}})) = -0.465,\\
&CVaR_{0.05}(\Gamma(k, \theta, \mu_{\Gamma}, \sigma_{\Gamma})) = -0.547,\\
&VaR_{0.05}(\mathcal{N}(\mu_{\mathcal{N}}, \sigma_{\mathcal{N}})) = VaR_{0.05}(-\Gamma(k, \theta, \mu_{\Gamma}, \sigma_{\Gamma})) = -3.710.
\end{aligned}
$$

In Fig. 4.7 and 4.8 we perform the experiment on a portfolio composed by assets

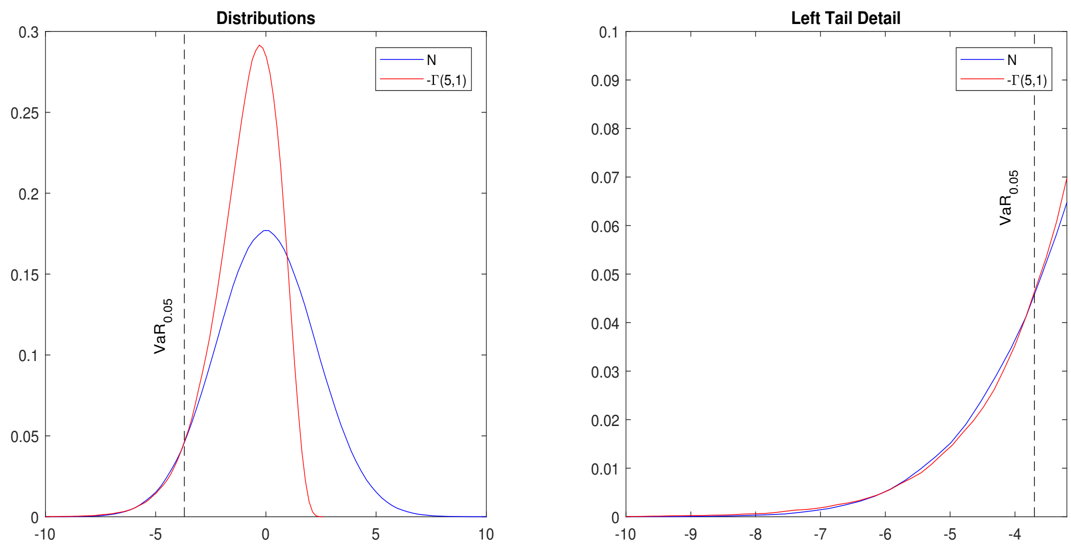


Figure 4.2: Probability functions and left tail detail

⟨pdf⟩

with distribution defined in eq. (4.12) and finally in Fig. 4.9 and 4.10 we introduce correlation among the assets: A, B and C will have correlation matrices Σ_+ and Σ_-.

Our aim is to reproduce the experiments performed by Bertsimas et al. (2004) in order to understand how the level of τ impacts on variability and make a preliminary study on the behavior of the $EVaR_\tau$ varying the parameters and the models.
Since there is an extensive body of financial literature based on the CVaR, we want to take as an example the first test on it and compare them with the EVaR in order make a preliminary analysis.
In their work Bertsimas et al. (2004) did an analysis of the CVaR comparing to others well known risk measures: with the method proposed by Rockafellar et al. (2006) evaluating the mean-CVaR portfolio problem is reduced to a LP problem,

this is one of the reasons why we perform the same test considering the fact that also the mean-EVaR portfolio problem is LP.

In Fig. 4.3, 4.5, 4.7 and 4.9 we can observe that in most of the cases, the minimum of the distance between the 10$^{\text{th}}$ percentile and the 90$^{\text{th}}$ percentile is often reached in a neighborhood of $\tau = 0.05$ and the maximum is reached for $\tau = 0.2$. Thus we study in details how are composed the portfolios minimizing $EVaR_{0.05}$ and $EVaR_{0.2}$: for completeness we evaluate also the minimum variance portfolio and the portfolio obtained minimizing the $CVaR_{0.05}$ (Fig. 4.4, 4.6, 4.8 and 4.10).

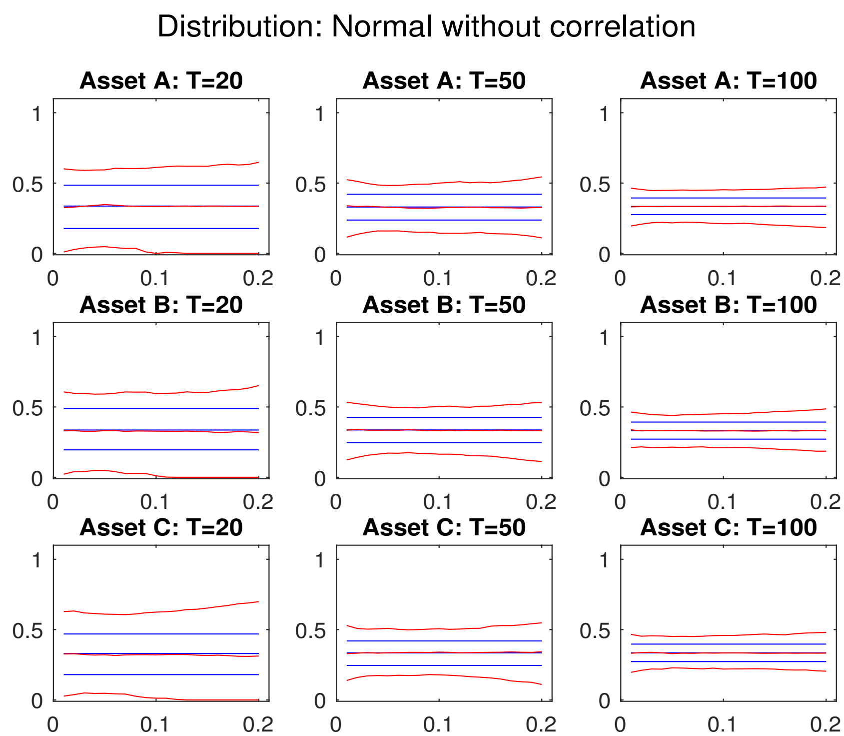


Figure 4.3: Variability of the weights of minimum $EVaR_\tau$ portfolios for $\tau = 0.01, \ldots, 0.2$ on simulated uncorrelated standard normal data. The red lines indicate respectively the 90$^{\text{th}}$ percentile, the median and the 10$^{\text{th}}$ percentile of the portfolio weights. The blue lines represent the same quantities for a minimal variance portfolio.

⟨fig1⟩

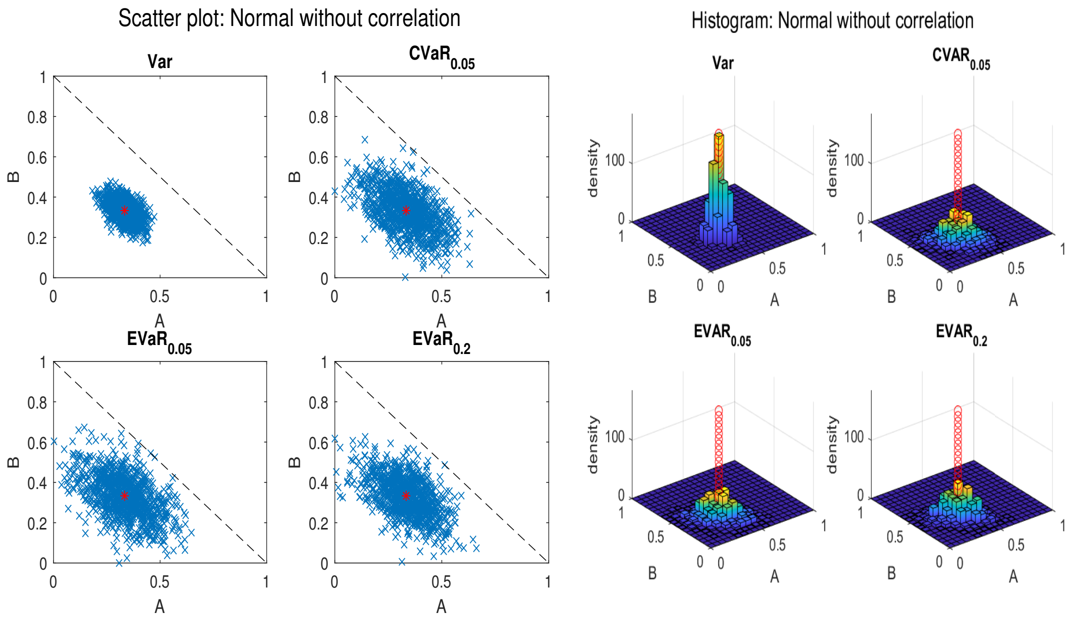


Figure 4.4: Scatter plot and histogram of the weights of the assets A and B, the weight of C is given by the relation $\sum \omega_i = 1$, the pattern is the same of Fig. 4.3 with 1000 simulations. The red points represent the theoretical value of the portfolio weights given by the minimal variance portfolio.

⟨fig2⟩

As we expect form theory we can observe that, in the standard normal case with uncorrelated data (Fig. 4.4) the cloud of points evaluated minimizing the $EVaR_{0.05}$ has the same shape of the one evaluated minimizing the $CVaR_{0.05}$. The cloud of points evaluated minimizing the $EVaR_{0.2}$ has the same shape but has more dispersion.

We now consider three asset with correlation matrices Σ_+ and Σ_-: in the first case all the asset has positive correlation, in particular

$$\rho_{A,B} = 0,3 \quad \rho_{A,C} = 0,7 \quad \rho_{B,C} = 0,5$$

Hence we expect that the resulting portfolio should have more weight on assets A and B at the expense of C.

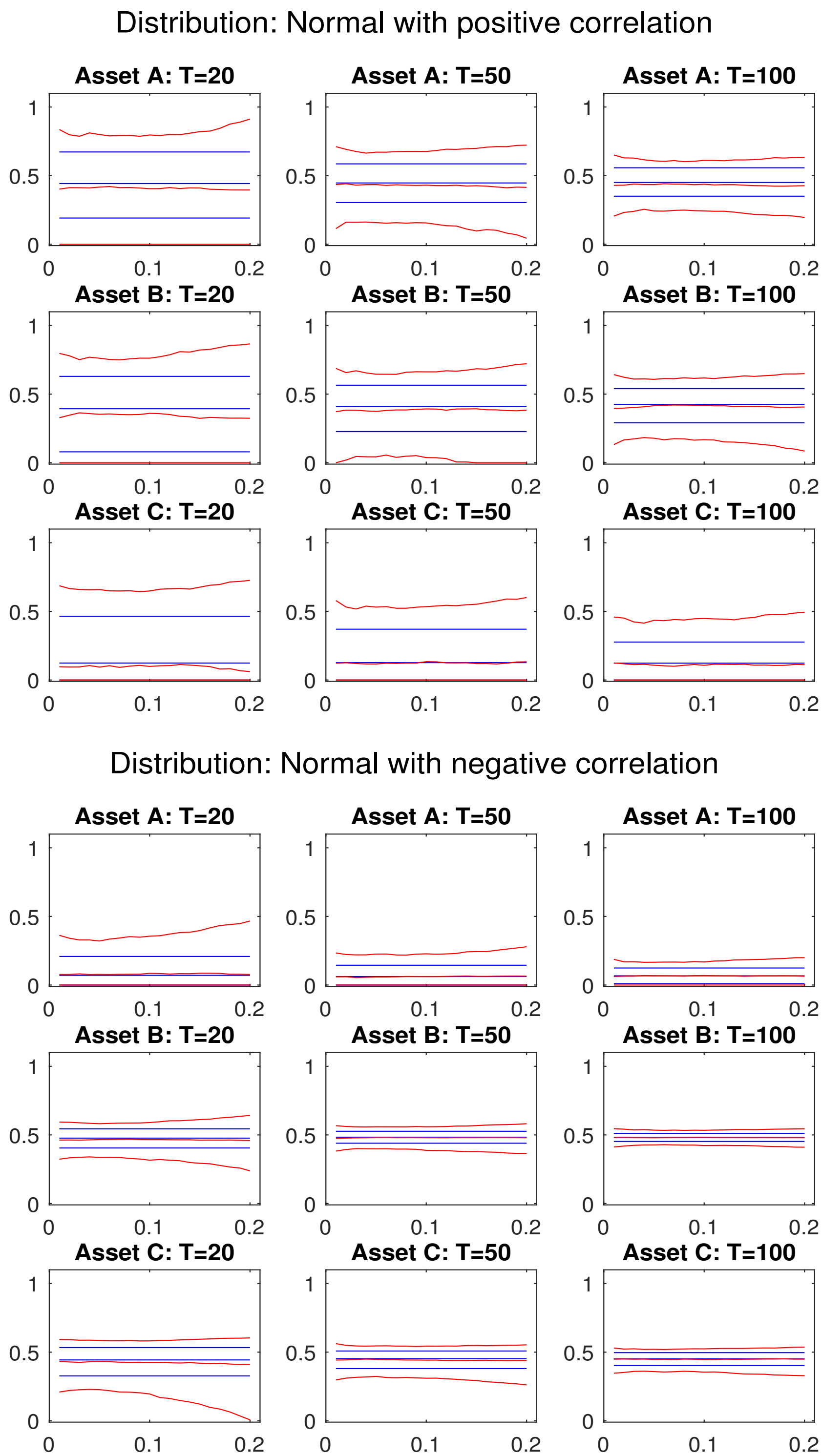


Figure 4.5: Variability of the weights of minimum $EVaR_\tau$ portfolios for $\tau = 0.01, \ldots, 0.2$ on simulated correlated standard normal data of length $T = 20, 50, 100$: the correlation is given by the matrices (4.11) The red lines indicate respectively the 90^{th} percentile, the median and the 10^{th} percentile of the portfolio weights. The blue lines represent the same quantities for a minimal variance portfolio.

⟨fig3⟩

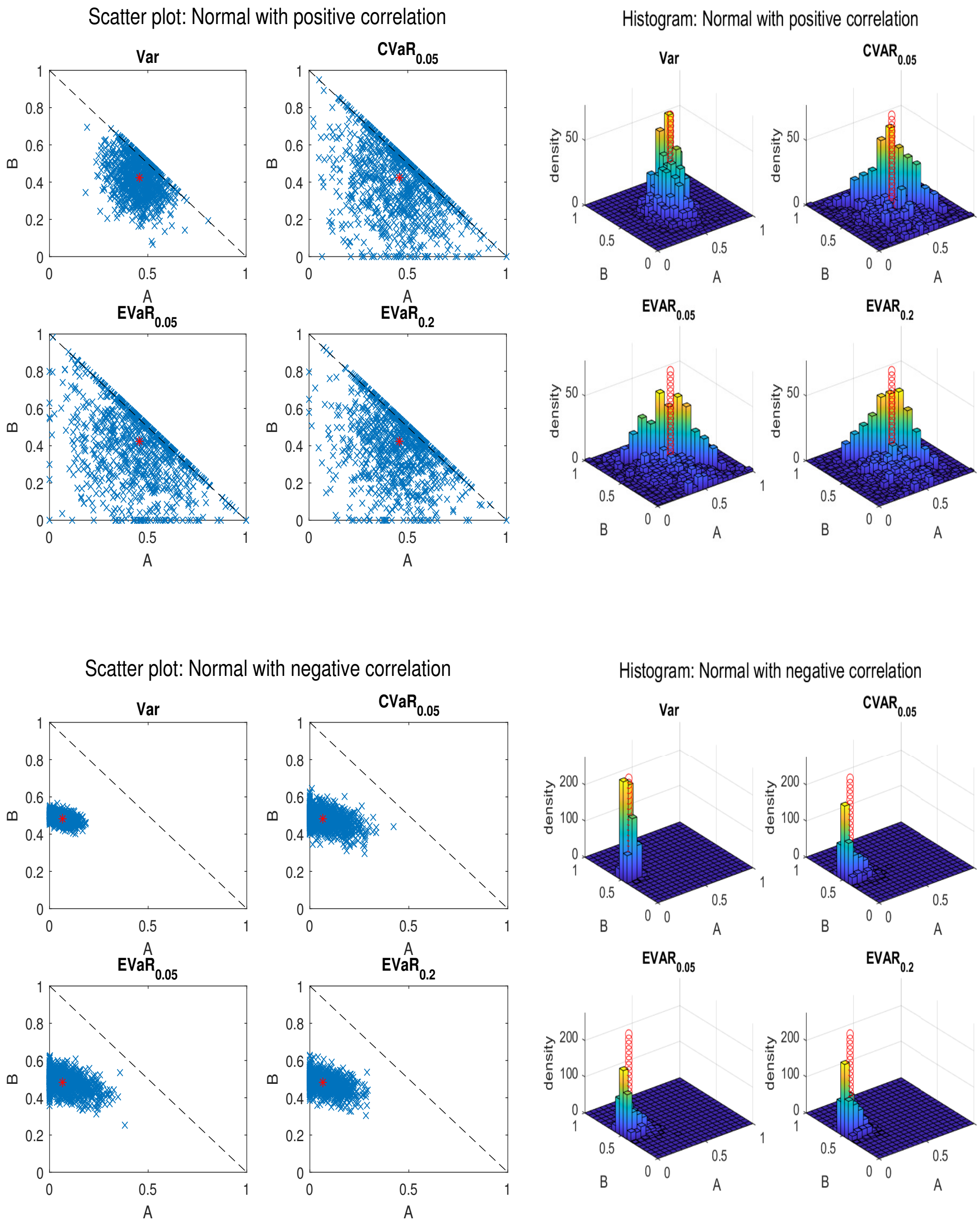


Figure 4.6: Scatter plot and histogram of the weights of the assets A and B, the weight of C is given by the relation $\sum \omega_i = 1$, the pattern is the same of Fig. 4.5 with 1000 simulations. The red points represent the theoretical value of the portfolio weights given by the minimal variance portfolio.

⟨fig4⟩

In the second case some asset has negative correlation, in particular

$$\rho_{A,B} = -0,3 \quad \rho_{A,C} = -0,7 \quad \rho_{B,C} = 0,5$$

Hence we expect that the resulting portfolio should have more weight on assets B and C at the expense of A.
Observing the results introducing correlation among the asset logreturns (Fig. 4.5 and 4.6) we can observe that:

- considering the correlation matrix Σ_+ the shapes of the clouds of points in the scatter plots gather on the diagonal, which means that buying the third asset does not decrease the level of the risk;

- considering the correlation matrix Σ_- the shapes of the clouds of points in the scatter plots gather on the vertical axis, which means that buying the first asset does not decrease the level of the risk.

Comparing the $EVaR_\tau$ and $CVaR_\alpha$ portfolios, we can infer similar conclusion to the uncorrelated case. Thus we can assume that, introducing a correlation through the data, evaluating the weights minimizing the $EVaR_{0.05}$ has no significant differences from evaluating the weights minimizing the $CVaR_{0.05}$.
As we expected from theory, when the asset has symmetric distributions with or without correlation, minimizing $EVaR_\tau$ gives similar results to minimizing $CVaR_\alpha$.

We now consider uncorrelated assets' distributions as described in Eq. 4.12, i.e.: asset A and B have a normal distribution with parameters $\mu_\mathcal{N} = 0,\ \sigma_\mathcal{N} = 2.25$ and asset C has a reflected gamma distribution with parameters $k = 5,\ \theta = 1,\ \mu_\Gamma = -0.912, \sigma_\Gamma = 1.5$.

As we expect from theory, in Fig. 4.7 we can observe that in general the weight of the asset C is lower with respect to the Markowitz case and, increasing the value of τ, the value of it decreases: this phenomenon is due to the fact that, by definition of the expectile in 3.1

$$EVaR_\tau = -e_\tau(X) = -\operatorname*{argmin}_{x\in\mathbb{R}} \left\{\tau\mathbb{E}\left[(X-x)_+^2\right] + (1-\tau)\mathbb{E}\left[(X-x)_-^2\right]\right\},$$

increasing τ, we prefer distribution with a fatter right tail.

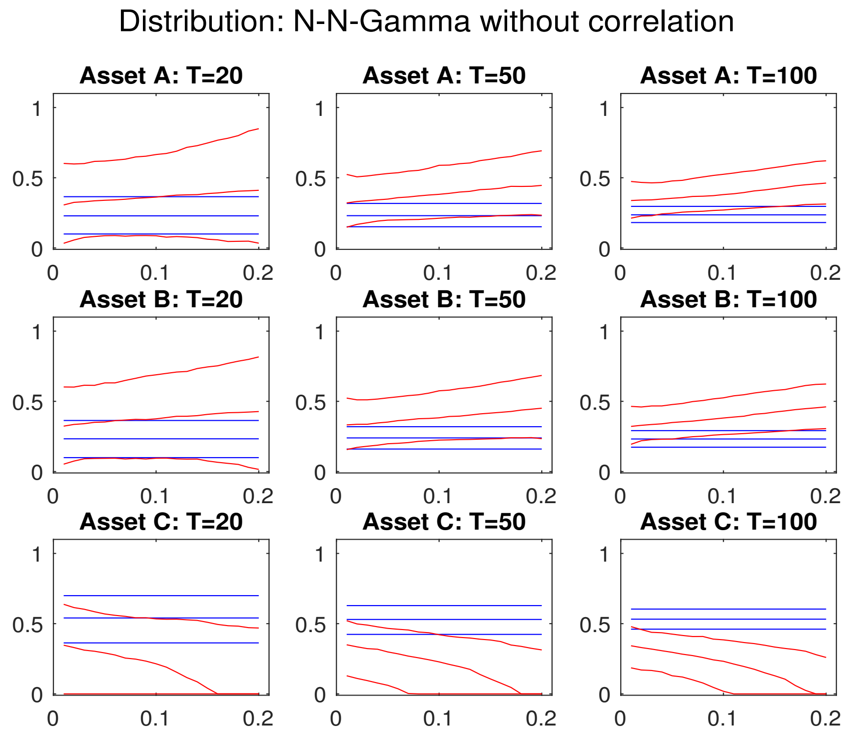


Figure 4.7: Variability of the weights of minimum $EVaR_\tau$ portfolios for $\tau = 0.01, \ldots, 0.2$ on simulated uncorrelated data of length $T = 20, 50, 100$: the assets' distributions are described in Eq. 4.12. The red lines indicate respectively the 90$^{\text{th}}$ percentile, the median and the 10$^{\text{th}}$ percentile of the portfolio weights. The blue lines represent the same quantities for a minimal variance portfolio.

⟨fig5⟩

Furthermore considering the stability of the algorithm, again the variability of the weights decreases when the sample size increases.

In Fig. 4.8, focusing on the case $T = 100$, we can observe that the scatter plot evaluated minimizing the $CVaR_{0.05}$ differs from the scatter plot evaluated minimizing $EVaR_{0.05}$: in fact in the second case the cloud of points is shifted towards the diagonal, this means that evaluating the portfolio's weights in this way, assets' distributions with a fatter right tail are preferred.

From the financial point of view, considering the density functions of the assets (Fig. 4.2) assets A and B are preferable to asset C because, having the same left tail (equal losses), they have a greater mean and a fatter right tail (greater profits).

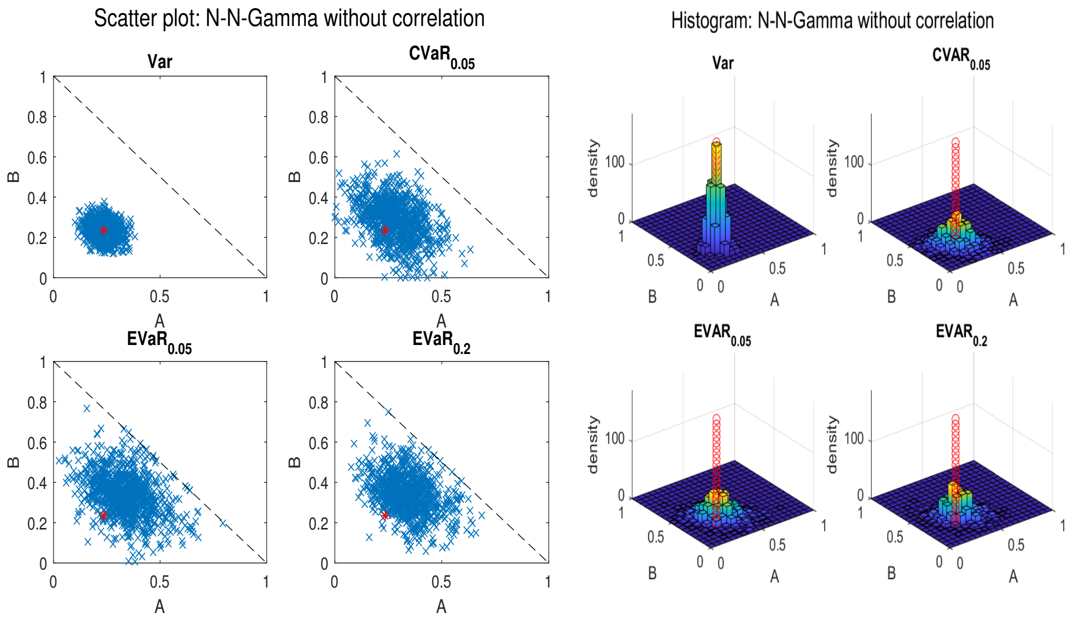


Figure 4.8: Scatter plot and histogram of the weights of the assets A and B, the weight of C is given by the relation $\sum \omega_i = 1$, the pattern is the same of Fig. 4.7 with 1000 simulations. The red points represent the theoretical value of the portfolio weights given by the minimal variance portfolio.

⟨fig6⟩

We now study the case of a portfolio composed by asset A and B with normal distribution and asset C with a reflected gamma distribution: furthermore we introduces the correlation matrices Σ_+ and Σ_- (Eq. 4.11).

The results (Fig. 4.9 and 4.10) are similar with respect to the uncorrelated case, moreover trend of giving the most of the weight to asset A and B is highlighted even more.

Comparing the Markowitz's portfolio and the $\text{CVaR}_{0.05}$ portfolio with the EVaR_τ ones we have opposite results both with positive and negative correlation: in the first case the histograms shows that Markowitz and CVaR prefer asset C (the bins lays on the first part of x axis), while EVaR avoid it (the bins lays on the diagonal); in the second case Markowitz and CVaR prefer assets B and C (the bins lays on the middle of y axis), while EVaR diversifies more shifting the bins to the center of the graph.

Considering the stability of the algorithm, again the variability of the portfolio weights decreases when the sample size increases.

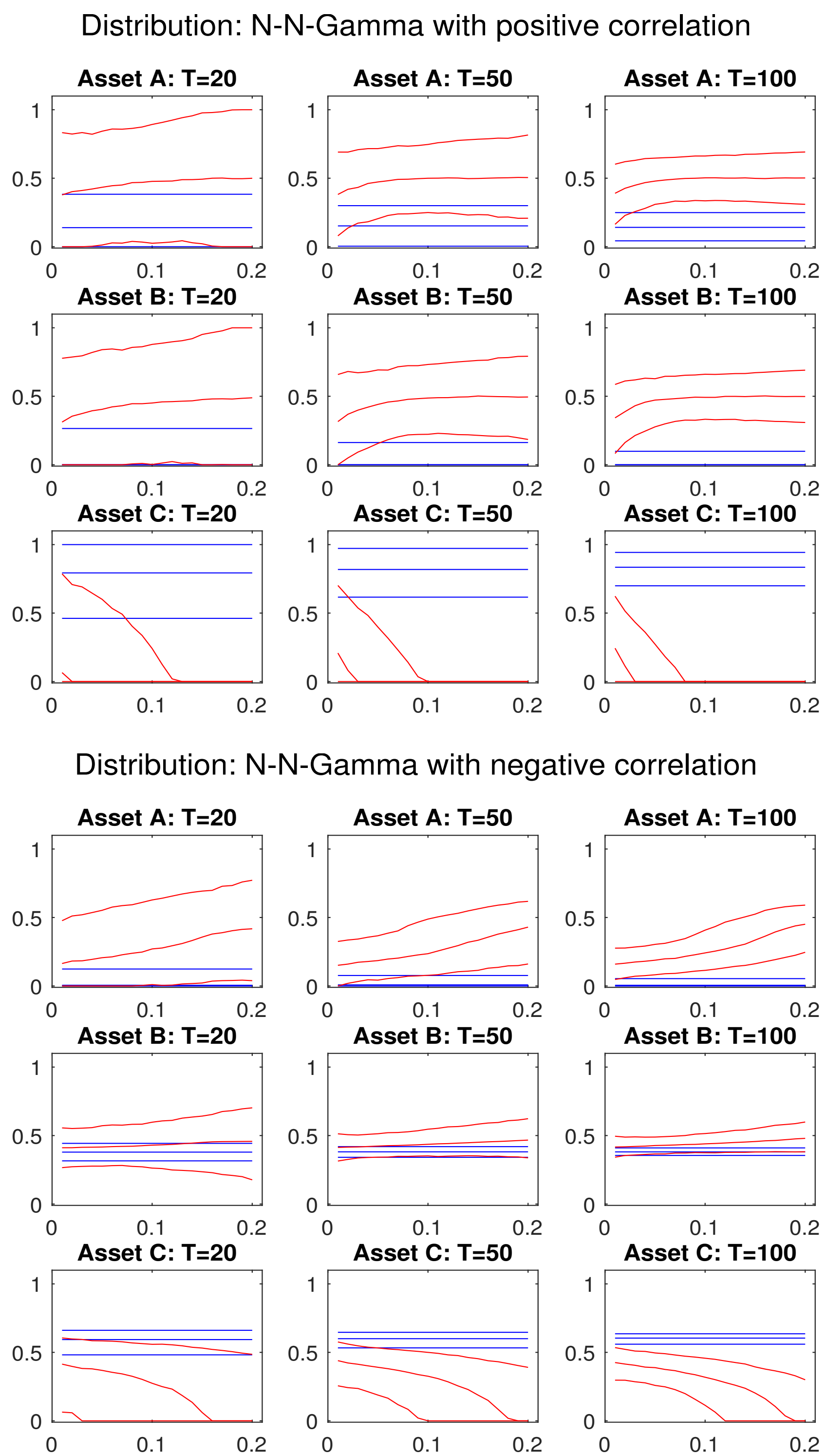


Figure 4.9: Variability of the weights of minimum $EVaR_\tau$ portfolios for $\tau = 0.01, \ldots, 0.2$ on simulated correlated data of length $T = 20, 50, 100$: the assets' distributions are described in eq. 4.12. The red lines indicate respectively the 90$^{\text{th}}$ percentile, the median and the 10$^{\text{th}}$ percentile of the portfolio weights. The blue lines represent the same quantities for a minimal variance portfolio.

⟨fig6a⟩

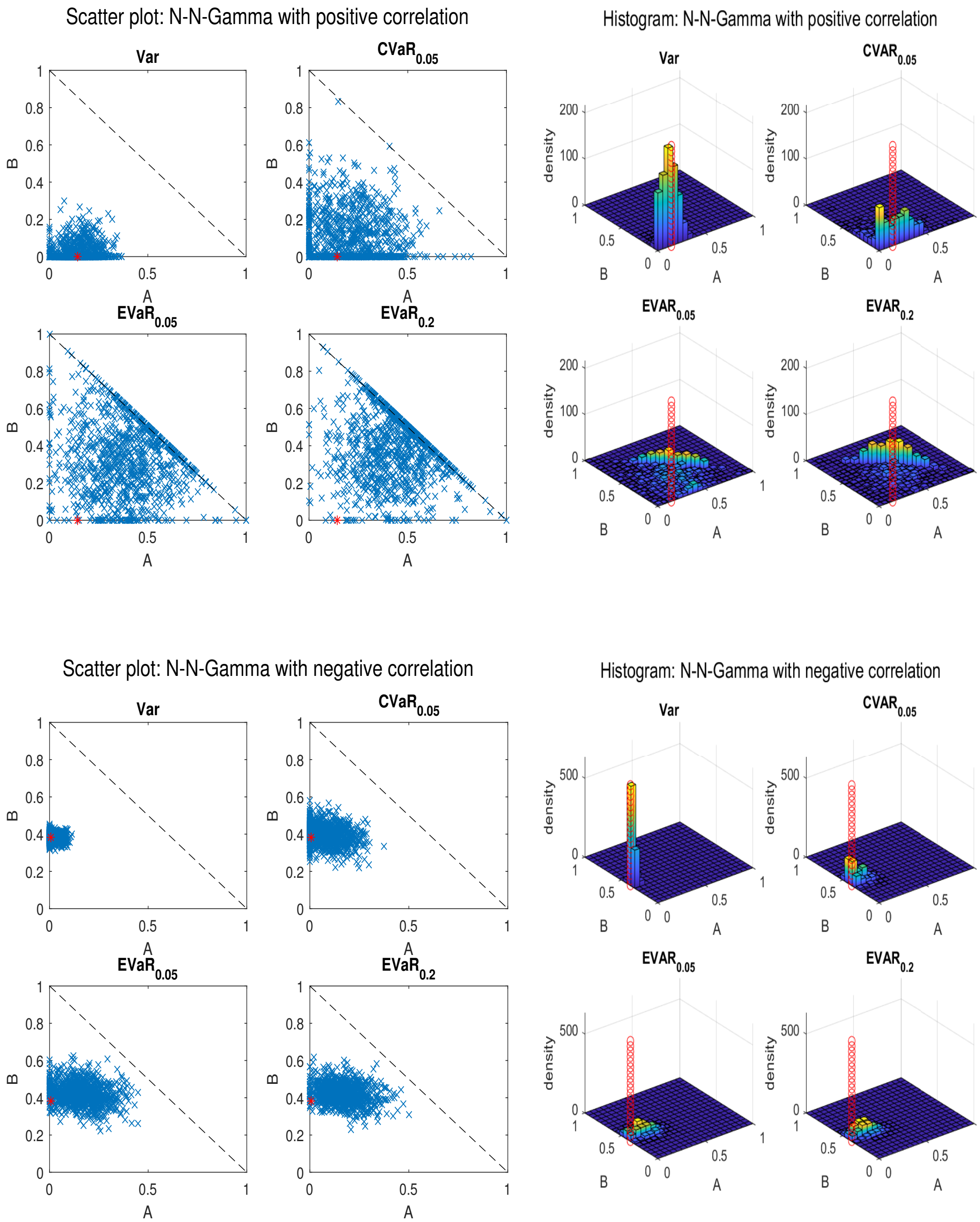


Figure 4.10: Scatter plot and histogram of the weights of the assets A and B, the weight of C is given by the relation $\sum \omega_i = 1$, the pattern is the same of Fig. 4.7 with 1000 simulations. The red points represent the theoretical value of the portfolio weights given by the minimal variance portfolio.

⟨fig6b⟩

4.4.2 Historical Data

⟨histEVaR⟩

In this section we consider the logreturns of the companies that constitute the $S\&P$ 500 index in two different periods: January 2011 - December 2013 and January 2014 - December 2016. We want to evaluate the set of optimal portfolios composed of the first five assets ordered by their weight in the $S\&P$ 500 index (Apple Inc., Microsoft Corporation, Exxon Mobil Corporation, Amazon.com Inc. and JPMorgan Chase & Co).

In the experiments we compare the optimal portfolios generated with the Markowitz portfolio theory, the mean-$CVaR_\alpha$ analysis and the mean-$EVaR_\tau$ analysis: for the sake of completeness we evaluate also the equally weighted portfolio.

In order to make the $CVaR_\alpha$ comparable with the $EVaR_\tau$, we consider α =0.01, 0.05, 0.1, 0.15, 0.2, 0.25, we evaluate the $CVaR_\alpha$ of a standard normal distribution using the closed formula

$$CVaR_\alpha(\mathcal{N}(0,1)) = -\frac{1}{\alpha}\phi(\Phi^{-1}(\alpha)),$$

and we find the τ that matches $EVaR_\tau(\mathcal{N}(0,1))$ with $CVaR_\alpha(\mathcal{N}(0,1))$ using the first order condition (3.3). In the case of the standard normal we have that

$$(1-\tau)\int_{-\infty}^{c}\Phi(s)ds = \tau\int_{c}^{\infty}(1-\Phi(s))ds,$$

where $c = EVaR_\tau(\mathcal{N}(0,1)) = CVaR_\alpha(\mathcal{N}(0,1))$.

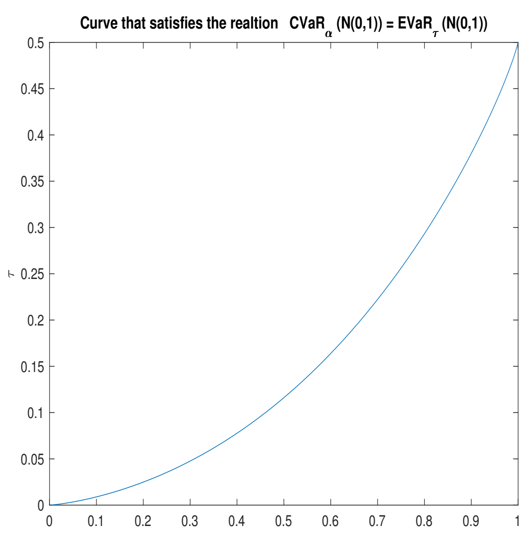


Figure 4.11: Relation between α and τ when the equation $EVaR_\tau(\mathcal{N}(0,1)) = CVaR_\alpha(\mathcal{N}(0,1))$ holds.

In particular we will consider the values in Tab. 4.1:

α	0.01	0.05	0.1	0.15	0.2	0.25
τ	0.0004	0.0035	0.0089	0.0161	0.0249	0.0354
c	-2.6652	-2.0627	-1.7550	-1.5544	-1.3998	-1.2711

Table 4.1: Correspondence between values of α and τ

⟨alphatau⟩

First we evaluate the optimal portfolios' frontiers evaluated considering the whole dataset over the two periods. Afterwards we make an *out-of-sample* analysis considering a rolling window of length 100 days giving equal probability 1/100 to the possible outcomes: every day the portfolio is rebalanced following the rules of Markowitz portfolio theory, mean-$CVaR_{0.2}$ analysis, mean-$EVaR_{0.05}$ analysis and equally weighted portfolio, then empirical distribution of the portfolios' logreturns is evaluated. Similar experiments are performed for a portfolio composed of the first 250 assets ordered by their weight in the $S\&P$ 500 index (see Tab. 4.2).

In Fig. 4.12 the asset prices of the title taken into account are showed, in Tab. 4.3 and 4.4 the statistics of the daily logreturns in the two different periods are reported.

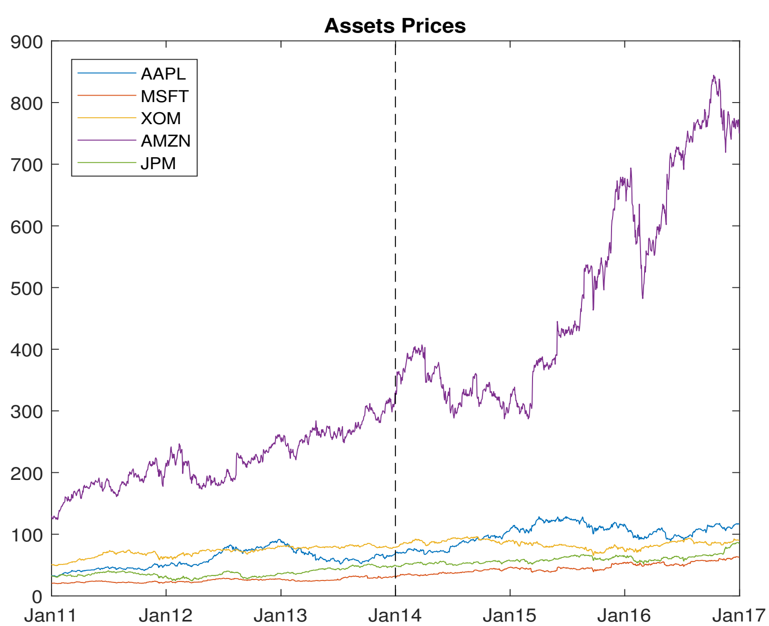


Figure 4.12: Asset prices of the first 5 asset of $S\&P$ 500 in the periods Jan11-Dec13 and Jan14-Dec16

⟨AssetPrice⟩

Symbol	Weight	Symbol	Weight	Symbol	Weight	Symbol	Weight	Symbol	Weight
AAPL	3,56	URI	0,42	BK	0,24	PCG	0,16	AAL	0,12
MSFT	2,51	CVS	0,41	SO	0,24	ECL	0,16	AVB	0,12
XOM	1,71	AVGO	0,41	MON	0,24	KR	0,16	EIX	0,12
AMZN	1,63	SBUX	0,41	PRU	0,23	CCI	0,16	FISV	0,12
FB	1,58	PCLN	0,4	BLK	0,23	DE	0,16	BAX	0,12
JPM	1,57	QCOM	0,39	FDX	0,23	AEP	0,16	ILMN	0,12
BRK-B	1,57	LLY	0,38	PYPL	0,23	LYB	0,16	PPL	0,12
JNJ	1,56	TXN	0,38	D	0,23	HUM	0,15	HCN	0,12
GE	1,32	COST	0,38	AMT	0,23	VLO	0,15	PCAR	0,12
WFC	1,29	TWX	0,37	CSX	0,23	AON	0,15	CCL	0,12
T	1,28	NKE	0,37	ADP	0,22	APD	0,15	CMI	0,12
GOOGL	1,24	UNH	0,37	KMB	0,22	MCK	0,15	WMB	0,12
GOOG	1,21	ABT	0,37	RTN	0,22	ALL	0,15	YUM	0,12
PG	1,19	ACN	0,37	KMI	0,22	SE	0,15	ADI	0,12
BAC	1,18	WBA	0,36	COF	0,22	SYF	0,14	ZBH	0,12
CVX	1,07	CHTR	0,35	ANTM	0,21	STI	0,14	DVN	0,12
VZ	1	MDLZ	0,35	AET	0,21	WM	0,14	ED	0,11
PFE	0,99	DOW	0,34	ESRX	0,21	AFL	0,14	WDC	0,11
CMCSA	0,9	DD	0,33	NOC	0,21	STT	0,14	EQR	0,11
MRK	0,89	LMT	0,33	CME	0,2	INTU	0,14	NWL	0,11
HD	0,85	MS	0,33	ITW	0,2	ALXN	0,14	VRTX	0,11
INTC	0,85	LOW	0,32	HPE	0,2	MAR	0,14	VTR	0,11
C	0,82	AIG	0,32	EMR	0,2	ZTS	0,14	PEG	0,11
DIS	0,81	NVDA	0,32	YHOO	0,2	STZ	0,14	APA	0,11
KO	0,81	NFLX	0,31	JCI	0,2	ISRG	0,14	PGR	0,11
V	0,8	COP	0,31	SYK	0,19	ATVI	0,14	AZO	0,11
PM	0,8	CB	0,31	AMAT	0,19	EQIX	0,14	IP	0,11
CSCO	0,79	TMO	0,31	APC	0,19	ROST	0,14	UAL	0,11
IBM	0,77	AXP	0,3	CI	0,19	DFS	0,14	APH	0,11
UTX	0,77	PNC	0,3	BBT	0,19	REGN	0,14	XEL	0,11
PEP	0,76	CL	0,3	BDX	0,19	TEL	0,13	IR	0,1
MO	0,71	BIIB	0,3	TGT	0,19	HPQ	0,13	CXO	0,1
AMGN	0,63	NEE	0,29	GIS	0,19	FIS	0,13	FCX	0,1
ORCL	0,61	ADBE	0,29	MMC	0,18	PPG	0,13	ROP	0,1
SLB	0,57	MET	0,29	DAL	0,18	ORLY	0,13	DLPH	0,1
MMM	0,54	EOG	0,29	NSC	0,18	SYY	0,13	BXP	0,1
MDT	0,52	SPG	0,29	PSX	0,18	MPC	0,13	DG	0,1
WMT	0,52	CAT	0,28	CTSH	0,18	PLD	0,13	OMC	0,1
MCD	0,52	DUK	0,27	EBAY	0,17	SRE	0,13	FITB	0,1
MA	0,51	KHC	0,27	BSX	0,17	MTB	0,13	PH	0,1
ABBV	0,5	OXY	0,26	ICE	0,17	BHI	0,13	NEM	0,1
BA	0,47	GD	0,26	TRV	0,17	MU	0,13	AMP	0,1
AGN	0,46	CRM	0,26	LUV	0,17	ADM	0,13	ROK	0,1
HON	0,46	DHR	0,26	SPGI	0,17	GLW	0,13	KEY	0,1
CELG	0,46	RAI	0,25	PX	0,17	CBS	0,13	EW	0,1
GS	0,45	TJX	0,25	PSA	0,16	CAH	0,13	NUE	0,1
UNP	0,45	HAL	0,25	PXD	0,16	EA	0,13	K	0,1
GILD	0,44	GM	0,25	EXC	0,16	WY	0,12	TSN	0,1
BMY	0,43	SCHW	0,24	ETN	0,16	SHW	0,12	CFG	0,1
USB	0,43	F	0,24	FOXA	0,16	HCA	0,12	LRCX	0,1

Table 4.2: First 250 asset of $S\&P$ 500 and their weight on January 2017

⟨weightsp⟩

	AAPL	**MSFT**	**XOM**	**AMZN**	**JPM**
Min	-1.319e-01	-1.210e-01	-6.388e-02	-1.353e-01	-9.888e-02
Max	8.502e-02	7.033e-02	5.087e-02	1.462e-01	8.101e-02
Mean	9.426e-04	5.804e-04	6.791e-04	1.313e-03	5.627e-04
Std	1.725e-02	1.432e-02	1.144e-02	2.036e-02	1.914e-02
Skew	-3.983e-01	-4.113e-01	-2.516e-01	2.038e-01	-1.369e-01
Kurt	8.463	1.049e+01	6.376	9.703	6.211

Table 4.3: Daily Logreturn during the period January 2011 - December 2013

⟨tablestat1⟩

	AAPL	**MSFT**	**XOM**	**AMZN**	**FB**
Min	-8.330e-02	-9.710e-02	-4.843e-02	-1.165e-01	-7.187e-02
Max	7.879e-02	9.941e-02	5.369e-02	1.322e-01	1.443e-01
Mean	7.361e-04	8.548e-04	1.091e-04	1.087e-03	1.278e-03
Std	1.516e-02	1.490e-02	1.192e-02	1.972e-02	1.955e-02
Skew	-2.187e-01	3.070e-01	5.055e-02	2.407e-01	5.987e-01
Kurt	6.821	1.089e+01	5.586	1.245e+01	9.297

Table 4.4: Daily Logreturn during the period January 2014 - December 2016

⟨tablestat2⟩

In Fig. 4.13 and 4.14 it is shown the comparison between the optimal portfolios of the different methods varying α and the respective τ: the portfolio is composed of the first five assets considering the historical distribution of the logreturns from January 2011 to December 2013 and from January 2014 to December 2016 respectively.

In Fig. 4.15 and 4.16 is represented the portfolio weights of the the optimal portfolios frontier, α is set to 0.2, so the respective τ is 0.025. The analysis is performed again in the periods from January 2011 to December 2013 and from January 2014 to December 2016 respectively.

As we expect, the optimal portfolios frontier is convex in all the cases and there are not substantial differences between the $CVaR_\alpha$ case and the $EVaR_\tau$ case by the fact that the distribution of the assets has low skewness.

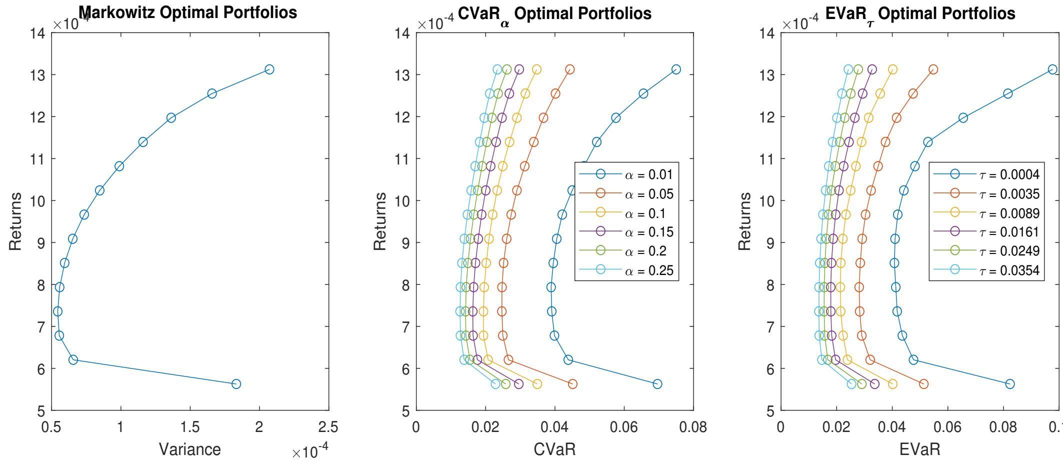


Figure 4.13: Optimal portfolios evaluated with the Markowitz model, the mean-$CVaR_\alpha$ analysis and the mean-$EVaR_\tau$ analysis for different values of α and τ during the period January 2011 December 2013.

⟨fig7⟩

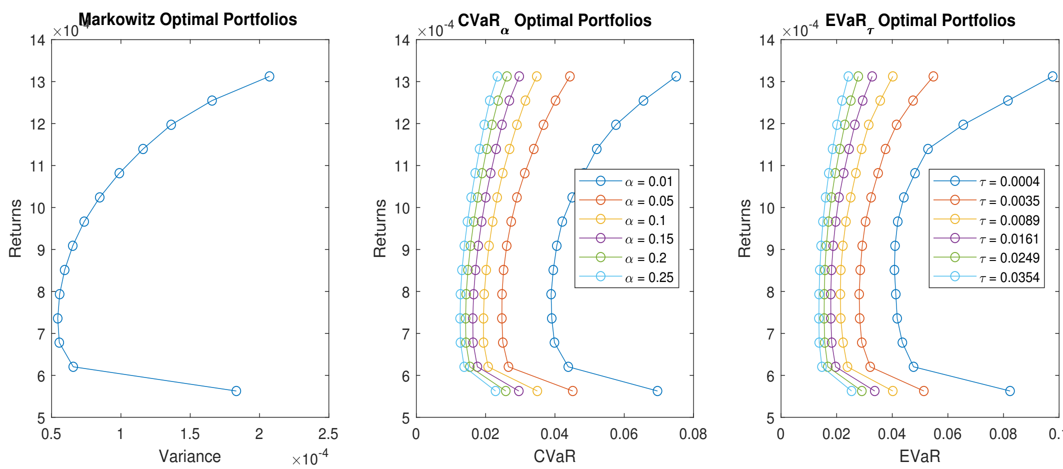


Figure 4.14: Optimal portfolios evaluated with the Markowitz model, the mean-$CVaR_\alpha$ analysis and the mean-$EVaR_\tau$ analysis for different values of α and τ during the period January 2014 December 2016.

⟨fig7a⟩

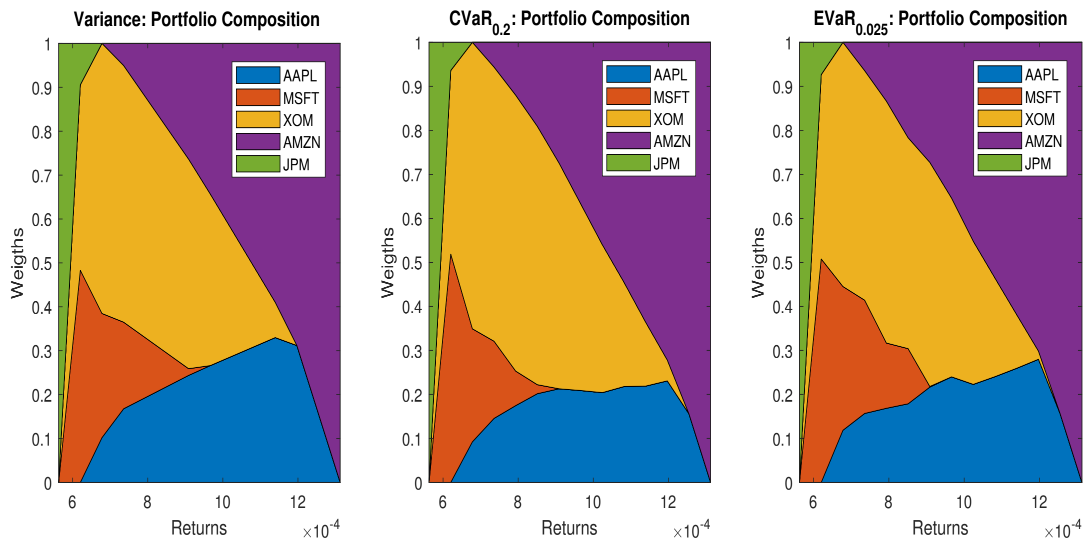


Figure 4.15: Portfolio weights of the optimal portfolios evaluated in the different methods during the period January 2011 December 2013.

⟨fig8⟩

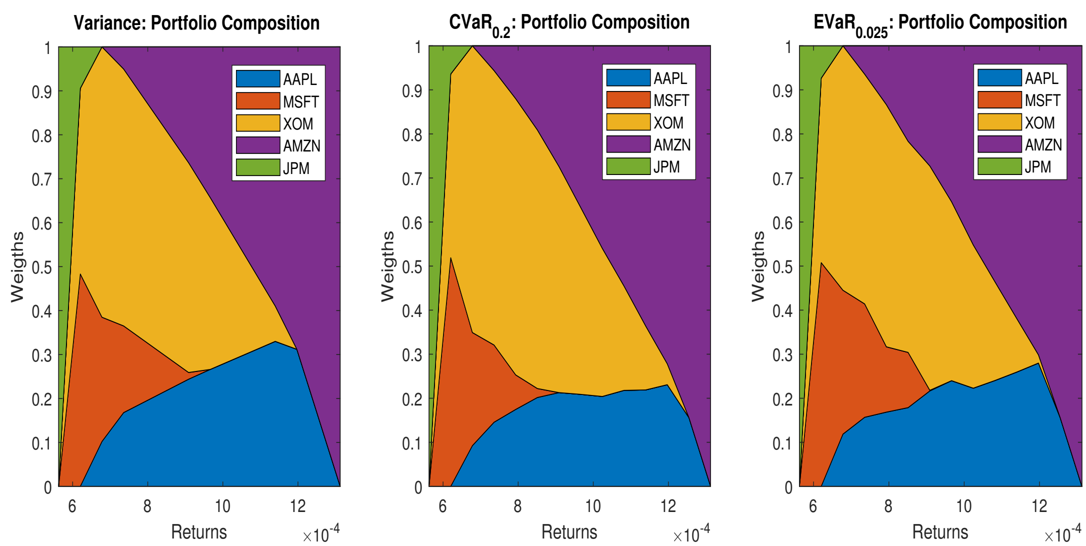


Figure 4.16: Portfolio weights of the optimal portfolios evaluated in the different methods during the period January 2014 December 2016.

⟨fig8a⟩

Let us consider again the portfolio composed by 5 assets. Each day we compute the realized logreturn of the mean-risk portfolio evaluated the day before. We consider variance, $CVaR_{0.2}$, $EVaR_{0.05}$ as risk measures and we compare them with the realization of the equally weighted portfolio.

In order to concentrate ourselves on the upper part of the efficient frontier, in addition to a non negativity constraint on the portfolio weights, we added a lower bound on expected logreturn. Each day, in accord with the formulation in equation (4.3), we impose:

$$\begin{aligned} \mu(w^i) &\geqslant \frac{1}{5}\min_j \mathbb{E}[r^i_j] + \frac{4}{5}\max_j \mathbb{E}[r^i_j] \\ w^i &\geqslant 0, \end{aligned} \tag{4.13}$$

constr1

where $\mathbb{E}[r^i_j]$ is the mean of the logreturns of the asset j evaluated in the rolling window prior to the day i.

The constraint in 4.13 is motivated be the fact usually, in portfolio optimization, we want a minimum logreturn threshold in order to make a profit, in particular we are interested in the high part of the optimal portfolio keeping an acceptable risk (see e.g. Fig 4.17).

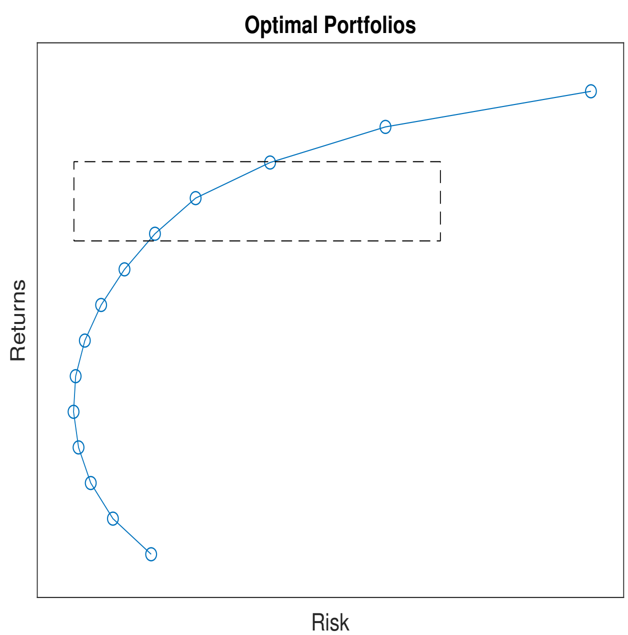


Figure 4.17: Region of target portfolio

⟨targetport⟩

1^{st} January 2011 - 31^{th} December 2013

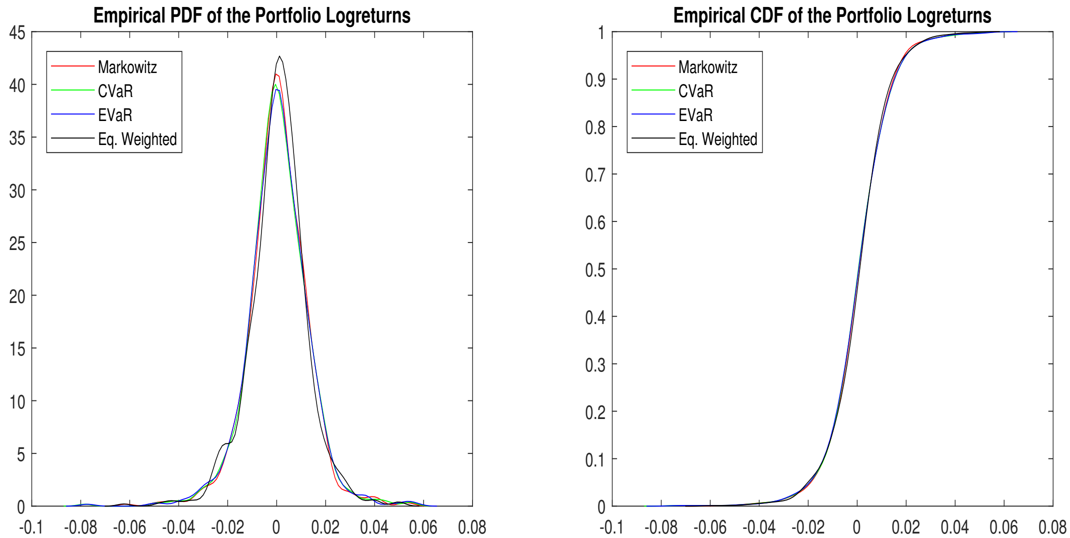


Figure 4.18: Empirical distribution of the 5-asset-portfolio logreturns computed out-of-sample with the rolling window approach over the period Jan11-Dec13: the probability and cumulative density function are smoothed by normal kernels.
⟨1511-13⟩

	Markowitz	**CVaR**	**EVaR**	**Eq. Weighted**
Min	-6.164e-02	-7.828e-02	-7.726e-02	-6.244e-02
Max	5.427e-02	5.458e-02	5.664e-02	5.046e-02
Mean	6.329e-04	6.202e-04	6.033e-04	6.788e-04
Std	1.210e-02	1.246e-02	1.251e-02	1.192e-02
Skew	-1.678e-01	-2.395e-01	-1.928e-01	-2.494e-01
Kurt	5.962	6.964	6.882	5.680

Table 4.5: Statistics of the 5-asset portfolio logreturn distributions over the period January 2014 - December 2016
⟨tablestat3⟩

1^{st} January 2014 - 31^{th} December 2016

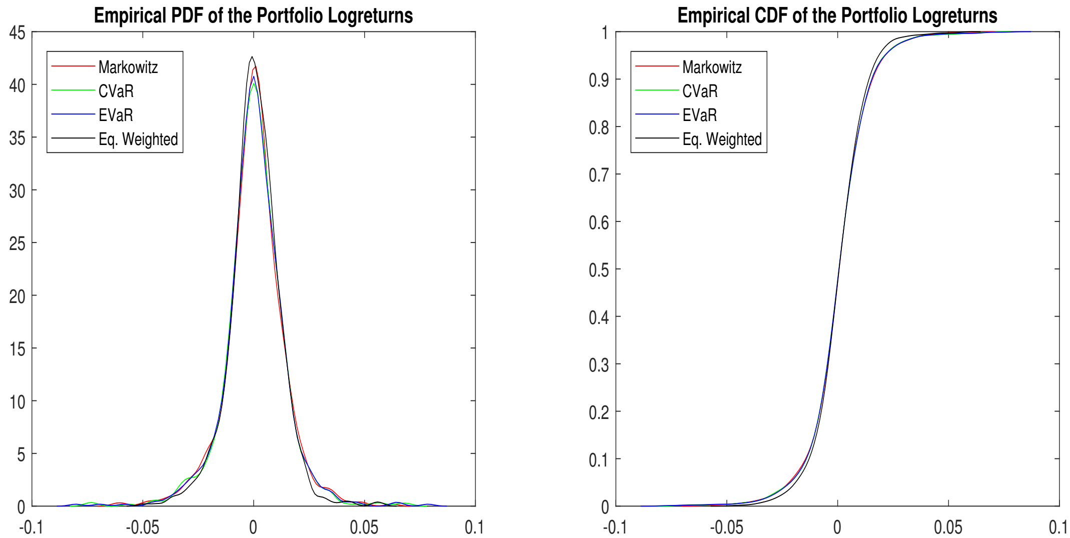


Figure 4.19: Empirical distribution of the 5-asset-portfolio logreturns computed out-of-sample with the rolling window approach over the period Jan14-Dec16: the probability and cumulative density function are smoothed by normal kernels.

⟨1514-16⟩

	Markowitz	**CVaR**	**EVaR**	**Eq. Weighted**
Min	-6.214e-02	-7.443e-02	-8.023e-02	-4.958e-02
Max	6.270e-02	6.968e-02	7.858e-02	5.675e-02
Mean	5.715e-04	5.600e-04	5.983e-04	6.280e-04
Std	1.319e-02	1.357e-02	1.364e-02	1.150e-02
Skew	-2.254e-01	-1.700e-01	-1.14804e-01	-3.00056e-03
Kurt	5.968	7.968	8.622	5.719

Table 4.6: Statistics of the 5-asset portfolio logreturn distributions over the period January 2011 - December 2013

⟨tablestat4⟩

In Fig. 4.18 and 4.19 the empirical distribution of the out-of-sample portfolios' logreturns is represented, and in Tab. 4.5 and 4.6 the extrema and the moments of the distributions are evaluated.

We can observe that the portfolio which minimize the EVaR has performance comparable to the other methods.

In the next experiment we perform a portfolio optimization over a large portfolio in order to assess the stability of the algorithm with a large number of variables and constraints. Considering a portfolio composed by the first 250 asset of the $S\&P$ 500 ordered by their weight, over a rolling window of 100 days, the LP problem in (4.10)

$$
\begin{aligned}
&\min_{u_i,v_i,w_j} y \quad \text{s.t.}\\
&\quad p_i y - u_i + v_i \geqslant p_i(Rw)_i \ \forall i\\
&\quad \sum_{i=1}^{n} u_i - \beta \sum_{i=1}^{n} v_i \geqslant 0\\
&\quad u_i, v_i \geqslant 0,\ \forall i = 1,\dots,n\\
&\quad \sum_{j=1}^{m} w_j = 1\\
&\quad \sum_{i=1}^{m} w_i \mathbb{E}[r_i] = \mu_p\\
&\quad w_j \geqslant 0 \ \forall j = 1,\dots,m
\end{aligned}
$$

is composed by 451 variables, 101 inequality constraints, 2 equality constraints and 250 lower bounds on the variables.

In Fig. 4.20 and 4.21 are shown the comparison between the optimal portfolio's frontiers of the different methods varying α and the respective τ: the portfolio is composed of the first 250 assets considering the historical distribution of the logreturns from January 2011 to December 2013 and from January 2014 to December 2016.
In Fig. 4.22 and 4.23 the empirical distribution of the out-of-sample portfolios' logreturns are represented and in Tab. 4.7 and 4.8 the extrema and the moments of the distributions are evaluated.

Again, the optimal portfolios frontier is convex in all the cases and there are not substantial differences between the $CVaR_\alpha$ case and the $EVaR_\tau$ case: hence our method can be used also for large portfolio problems.
Moreover we can observe that, considering large portfolios, the $EVaR$ method has best performance considering the realized logreturn: another matter should be done for the case of equally weighted portfolio because it is strictly path dependent and does not consider any constraints on risk.

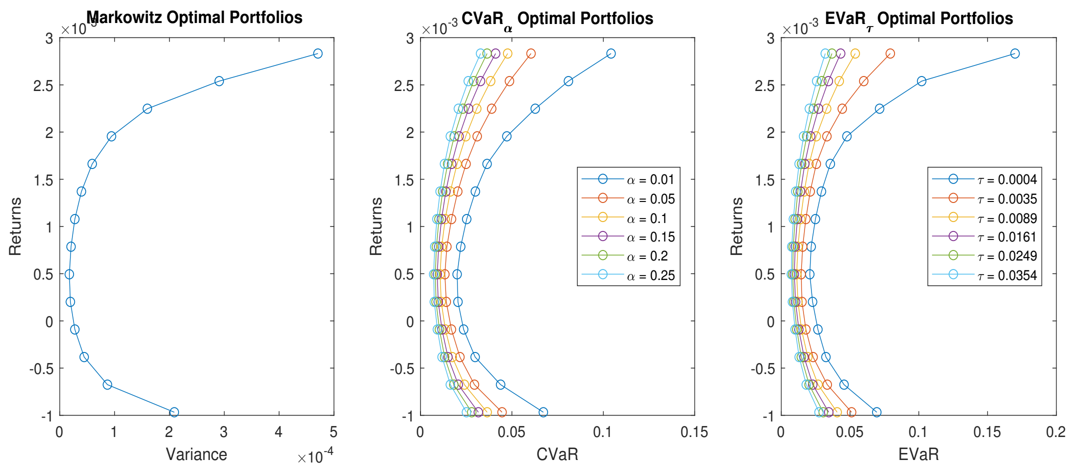


Figure 4.20: Optimal portfolio's frontiers of the portfolios composed of the first 250 assets of the $S\&P$ 500 evaluated with the Markowitz portfolio theory, the mean-$CVaR_\alpha$ analysis and the mean-$EVaR_\tau$ analysis for different values of α and τ over the period Jan2011-Dec2013.

⟨fig10⟩

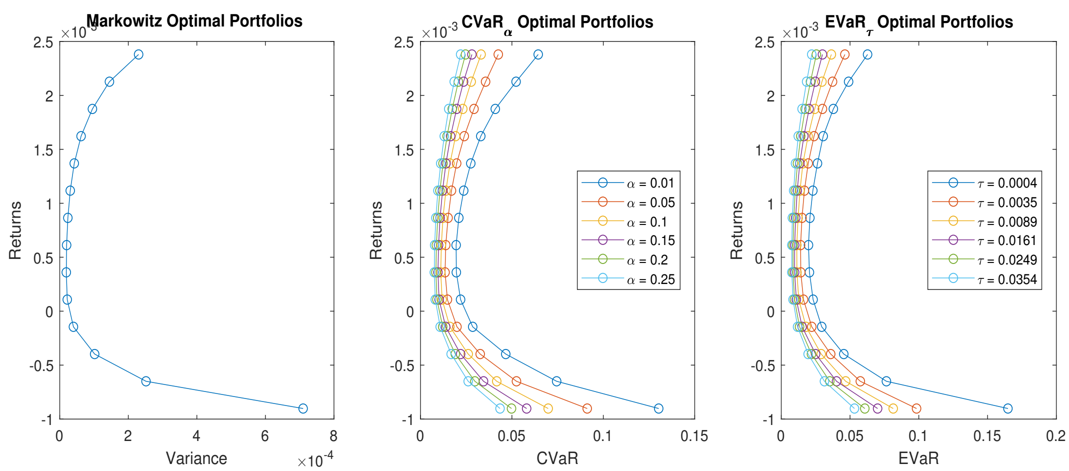


Figure 4.21: Optimal portfolio's frontiers of the portfolios composed of the first 250 assets of the $S\&P$ 500 evaluated with the Markowitz portfolio theory, the mean-$CVaR_\alpha$ analysis and the mean-$EVaR_\tau$ analysis for different values of α and τ over the period Jan2014-Dec2016.

⟨fig10a⟩

1^{st} January 2011 - 31^{th} December 2013

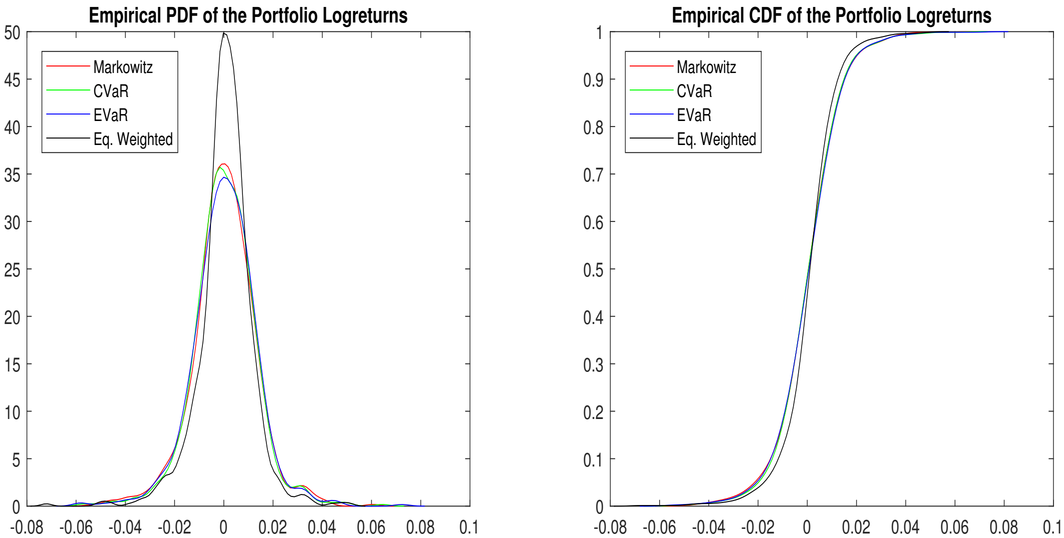


Figure 4.22: Empirical distribution of the 250-asset-portfolio logreturns computed out-of-sample for the period Jan11-Dec13: the probability and cumulative density function are smoothed by normal kernels.

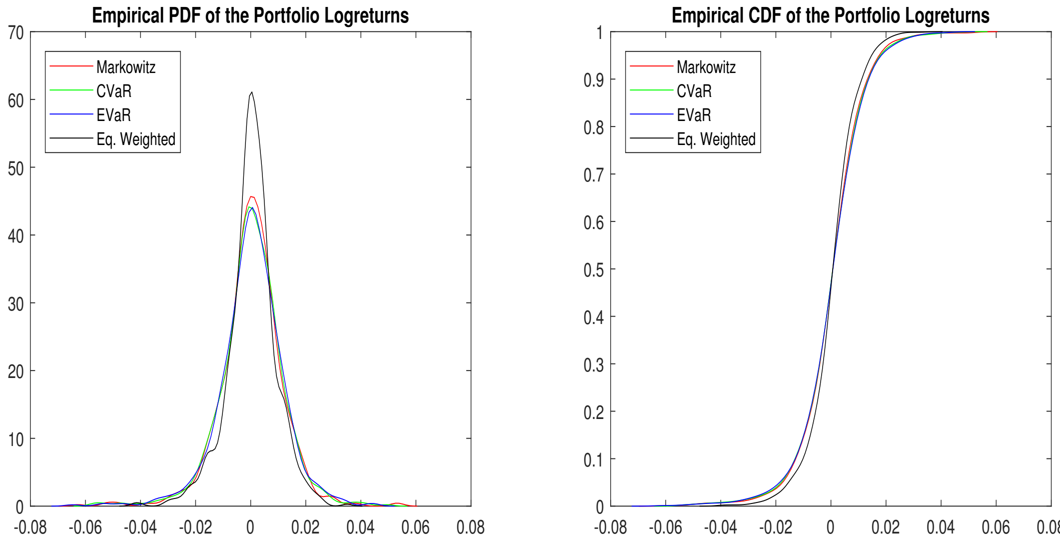
⟨125011-13⟩

	Markowitz	**CVaR**	**EVaR**	**Eq. Weighted**
Min	-5.813e-02	-5.759e-02	-5.882e-02	-7.208e-02
Max	6.010e-02	6.432e-02	7.234e-02	5.109e-02
Mean	2.239e-04	4.394e-04	4.560e-04	6.836e-04
Std	1.298e-02	1.267e-02	1.295e-02	1.100e-02
Skew	-2.551e-01	-1.796e-02	-8.106e-02	-5.835e-01
Kurt	5.306	5.824	6.062	8.416

Table 4.7: Statistics of the 255-asset portfolio logreturn distributions over the period January 2011 - December 2013

⟨tablestat5⟩

1^{st} January 2014 - 31^{th} December 2016



[H]

Figure 4.23: Empirical distribution of the 250-asset-portfolio logreturns computed out-of-sample for the period Jan14-Dec16: the probability and cumulative density function are smoothed by normal kernels.

⟨125014-16⟩

	Markowitz	**CVaR**	**EVaR**	**Eq. Weighted**
Min	-6.272e-02	-5.684e-02	-6.462e-02	-4.216e-02
Max	5.338e-02	4.939e-02	4.452e-02	3.503e-02
Mean	3.058e-04	3.627e-04	3.930e-04	4.118e-04
Std	1.140e-02	1.151e-02	1.162e-02	8.810e-03
Skew	-4.386e-01	-5.22420e-01	-6.00888e-01	-4.358e-01
Kurt	7.183	6.497	6.499	5.137

Table 4.8: Statistics of the 255-asset portfolio logreturn distributions over the period January 2014 - December 2016

⟨tablestat6⟩

Summarizing, we introduced the Expectile Value at Risk at level τ as a new risk measure and we proved that finding the portfolio which minimize the $EVaR_\tau$ is a linear programming problem with $2n + d + 1$ variables and $n + 1$ constraints, where n is the length of the dataset and d is the number of the assets.

During the simulation we evinced that the EVaR_τ behaves similarly to the CVaR_α when we consider symmetric distributions: on the other hand, considering distributions with the same left tail but different right tail, CVaR_α does not distinguish the distribution focusing only on the left tail while EVaR_τ prefers distribution with a greater right tail.

In the out-of-sample tests we did not find substantial differences in using $EVaR_\tau$, $CVaR_\alpha$ and variance as quantities to minimize in order to allocate small portfolios: the distributions of the assets are too similar to notice a difference. Considering large portfolios the mean-$EVaR_\tau$ approach seems to have better performances, in fact we have a greater realized logreturn with a comparable variance.

Chapter 5

Interexpectile Difference as Deviation Measure

⟨interexpectile⟩ In the recent paper by Bellini et al. (2018), it has been suggested that the interexpectile difference might be an interesting measure of variability. Indeed, in comparison with the very well known interquartile or more generally interquantile differences, it has the additional property of being consistent with respect to the convex order (also known as Second Order Stochastic Dominance with equal means), which is a highly desirable property for a deviation measure as we pointed out in Example 30 at Section 4.1.

⟨interexp⟩ **Definition 31.** *We define*

$$\Delta_\tau = e_{1-\tau}(X) - e_\tau(X) = e_{1-\tau}(X) + e_{1-\tau}(-X) = -e_\tau(-X) - e_\tau(X),$$

for $X \in L^1$ *and* $\tau \in (0, 1/2)$.

The following properties can be proved straightforwardly from the corresponding properties of expectiles (see Bellini et al. (2017))

Proposition 32. *Let* $X \in L^1$, $\tau \in (0, 1/2)$ *and* Δ_τ *as in (31). Then:*

a) $\Delta_\tau(X) \geqslant 0$ *and* $\Delta_\tau(X) = 0$ *if and only if* $X = c$ *P-a.s.*

b) $\Delta_\tau(X) = \Delta_\tau(-X)$

c) $\Delta_\tau(X)$ *is strictly increasing and continuous in* τ

d) $\Delta_\tau(X) \to 0^+$ *for* $\tau \to 1/2^-$ *and* $\Delta_\tau(X) \to \operatorname{ess\,sup}(X) - \operatorname{ess\,inf}(X)$ *for* $\tau \to 0^+$

e) $\Delta_\tau(X+h) = \Delta_\tau(X)$, *for* $h \in \mathbb{R}$

f) $\Delta_\tau(\lambda X) = \lambda \Delta_\tau(X)$, *for* $\lambda \geqslant 0$

g) $\Delta_\tau(X+Y) \leqslant \Delta_\tau(X) + \Delta_\tau(Y)$

h) $X \leqslant_{cx} Y \Rightarrow \Delta_\tau(X) \leqslant \Delta_\tau(Y)$,

where $\leqslant_{cx}$ *denotes the usual convex order.*

Recalling that $EVaR_\tau(X) = \max_{\varphi \in \mathcal{M}_\tau} \mathbb{E}[-X\varphi]$, we can write

$$\begin{aligned}
\Delta_\tau(X) &= -e_\tau(-X) - e_\tau(X) = EVaR_\tau(-X) + EVaR_\tau(X) \\
&= \max_{\varphi \in \mathcal{M}_\tau} \mathbb{E}[X\varphi] + \max_{\varphi \in \mathcal{M}_\tau} \mathbb{E}[-X\varphi], \qquad (5.1) \ \boxed{\texttt{deltatau}} \\
\mathcal{M}_\tau &= \left\{ \varphi \in L^\infty,\ \varphi \geqslant 0 \text{ a.s.}, \mathbb{E}[\varphi] = 1, \frac{\operatorname{ess\,sup} \varphi}{\operatorname{ess\,inf} \varphi} \leqslant \frac{1-\tau}{\tau} \right\}.
\end{aligned}$$

5.1 Linear Programming Formulation

By discretizing the maximization problem in (5.1) and assuming that X is a discrete random variable with $\mathbb{P}(X = x_i) = p_i$ for $i = 1, \dots, n$, we obtain:

$$\Delta_\tau(X) = \left\{ \begin{array}{lcl}
\max\limits_{f_i, m} \sum\limits_{i=1}^n x_i f_i p_i \quad \text{s.t.} & & \max\limits_{f_i, m} \sum\limits_{i=1}^n -x_i f_i p_i \quad \text{s.t.} \\
\sum\limits_{i=1}^n f_i p_i = 1 & + & \sum\limits_{i=1}^n f_i p_i = 1 \\
m \leqslant f_i \leqslant \beta m,\ \forall i & & m \leqslant f_i \leqslant \beta m,\ \forall i \\
m \geqslant 0,\ f_i \geqslant 0, \forall i = 1, \dots, n & & m \geqslant 0,\ f_i \geqslant 0, \forall i = 1, \dots, n
\end{array} \right\}$$

The corresponding dual is:

$$\Delta_\tau(X) = \left\{ \begin{array}{lcl}
\min\limits_{u_i, v_i} y \quad \text{s.t.} & & \min\limits_{u_i, v_i} y \quad \text{s.t.} \\
p_i y - u_i + v_i \geqslant -p_i x_i\ \forall i & & p_i y - u_i + v_i \geqslant p_i x_i\ \forall i \\
\sum\limits_{i=1}^n u_i - \beta \sum\limits_{i=1}^n v_i \geqslant 0 & + & \sum\limits_{i=1}^n u_i - \beta \sum\limits_{i=1}^n v_i \geqslant 0 \\
u_i, v_i \geqslant 0, \forall i = 1, \dots, n & & u_i, v_i \geqslant 0, \forall i = 1, \dots, n
\end{array} \right\}$$

Grouping the minimum functions and renaming $(y,\ u,\ v) = (y^1,\ u^1,\ v^1)$ in the first minimization problem and $(y,\ u,\ v) = (y^2,\ u^2,\ v^2)$ in the second one, we obtain that the Interexpectiles Difference of a discrete random variable X is given by:

$$
\begin{aligned}
&\Delta_\tau = \min_{u_i,v_i} y^1 + y^2 \quad \text{s.t.}\\
&p_i^1 y^1 - u_i^1 + v_i^1 \geqslant -p_i^1 x_i \ \forall i\\
&\sum_{i=1}^n u_i^1 - \beta \sum_{i=1}^n v_i^1 \geqslant 0\\
&p_i^2 y^2 - u_i^2 + v_i^2 \geqslant p_i^2 x_i \ \forall i\\
&\sum_{i=1}^n u_i^2 - \beta \sum_{i=1}^n v_i^2 \geqslant 0\\
&u_i^1, v_i^1, u_i^2, v_i^2 \geqslant 0, \ \forall i = 1, \dots, n
\end{aligned}
\tag{5.2}
$$

DeltaTau

5.1.1 Portfolio Problem

Suppose now that the outcomes of the random variable X are portfolio logreturns, so we have that $X = Rw$, where R is the matrix of the historical logreturn and $w = (w_1, \dots, w_m)$ are the weights. Finding the portfolio which minimize the Δ_τ means finding w such that

$$
\begin{aligned}
w \;=\; & \operatorname*{argmin}_{w} \Delta_\tau(Rw) \quad s.t.\\
& \sum w = 1\\
& \text{portfolio constraints}
\end{aligned}
$$

The problem in (5.2) becomes

$$
\begin{aligned}
&w = \operatorname*{argmin}_{w_j}\{\min_{u_i,v_i} y^1 + y^2\} \quad \text{s.t.}\\
&p_i^1 y^1 - u_i^1 + v_i^1 \geqslant -p_i(R\overline{w})_i \ \forall i = 1, \dots, n\\
&\sum_{i=1}^n u_i^1 - \beta \sum_{i=1}^n v_i^1 \geqslant 0\\
&p_i^2 y^2 - u_i^2 + v_i^2 \geqslant p_i(R\overline{w})_i \ \forall i = 1, \dots, n\\
&\sum_{i=1}^n u_i^2 - \beta \sum_{i=1}^n v_i^2 \geqslant 0\\
&u_i^1, v_i^1, u_i^2, v_i^2 \geqslant 0, \ \forall i = 1, \dots, n\\
&\left.\begin{array}{l}
\sum_{j=1}^m w_j = 1\\
\text{Shortselling} \quad w_j \geqslant 0 \ \forall j = 1, \dots, m\\
\text{Efficient Frontier} \quad \sum_{i=1}^m w_i \mathbb{E}[r_i] = \mu_p
\end{array}\right\} \Longrightarrow \text{Portfolio Constraints}
\end{aligned}
$$

5.2 Numerical Examples

In the first part of this section we consider optimal portfolios on simulated data: our purpose is to compare the portfolio problem using the Δ_τ as risk measure with the usual risk measures used in literature. In the second part of the section we test the method on historical data.
All the experiments are performed using MatLab and the optimization problems are solved with the functions `linprog`, `quadprog` and `fmincon`. The solutions to Linear and Quadratic Programming problems are checked with GAMS (General Algebraic Modeling System).

5.2.1 Simulated Data

As toy problem we consider again the pattern of section 4.4.1, i.e. a portfolio composed of three assets varying the distributions and the correlations matrix. In the first example we suppose that the logreturn of three assets A, B and, C are described by uncorrelated normal distributions: in order to assess the variability of the optimal portfolio, we simulate samples of length $T = 20, 50, 100$ and compute the distribution of minimal Δ_τ portfolios for τ ranging from 5% to 45%. The results are reported in Fig. 5.1; the three lines correspond, respectively from top to bottom, to the 10^{th} percentile, the median and the 90^{th} percentile of the portfolio weights.
If $iqr_{0.25}$ is the well-known interquartile range, we will consider its generalization i.e. the interquantile range of level alpha defined as $iqr_\alpha = q_{1-\alpha} - q_\alpha$. Since minimizing iqr_α is a NLP problem, in order to find a solution we used MatLab `fmincon` with objective function

```
prctile(x,1-alpha)-prctile(x,alpha).
```

In Fig. 5.2 we compare the portfolios originated by the minimization of variance, $\text{CVaR}_{0.20}$, $iqr_{0.25}$ and $\Delta_{0.25}$ through the scatter plot of the weights of A and B and the relative histogram (the weight of C is given by $1 - w_A - w_B$).
In Fig. 5.3 and 5.4 we introduce correlation among the assets: A, B and C are described by correlated normal distribution with correlation matrix

$$\Sigma_+ = \begin{pmatrix} 1 & 0.3 & 0.5 \\ 0.3 & 1 & 0.7 \\ 0.5 & 0.7 & 1 \end{pmatrix} \quad \Sigma_- = \begin{pmatrix} 1 & -0.3 & 0.5 \\ -0.3 & 1 & -0.7 \\ 0.5 & -0.7 & 1 \end{pmatrix} \tag{5.3}$$?corrmat2?

Again in Fig. 5.5 and 5.6 we consider uncorrelated asymmetric distributions of the following types: asset A and B have a normal distribution, asset C has a reflected standardized gamma distribution.
The parameters of the distributions are the same as eq. 4.12 i.e.:

$$\begin{aligned}
&\text{Asset A} \sim \mathcal{N}(\mu_{\mathcal{N}}, \sigma_{\mathcal{N}}), && \text{with} \;\; \mu_{\mathcal{N}} = 0, \;\; \sigma_{\mathcal{N}} = 2.25 \\
&\text{Asset B} \sim \mathcal{N}(\mu_{\mathcal{N}}, \sigma_{\mathcal{N}}), && \text{with} \;\; \mu_{\mathcal{N}} = 0, \;\; \sigma_{\mathcal{N}} = 2.25 \\
&\text{Asset C} \sim -\Gamma(k, \theta, \mu_{\Gamma}, \sigma_{\Gamma}), && \text{with} \;\; k = 5, \;\; \theta = 1, \;\; \mu_{\Gamma} = -0.912, \sigma_{\Gamma} = 1.5
\end{aligned}$$

As in section 4.4.1, the assets have a similar left tail with equal $\text{VaR}_{0.05}$ but different right tail (Fig. 4.2).
Finally in Fig. 5.7 and 5.8 we introduce correlation among the asset: we perform tests with correlation matrices Σ_+ and Σ_-.

As in section 4.4.1, we want to reproduce the experiment performed by Bertsimas et al. (2004) in order to study how the portfolio change with respect to the level τ and to compare the resalts with other well-known variability risk measures such as variance and interquantile difference.
In Fig. 5.1, 5.3, 5.5 and 5.7 we can observe that in all cases the variability of the portfolio is similar to the portfolio found minimizing variance, so we do not expect substantial differences in the portfolio weights varying τ or considering the Markowitz portfolio.

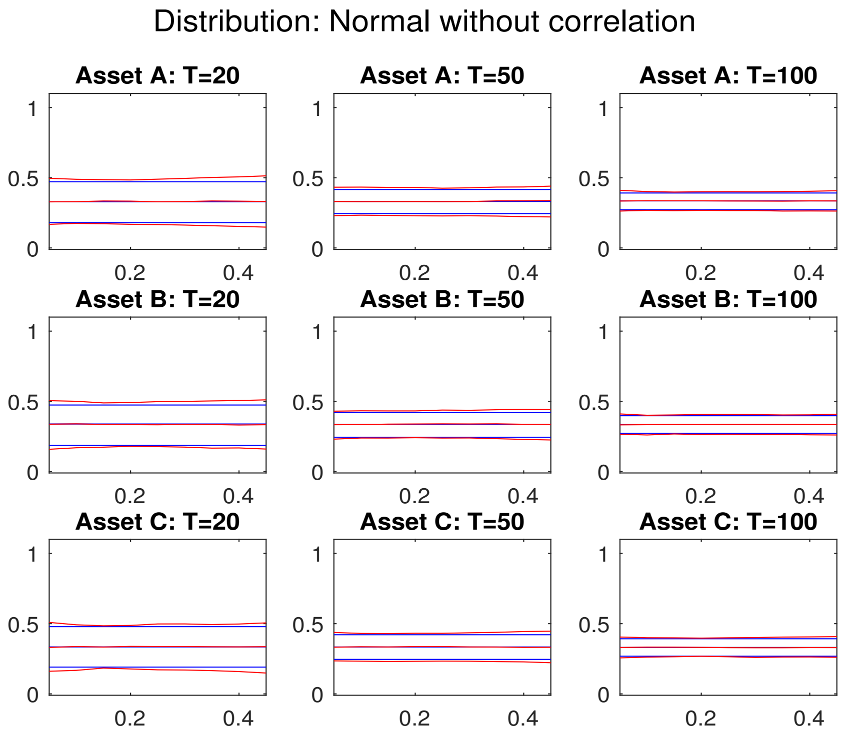


Figure 5.1: Variability of the weights of minimum Δ_τ portfolios for $\tau = 0.05, \ldots, 0.45$ on simulated uncorrelated standard normal data of length $T = 20, 50, 100$. The red lines indicate respectively the 90^{th} percentile, the median and the 10^{th} percentile of the portfolio weights. The blue lines represent the same quantities for a minimal variance portfolio.

⟨3N-uncorr⟩

In Fig. 5.2, 5.4, 5.6 and 5.8 we compare the portfolio found minimizing variance, minimizing interquartile range, minimizing $EVaR_{0.05}$ and minimizing $\Delta_{0.25}$.

In the uncorrelated normal case (Fig. 5.2) we can observe that the dispersion of the cloud of points of $\Delta_{0.25}$ case is again comparable with the variance case and has less variability than the $EVaR_{0.05}$ and $iqr_{0.25}$ cases. Despite the high dispersion, the portfolio evaluated minimizing $iqr_{0.25}$ is more concentrated to the center of the distribution.

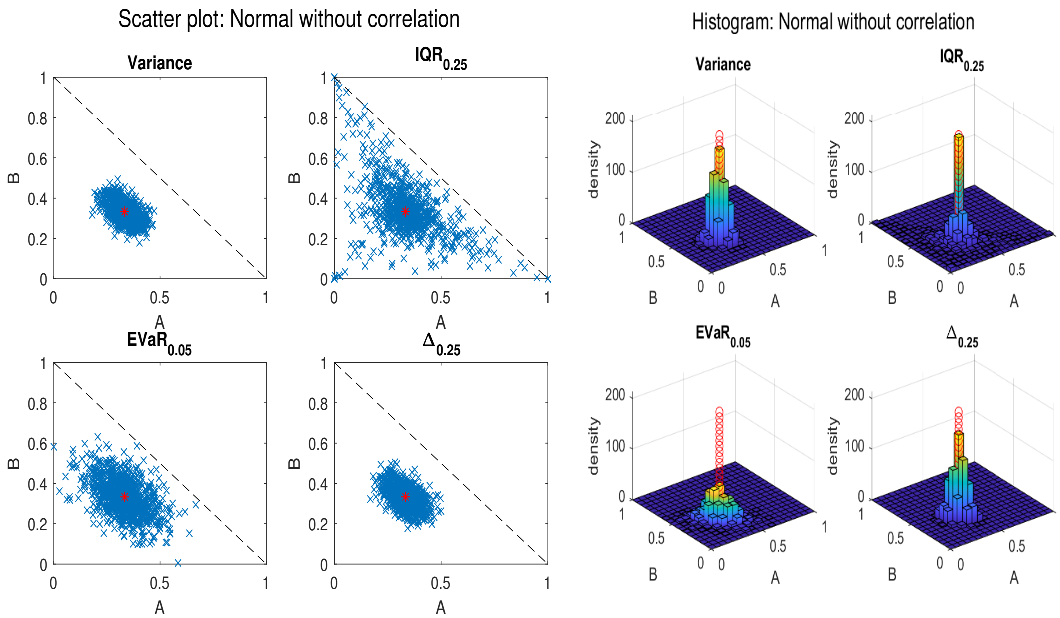


Figure 5.2: Scatter plot and histogram of the weights of the assets A and B, the weight of C is given by the relation $\sum \omega_i = 1$, the pattern is the same of Fig. 5.1 with $T = 100$. The red points represent the theoretical value of the portfolio weights given by the minimal variance portfolio.

⟨3N-uncorr_hist⟩

Introducing correlation among the asset logreturns, in Fig. 5.3 we can observe that the Δ_τ minimization behaves again as the variance minimization: varying τ and the sample size, the 10^{th} percentiles, the 90^{th} percentiles and the medians almost coincide.

Hence, when we consider correlated normal distributions, minimizing Δ_τ gives results extremely similar to the Markowitz portfolio.

Finally, considering the sample size, the variability decreases for higher value of T, so the algorithm seems pretty stable.

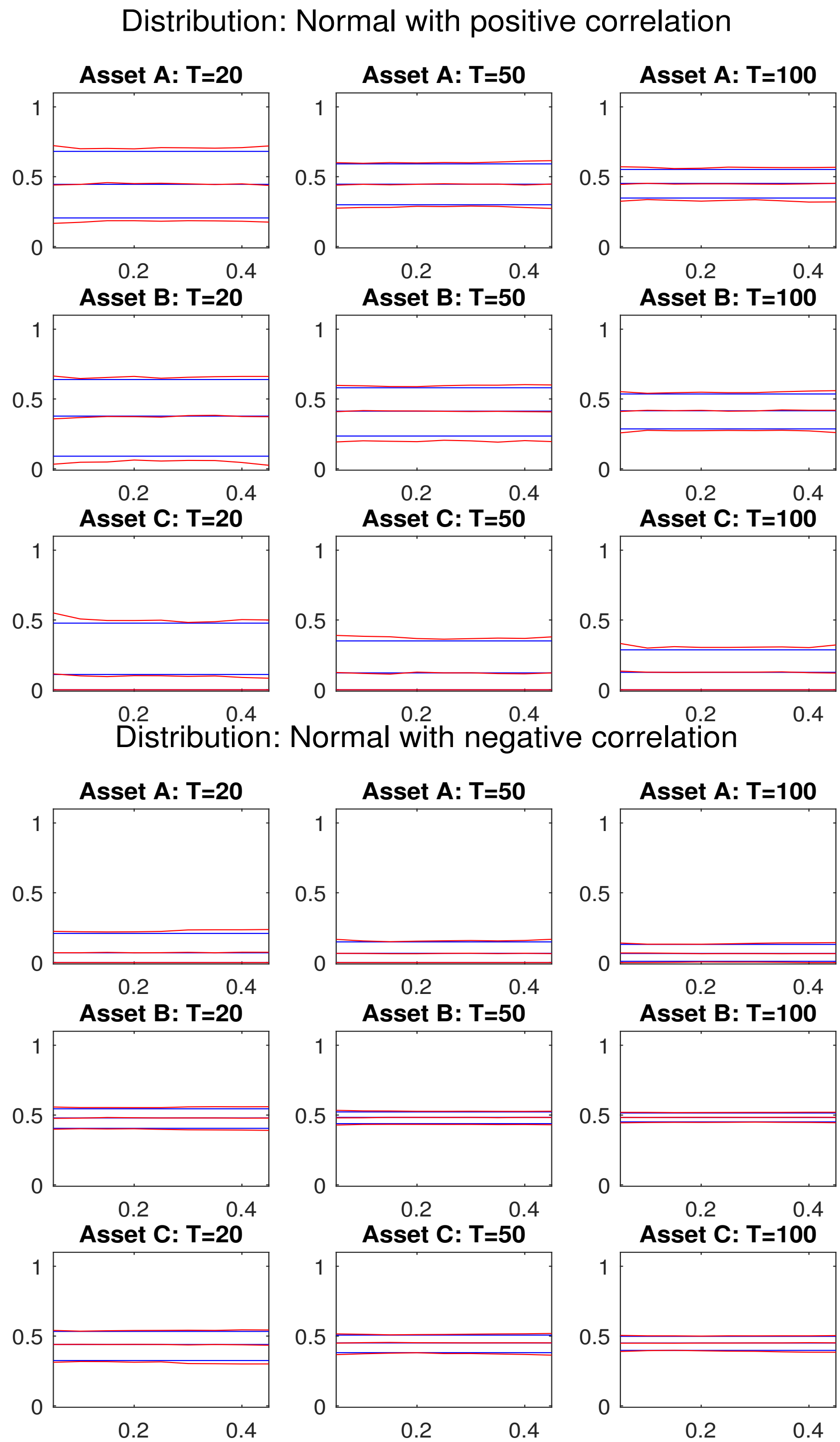


Figure 5.3: Variability of the weights of minimum Δ_τ portfolios for $\tau = 0.05, \ldots, 0.45$ on simulated correlated standard normal data of length $T = 20, 50, 100$. The red lines indicate respectively the 90^{th} percentile, the median and the 10^{th} percentile of the portfolio weights. The blue lines represent the same quantities for a minimal variance portfolio.

⟨3N-corr⟩

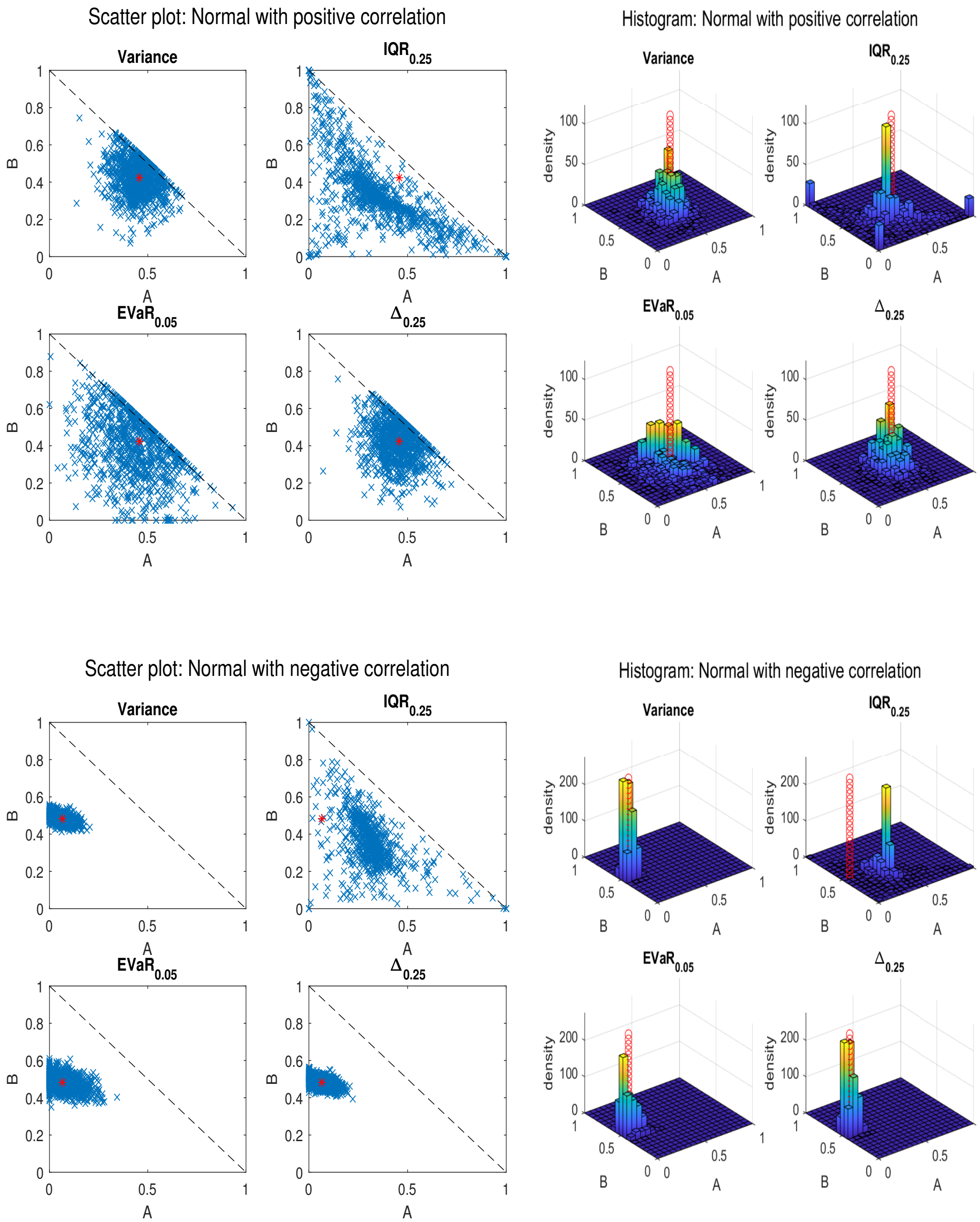


Figure 5.4: Scatter plot and histogram of the weights of the assets A and B, the weight of C is given by the relation $\sum \omega_i = 1$, the pattern is the same of Fig. 5.3 with 1000 simulations. The red points represent the theoretical value of the portfolio weights given by the minimal variance portfolio.

⟨3N-corr_hist⟩

Considering $\tau = 0.25$ and comparing $\Delta_{0.25}$ with variance, $EVaR_{0.05}$ and $iqr_{0.25}$ (Fig. 5.4) we can observe that:

- when we consider the correlation matrix Σ_+ in the $EVaR_{0.05}$ case the cloud of points lays mostly on the diagonal, on the contrary minimizing the interquartile range we find portfolios that are close to the center of the graph or that gives all the weight to a single asset;

- when we consider the correlation matrix Σ_- variance, Δ_τ and the $EVaR_{0.05}$ behave similarly indeed the cloud of points lays mostly on the y-axis giving weight to asset B and C avoiding A, on the contrary minimizing the interquartile range we find again portfolios that are close to center of the graph.

Observing the histogram of the portfolio generated minimizing the variance, we evince that it is very similar to the histogram of the portfolio generated minimizing Δ_τ.

Minimizing $iqr_{0.25}$ is another matter entirely: negative correlation seems not to affect the portfolio composition, while positive correlation produces some portfolios with all the weight concentrated in a single asset.
These phenomena can be explained by the lack of the subadditivity property of $iqr_{0.25}$, so diversification does not necessary decrease the risk.

Let now consider uncorrelated asymmetrical distributions (Fig. 4.2): asset A and B again have a normal distribution with $\mu_{\mathcal{N}} = 0$ and $\sigma_{\mathcal{N}} = 2.25$, asset C has a reflected Γ distribution with parameters $k = 5,\ \theta = 1,\ \mu_\Gamma = -0.912,\ \sigma_\Gamma = 1.5$. In Fig. 5.5 we can observe the same results obtained in the standard normal case with and without correlation, i.e.:

- the 10^{th} percentiles, the 90^{th} percentiles and the medians almost coincide in all the cases;

- the portfolio composition remain unchanged varying τ;

- the variability of the portfolio decrease when the sample size increase.

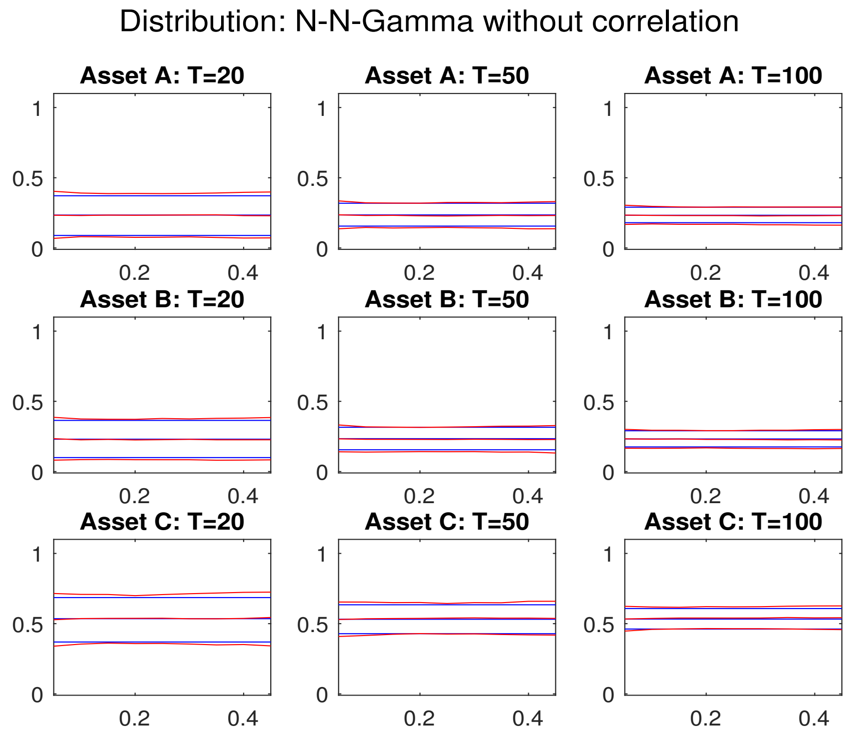


Figure 5.5: Variability of the weights of minimum Δ_τ portfolios for $\tau = 0.05, \ldots, 0.45$ on simulated uncorrelated data of length $T = 20, 50, 100$: the assets' distributions are described in Eq. 4.12. The red lines indicate respectively the 90$^{\text{th}}$ percentile, the median and the 10$^{\text{th}}$ percentile of the portfolio weights. The blue lines represent the same quantities for a minimal variance portfolio.

⟨3NGG-uncorr⟩

From Fig. 5.6 we can observe that the histogram generated minimizing $\Delta_{0.25}$ is again very similar to the histogram generated minimizing variance and differs from the histogram generated minimizing $EVaR_{0.05}$ which gather on the diagonal. The case of $iqr_{0.25}$ does not differ from the standard normal case.

Finally, we introduce correlation among the data: we consider again separately the correlation matrices Σ_+ and Σ_-. In Fig. 5.7 we observe again that the portfolios generated minimizing Δ_τ and variance are extremely similar, that varying τ the portfolio composition does not change and that the variability decreases when the sample size increase.

Considering $\tau = 0.25$ and comparing $\Delta_{0.25}$ with variance, $EVaR_{0.05}$ and $iqr_{0.25}$ (Fig. 5.8) we can observe that:

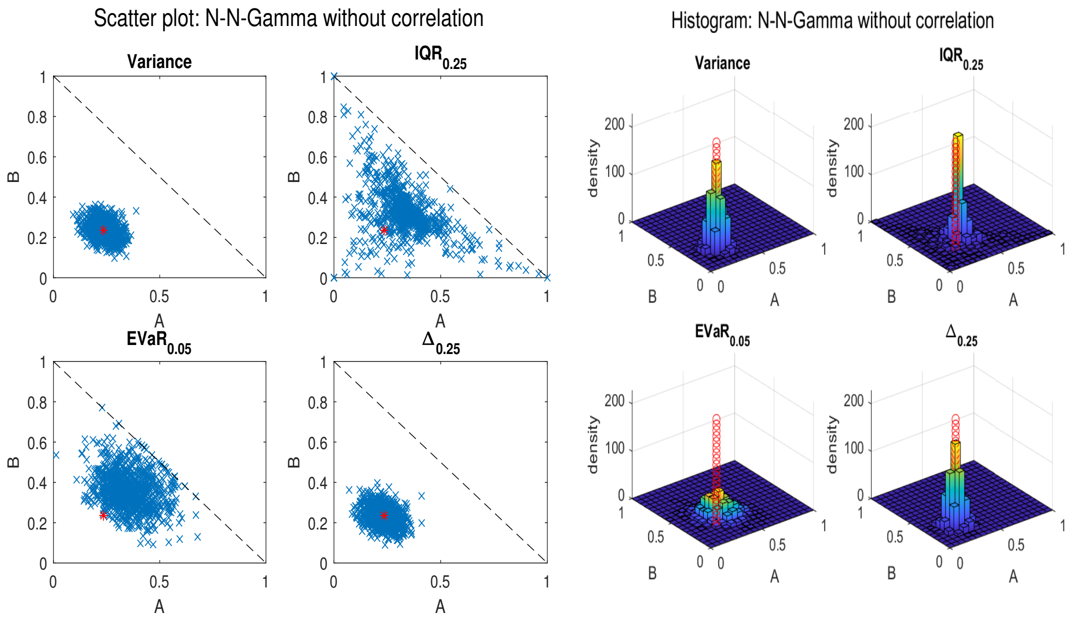


Figure 5.6: Scatter plot and histogram of the weights of the assets A and B, the weight of C is given by the relation $\sum \omega_i = 1$, the pattern is the same of Fig. 5.5 with 1000 simulations. The red points represent the theoretical value of the portfolio weights given by the minimal variance portfolio.

⟨3NGG-uncorr_hist⟩

- when Σ_+ is considered, minimizing $EVaR_{0.05}$ the cloud of points lays mostly on the diagonal, minimizing variance and Δ_τ we find portfolios that lays on the x-axis;

- when Σ_- is considered, minimizing $EVaR_{0.05}$ we find portfolios that are coser to the center of the graph, minimizing variance and in Δ_τ the cloud of points lays mostly on the on the y-axis.

The portfolio found minimizing $iqr_{0.25}$ again does not differ from the correlated normal case.

In conclusion, considering all the asset's distributions presented in this section, we observe that the portfolios generated minimizing Δ_τ are extremely similar to the portfolios generated minimizing variance. As we expected from theory, minimizing Δ_τ consists in minimizing the quantity $-e_\tau(-X) - e_\tau(X)$, hence we are minimizing the risk measure both the gain and the loss of the distribution, which is very similar to minimizing variance. Furthermore since minimizing Δ_τ is a LP problem, when we consider large portfolios our method is preferable to minimizing variance from a computational point of view.

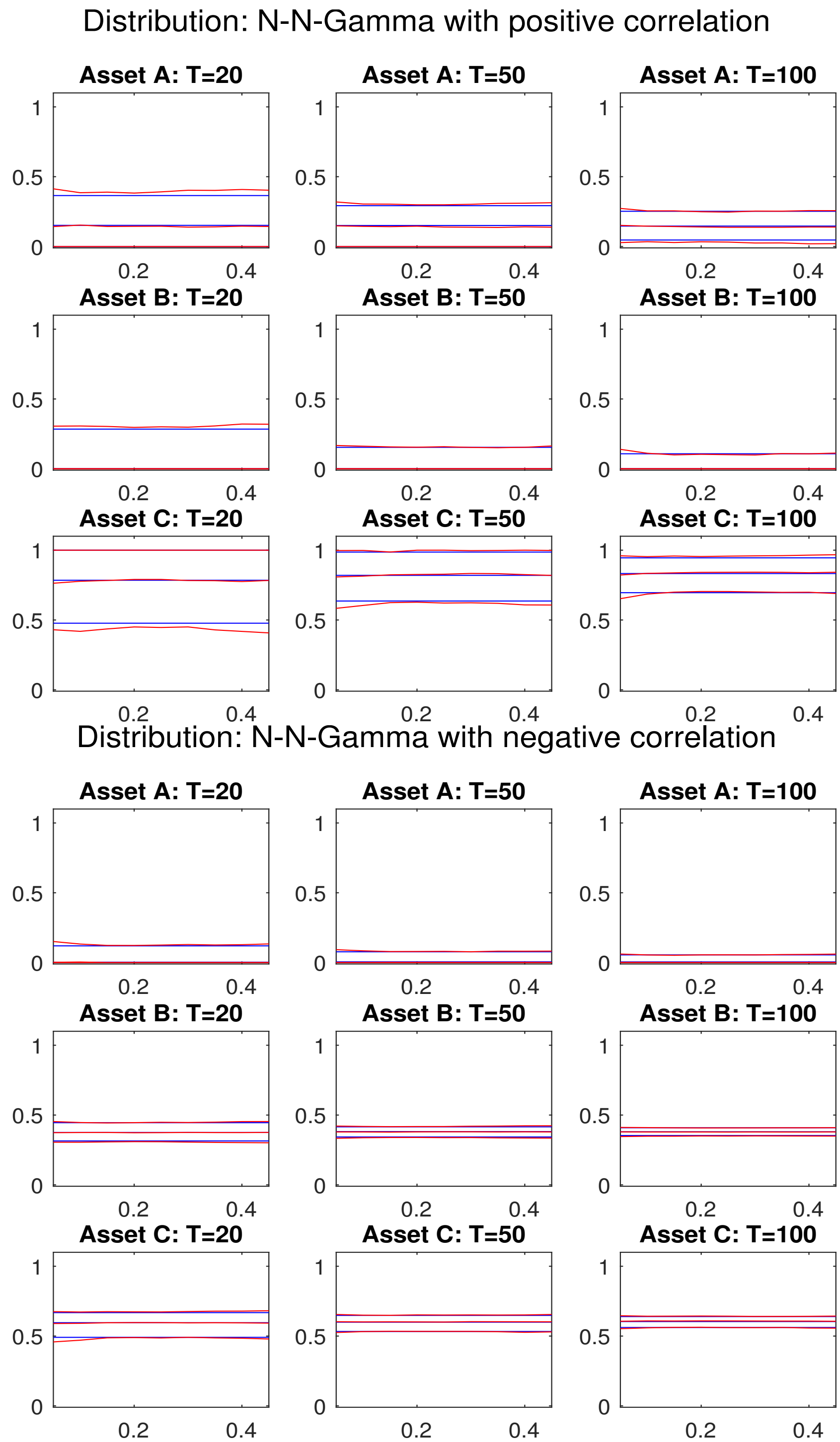


Figure 5.7: Variability of the weights of minimum Δ_τ portfolios for $\tau = 0.05, \ldots, 0.45$ on simulated correlated data of length $T = 20, 50, 100$ with correlation matrices Σ_+ and Σ_-: the assets' distributions are described in Eq. 4.12. The red lines indicate respectively the 90$^{\text{th}}$ percentile, the median and the 10$^{\text{th}}$ percentile of the portfolio weights. The blue lines represent the same quantities for a minimal variance portfolio.

⟨3NGG-corr⟩

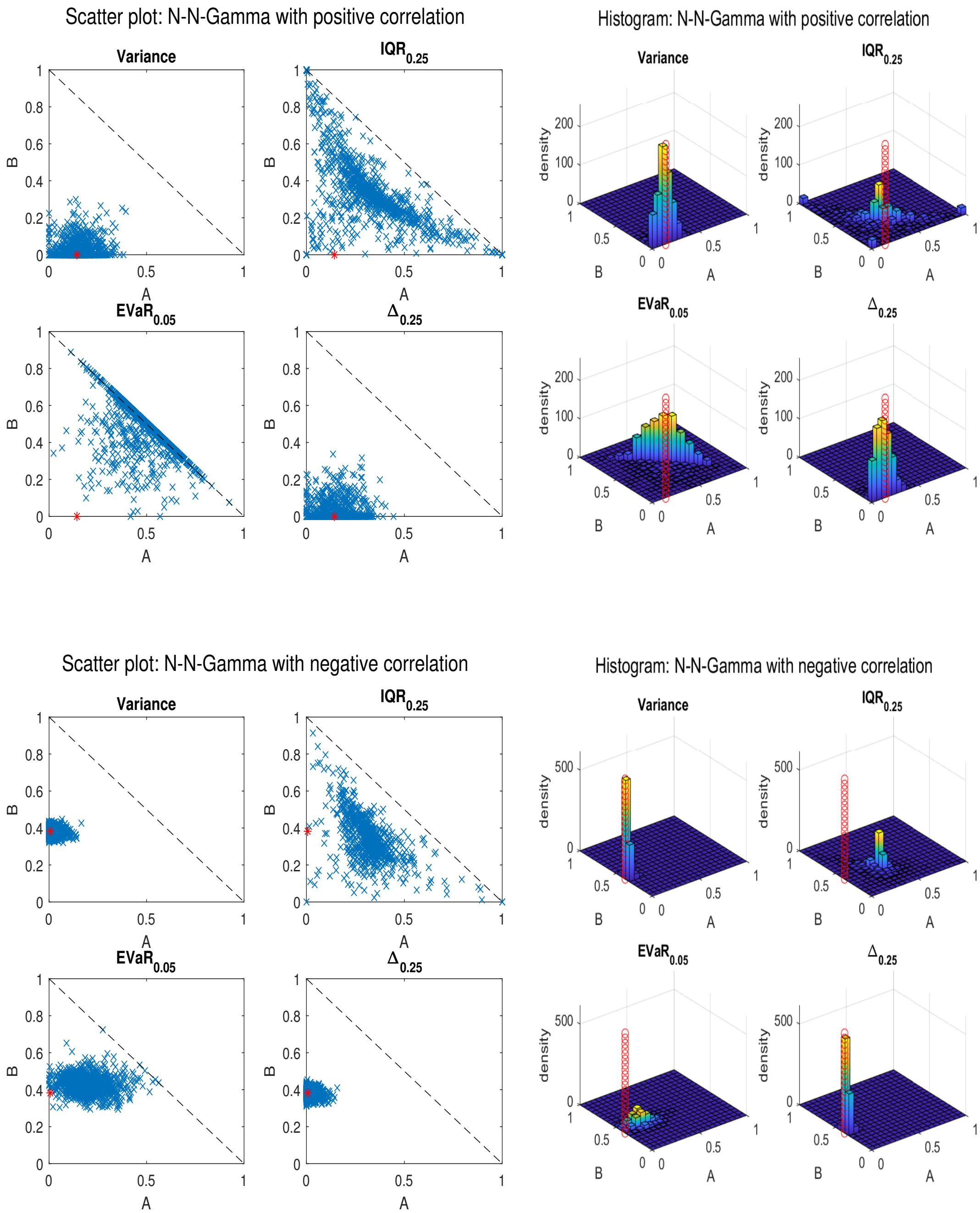


Figure 5.8: Scatter plot and histogram of the weights of the assets A and B, the weight of C is given by the relation $\sum \omega_i = 1$, the pattern is the same of Fig. 5.7 with 1000 simulations. The red points represent the theoretical value of the portfolio weights given by the minimal variance portfolio.

⟨3NGG-corr_hist⟩

5.2.2 Historical Data

stInterexpectile⟩

In this section we perform tests on Δ_τ using the same dataset of section 4.4.2, i.e. using the logreturns of the companies that constitute the $S\&P$ 500 index in two different periods: January 2011 - December 2013 and January 2014 - December 2016.

First we evaluate the optimal portfolios' frontiers evaluated considering the whole dataset over the two periods. Afterwards, we make an *out-of-sample* analysis considering a rolling window of length 100 days, giving equal probability 1/100 to each possible outcome. Every day the portfolio is rebalanced following the rules of Markowitz portfolio theory, mean-$iqr_{0.25}$ analysis, mean-$\Delta_{0.25}$ analysis and comparing them with the equally weighted portfolio. Afterwards the empirical distribution of the portfolios' logreturns is evaluated.

In Fig. 5.9 and 5.10 the comparison between the optimal portfolios' frontiers evaluates with the different methods are presented: the portfolio is composed again of the first five assets of the $S\&P$ 500 ordered by their weight in the index (Apple Inc., Microsoft Corporation, Exxon Mobil Corporation, Amazon.com Inc. and JPMorgan Chase & Co) considering the historical distribution of the logreturns from 1^{st} January 2011 to 31^{th} December 2013 and 1^{st} January 2014 to 31^{th} December 2016 respectively.

As we expect, the efficient frontier is not convex in the case of the iqr_α because it is not a coherent risk measure due to the lack of subadditivity, on the contrary the efficient frontiers of the mean-Δ_τ portfolios are convex for all levels of τ.

In Fig. 5.11 and 5.12 the portfolio weights of the efficient frontiers are represented, α is set to 0.2, so the respective τ is 0.025: the graphic of the mean-$\Delta_{0.025}$ portfolio weights is again similar to graphic of the Markowitz portfolio weights, while minimizing $iqr_{0.25}$ the portfolio composition is not regular with respect to the different values of the expected logreturn of the portfolio.

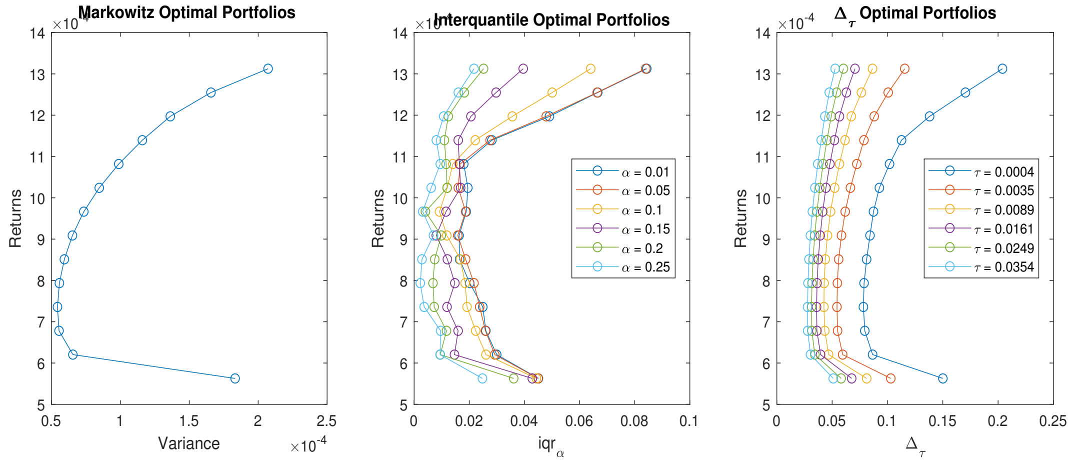


Figure 5.9: Optimal portfolios evaluated with the Markowitz model, the mean-iqr_α analysis and the mean-Δ_τ analysis for different values of α and τ during the period January 2011 December 2013.

⟨25ef⟩

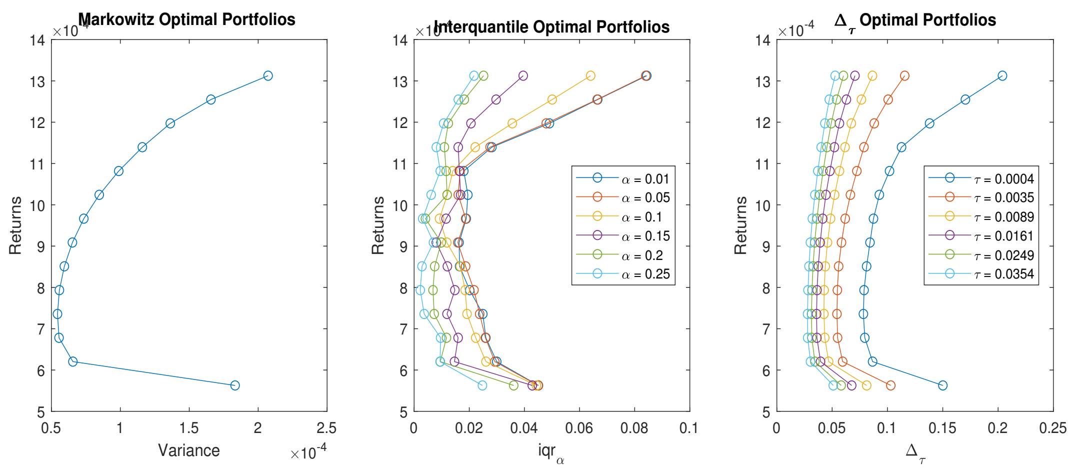


Figure 5.10: Optimal portfolios evaluated with the Markowitz model, the mean-iqr_α analysis and the mean-Δ_τ analysis for different values of α and τ during the period January 2014 December 2016.

⟨25efa⟩

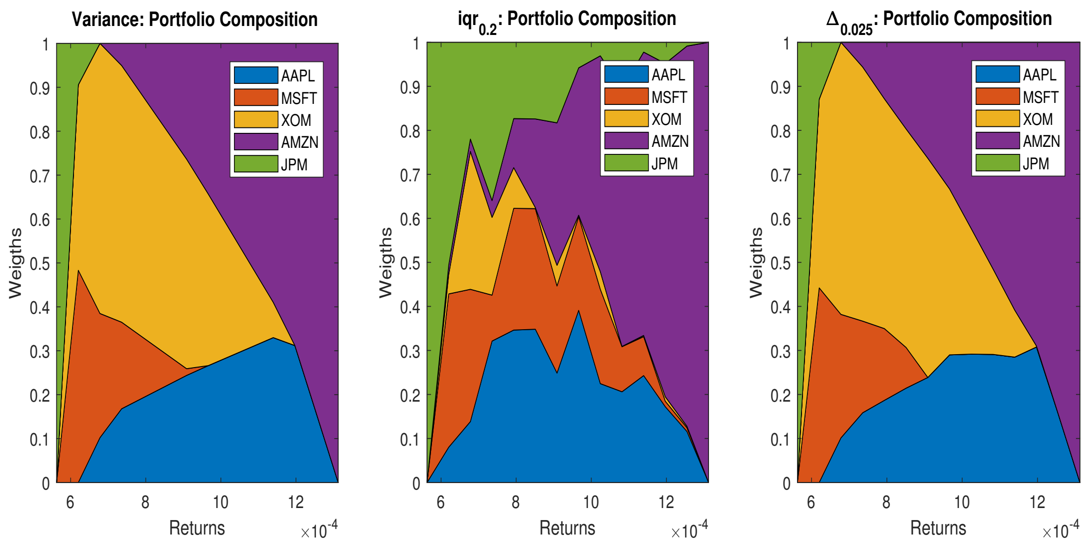


Figure 5.11: Portfolio weights of the optimal portfolios evaluated in the different methods during the period January 2011 December 2013.

⟨25efpw⟩

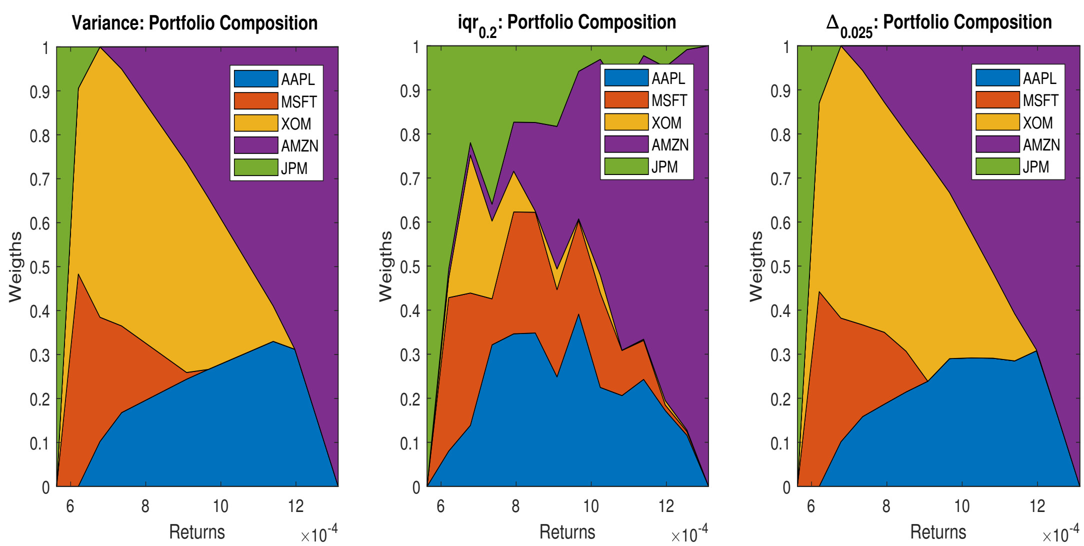


Figure 5.12: Portfolio weights of the optimal portfolios evaluated in the different methods during the period January 2014 December 2016.

⟨25efpwa⟩

In Fig. 5.13 and 5.14 the empirical distribution of the out-of-sample portfolios' logreturns is represented, and in Tab. 5.1 and 5.2 the extrema and the moments of the distributions are evaluated.

In this analysis we consider again two different periods: from 1^{st} January 2011 to 31^{th} December 2013 and from 1^{st} January 2014 to 31^{th} December 2016.

Each day we compute the realized logreturn of the mean-risk portfolio evaluated the day before. We consider variance, $iqr_{0.25}$, $\Delta_{0.25}$ as risk measures and we compare the relative portfolios with the the equally weighted portfolio.

In order to concentrate ourselves on the upper part of the efficient frontier, as we did in Section 4.4.2 we added a lower bound on the expected logreturn in addition to a non-negativity constraint. So each day we solve the LP problem

$$\begin{aligned} &\min_{w \in \Pi} \rho(w^i) \quad \text{s.t.} \\ &\quad \mu(w^i) \geqslant \frac{1}{5} \min_j \mathbb{E}[r_j^i] + \frac{4}{5} \max_j \mathbb{E}[r_j^i] \\ &\quad w^i \geqslant 0, \end{aligned} \tag{5.4}$$

constr2

where w^i are the portfolio weights evaluated in the rolling window prior to the day i and $\mathbb{E}[r_j^i]$ is the mean of the logreturns of the asset j evaluated in the rolling window prior to the day i.

The constraint in eq. 5.4 is arbitrary: we imposed it to investigate the high region of the optimal portfolios' frontier (see Fig. 4.17): from a financial point of view, we are minimizing the risk of portfolio with high performances.

Looking at Tab. 5.1 and 5.2 we can observe that the portfolio which minimize Δ_τ has performance comparable with the Markowitz portfolio in fact the relative empirical distributions are very similar: the mean-$iqr_{0.25}$ portfolio is another matter, it has the best realized logreturn but also the greatest standard deviation.

The tests performed on historical data confirm what we expected from theory, i.e., minimizing Δ_τ seems to be a viable alternative to the Markowitz portfolio considering the fact that the first is a LP problem, on the contrary the second is a QP problem.

1^{st} January 2011 - 31^{th} December 2013

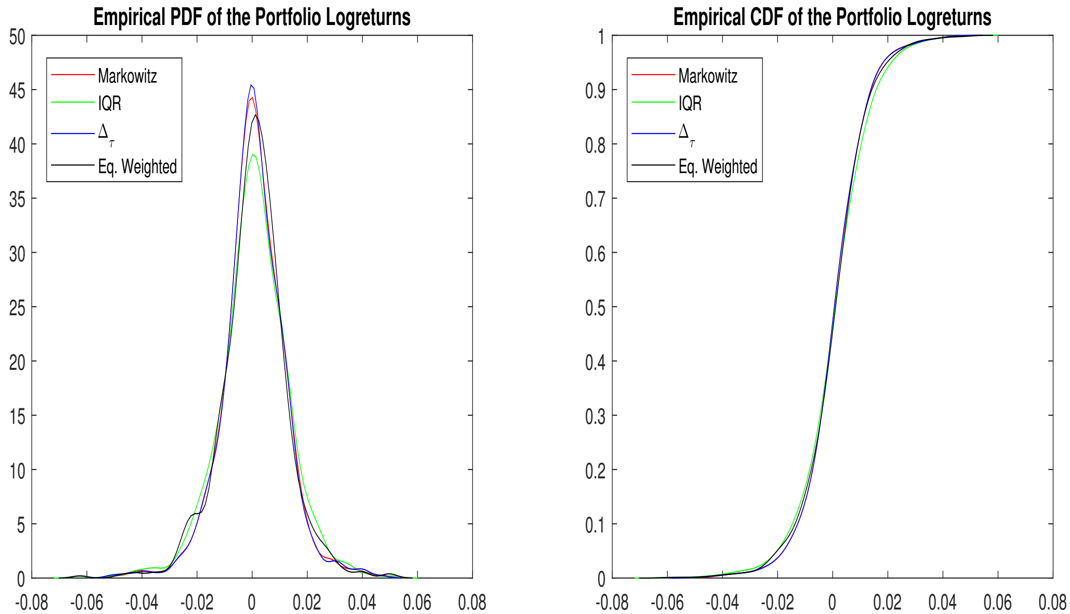


Figure 5.13: Empirical distribution of the 5-asset-portfolio logreturns computed out-of-sample with the rolling window approach over the period Jan11-Dec13: the probability and cumulative density function are smoothed by normal kernels.

⟨2511-13⟩

	Markowitz	$iqr_{0.25}$	Δ_τ	**Eq. Weighted**
Min	-5.08e-02	-6.30e-02	-5.12e-02	-6.24e-02
Max	4.76e-02	5.13e-02	4.68e-02	5.04e-02
Mean	6.77e-04	7.50e-04	6.62e-04	6.78e-04
Std	1.12e-02	1.25e-02	1.13e-02	1.19e-02
Skew	-1.48e-01	-2.89e-01	-1.43e-01	-2.49e-01
Kurt	5.38	5.15	5.53	5.68

Table 5.1: Statistics of the 5-asset portfolio logreturn distributions over the period January 2011 - December 2013

⟨tablestat21⟩

1^{st} January 2011 - 31^{th} December 2013

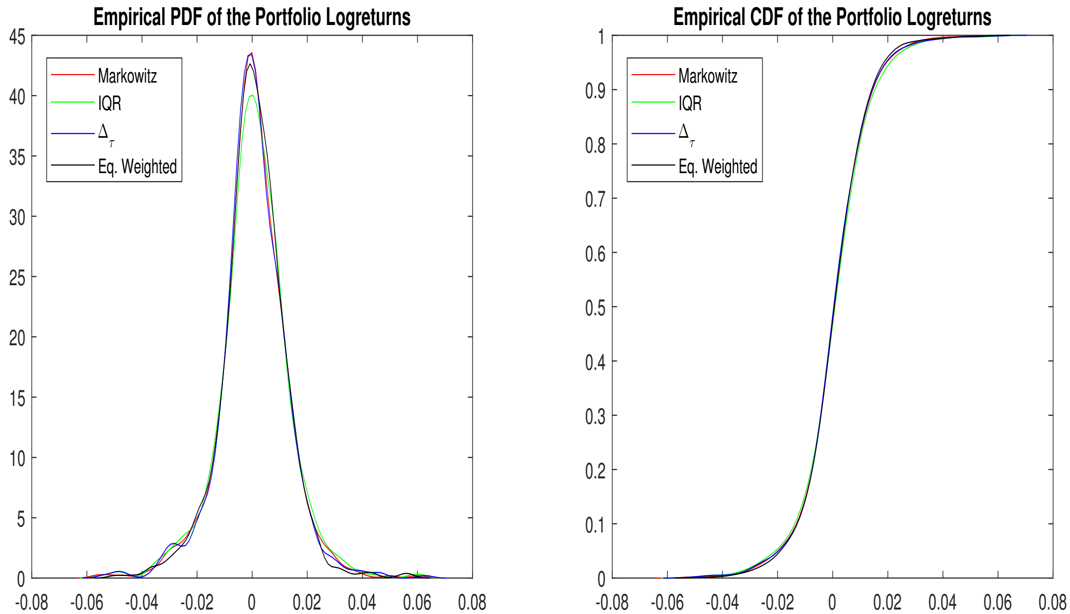


Figure 5.14: Empirical distribution of the 5-asset-portfolio logreturns computed out-of-sample with the rolling window approach over the period Jan14-Dec16: the probability and cumulative density function are smoothed by normal kernels.

⟨2514-16⟩

	Markowitz	$iqr_{0.25}$	Δ_τ	**Eq. Weighted**
Min	-5.63e-02	-5.41e-02	-5.38e-02	-4.95e-02
Max	6.01e-02	6.13e-02	6.27e-02	5.67e-02
Mean	5.55e-04	7.52e-04	5.01e-04	6.28e-04
Std	1.19e-02	1.26e-02	1.20e-02	1.15e-02
Skew	-1.68e-01	-8.92e-02	-8.23e-02	-3.00e-03
Kurt	5.79	5.72	6.06	5.71

Table 5.2: Statistics of the 5-asset portfolio logreturn distributions over the period January 2014 - December 2016

⟨tablestat22⟩

In the following experiment we will consider a portfolio composed by the first 250 asset of $S\&P$ 500 ordered by their weight in the index: what we want to test is the stability of the algorithm for large portfolios.
In Fig. 5.15 it is shown the comparison between the efficient frontiers of the different methods varying α and the respective τ: the portfolio is composed of the first 250 assets considering the historical distribution of the logreturns from 1^{st} January 2012 to 31^{th} December 2016.
Again, the efficient frontier is not convex in the case of the iqr_{α}.

In Fig. 5.17 and 5.18 the empirical distribution of the out-of-sample portfolios' logreturns are represented and in Tab. 5.3 and 5.4, the extrema and the moments of the distributions are evaluated.

Considering large portfolio, minimizing Δ_{τ} gives again results comparable with the Markowitz portfolio: the empirical distribution and the statistics are similar. Our algorithm works also with large portfolios and gives again the results expected from theory: from a computational point of view minimizing Δ_{τ} is preferable to minimizing variance dealing with a LP problem and not with a QP problem.

Minimizing $iqr_{0.25}$ gives portfolios with higher realized logreturns but on the other hand also the variance is higher. Further analysis should be performed on this risk measure but it is not the aim of this work.

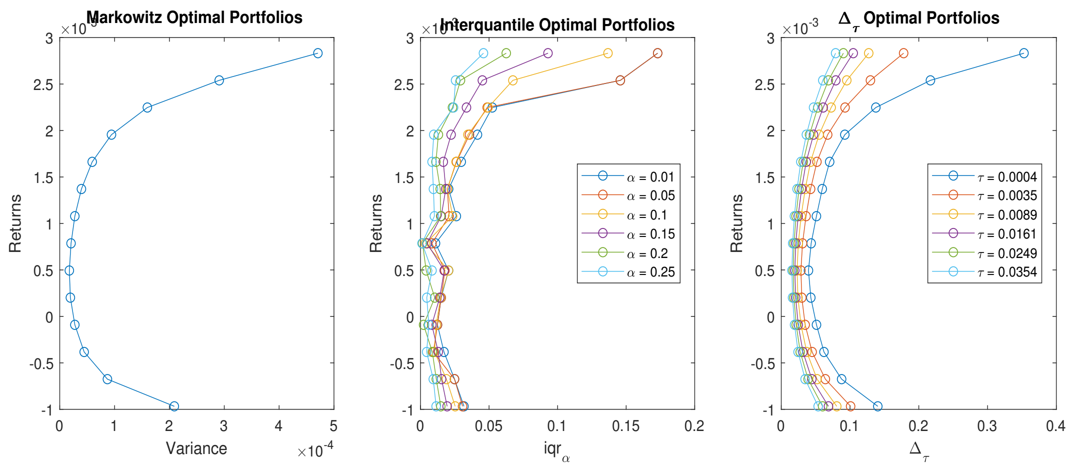


Figure 5.15: Optimal portfolios, composed of the first 250 assets of the $S\&P$ 500, evaluated with the Markowitz portfolio theory, the mean-iqr_α analysis and the mean-Δ_τ analysis for different values of α and τ during the period January 2011 December 2013.

⟨2250ef⟩

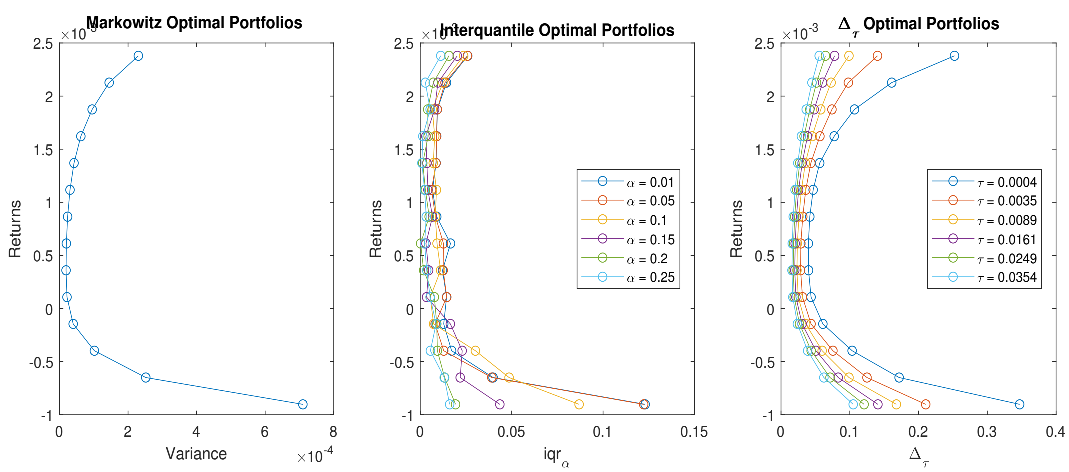


Figure 5.16: Optimal portfolios, composed of the first 250 assets of the $S\&P$ 500, evaluated with the Markowitz portfolio theory, the mean-iqr_α analysis and the mean-Δ_τ analysis for different values of α and τ during the period January 2014 December 2016.

?⟨2250efa⟩?

1^{st} January 2011 - 31^{th} December 2013

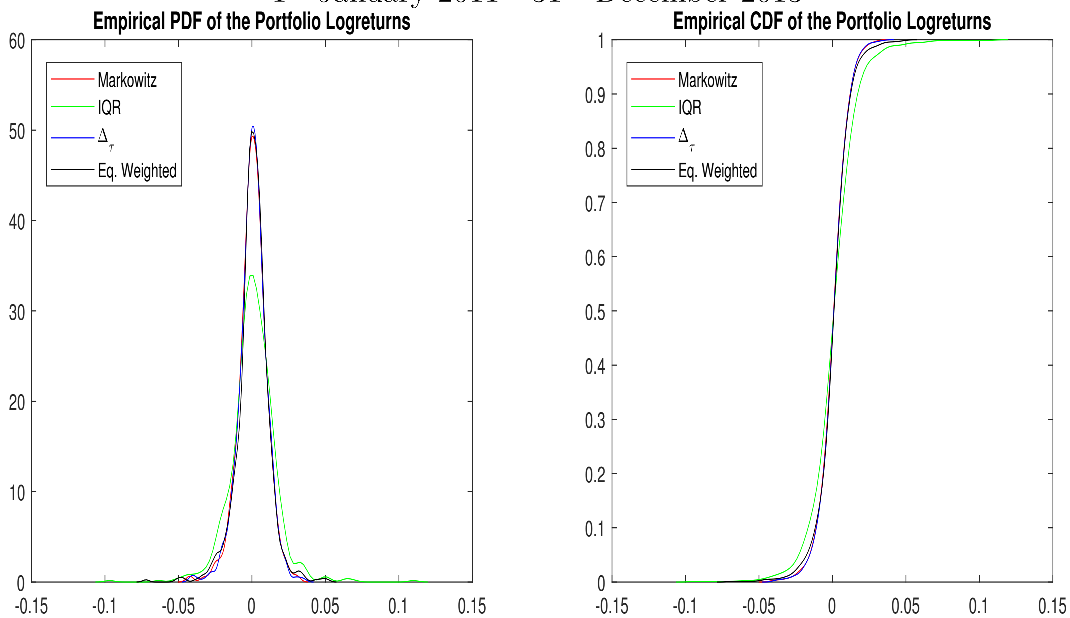


Figure 5.17: Empirical distribution of the 250-asset-portfolio logreturns computed out-of-sample for the period Jan11-Dec13: the probability and cumulative density function are smoothed by normal kernels.
⟨225011-13⟩

	Markowitz	$iqr_{0.25}$	Δ_τ	**Eq. Weighted**
Min	-4.22e-02	-9.67e-02	-4.12e-02	-7.20e-02
Max	3.16e-02	1.09e-01	3.54e-02	5.10e-02
Mean	4.79e-04	8.98e-04	6.07e-04	6.83e-04
Std	9.56e-03	1.53e-02	9.45e-03	1.10e-02
Skew	-4.47e-01	2.19e-01	-4.42e-01	-5.83e-01
Kurt	4.94	10.199	5.06	8.41

Table 5.3: Statistics of the 250-asset portfolio logreturn distributions over the period January 2014 - December 2016
⟨tablestat23⟩

1^{st} January 2014 - 31^{th} December 2016

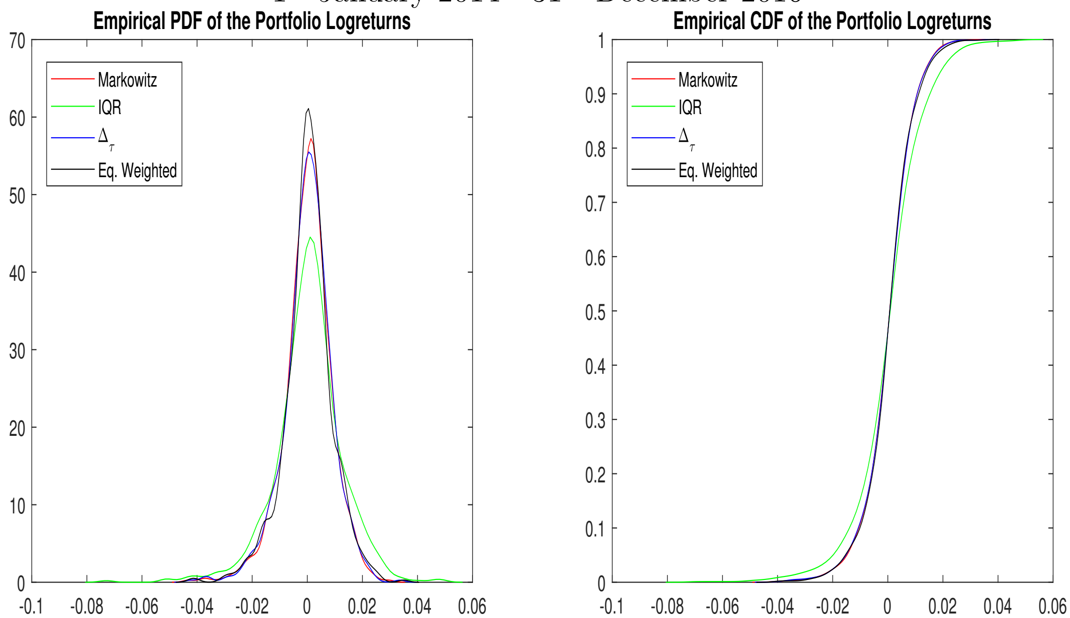


Figure 5.18: Empirical distribution of the 250-asset-portfolio logreturns computed out-of-sample for the period Jan14-Dec16: the probability and cumulative density function are smoothed by normal kernels.

⟨225014-16⟩

	Markowitz	$iqr_{0.25}$	Δ_τ	**Eq. Weighted**
Min	-4.42e-02	-7.28e-02	-4.23e-02	-4.21e-02
Max	2.95e-02	4.86e-02	3.36e-02	3.50e-02
Mean	2.44e-04	5.09e-04	2.61e-04	4.11e-04
Std	8.64e-03	1.23e-02	8.79e-03	8.81e-03
Skew	-5.90e-01	-5.78e-01	-5.91e-01	-4.35e-01
Kurt	5.20	6.38	5.15	5.13

Table 5.4: Statistics of the 250-asset portfolio logreturn distributions over the period January 2014 - December 2016

⟨tablestat24⟩

Chapter 6

Robust Optimization of Expectiles with a worst-case approach

⟨capmeanWEVaR⟩

6.1 Robust Risk Measures

In modern empirical portfolio theory assets are sometimes supposed to follow discrete distributions defined by historical data: in fact if $r_1, \ldots, r_t$ are the last t logreturns of a certain asset X, in general it is assumed that

$$\mathbb{P}(X = r_i) = p_i, \quad i = 1, \ldots, T.$$

with $\sum_i p_i = 1$.
p_i represents the frequency of the event r_i and in general is chosen to be $1/T$.
This assumption is very convenient due to its simplicity, but on the other hand it does not take into account features strictly related to the market, such as the leverage or the model uncertainty.
In recent years, several authors paid more attention to the robustification of a risk measures: Ben-Tal et al. (2000) formulated the *Robust Counterpart* approach and illustrated it through multi-stage asset allocation problem; Lobo and Boyd (2000), Costa and Paiva (2002), Goldfarb and Iyengar (2003) gave a robust version of the Markowitz portfolio problem; Zhu and Fukushima (2009) investigated the *worst-case CVaR*.
Robustification of a risk measure means adding uncertainty to the distribution

considering a perturbation of the probability p_i of the events $X = x_i$.

Definition 33. *Let $\rho\colon \mathcal{X} \to \overline{\mathbb{R}}$ be a risk measure and $\mathcal{P}$ a class of probability measure. The worst-case ρ is defined as:*

$$\rho_W(X) := \sup_{\mathbb{P}\in\mathcal{P}} \rho(X), \quad X \in \mathcal{X}$$

Zhu and Fukushima (2009) showed that the robustification preserves the property of coherence. Indeed:

Proposition 34. *If ρ is a coherent risk measure under a probability measure $\mathbb{P}$, then the corresponding ρ_W associated to a class of probability measure $\mathcal{P}$ remains a coherent risk measure.*

Proof. We will prove that all properties in Definition 2 are satisfied.

subadditivity If $X, Y \in \mathcal{X}$, then

$$\begin{aligned}
\rho_W(X+Y) &= \sup_{\mathbb{P}\in\mathcal{P}} \rho(X+Y) \\
&\leq \sup_{\mathbb{P}\in\mathcal{P}} [\rho(X)+\rho(Y)] \\
&\leq \sup_{\mathbb{P}\in\mathcal{P}} \rho(X) + \sup_{\mathbb{P}\in\mathcal{P}} \rho(Y) \\
&= \rho_W(X) + \rho_W(Y)
\end{aligned}$$

Positive homogeneity If $a \geqslant 0$ and $X \in \mathcal{X}$, then

$$\rho_W(aX) = \sup_{\mathbb{P}\in\mathcal{P}} \rho(aX) = \sup_{\mathbb{P}\in\mathcal{P}} a\rho(X) = a \sup_{\mathbb{P}\in\mathcal{P}} \rho(X) = a\rho_W(X)$$

Translation invariance If $a \in \mathbb{R}$ and $X \in \mathcal{X}$, then

$$\rho_W(X+a) = \sup_{\mathbb{P}\in\mathcal{P}} \rho(X+a) = \sup_{\mathbb{P}\in\mathcal{P}} [\rho(X)-a] = \sup_{\mathbb{P}\in\mathcal{P}} \rho(X) - a = \rho_W(X) - a$$

Monotonicity If $X, Y \in \mathcal{X}$ and $X \leqslant Y$ for each event $\omega \in \Omega$, then $\rho(X) \geqslant \rho(Y)$ for any probability measure $\mathbb{P} \in \mathcal{P}$, which implies

$$\rho_W(X) = \sup_{\mathbb{P}\in\mathcal{P}} \rho(X) \geqslant \sup_{\mathbb{P}\in\mathcal{P}} \rho(Y) = \rho_W(Y).$$

$\square$

The problem is now to properly define $\mathcal{P}$ in order to add uncertainty to the distribution: as pointed out by many authors, the uncertainty can be added in different ways (see e.g. Costa and Paiva (2002), Goldfarb and Iyengar (2003), Halldórsson and Tütüncü (2003) and El Ghaoui et al. (2003)).
In this work we will focus on box uncertainty: the class of probability measure $\mathcal{P}$ is defined as

$$\mathcal{P} := \{p_i : p_i = p_i^0 + \eta_i, \ \sum_i \eta_i = 0, \ \underline{\eta_i} \leqslant \eta_i \leqslant \overline{\eta_i}\} \tag{6.1}$$ Box

where $\underline{\eta_i}$ and $\overline{\eta_i}$ are given constants and p_i^0 is the nominal probability for the event $X = r_i$ that represents the most likely distribution. The condition $\sum_i \eta_i = 0$ ensures that the elements belonging $\mathcal{P}$ are probability distributions.
As pointed out by Tütüncü and Koenig (2004) and Bertsimas et al. (2011), there are three main class of uncertainty sets:

- Box Uncertainty: the set of probability measure is defined in eq. 6.1;
- Ellipsoidal Uncertainty: the set of probability measure is defined as

$$\mathcal{P} := \{p \in \mathbb{R}^N : p = p^0 + A\eta, \ e^T A\eta = 0, \ p^0 + A\eta \geqslant 0, \ \sqrt{\eta^T \eta} \leqslant 1\}; \tag{6.2}$$ ellun

- Tail Uncertainty: the set of probability measure is defined as

$$\mathcal{P} := \{p \in \mathbb{R}^N : p = \sum_{i=1}^N \eta_i p_i, \ \sum_{i=1}^N \eta_i = 1, \ \eta_i \leqslant \frac{1}{N(1-\alpha)}, \ i = 1, \dots, N\}. \tag{6.3}$$ tailun

In this chapter a robust version of the *Expectile Value-at-Risk* is provided: we consider Box Uncertainty because it is the easiest to implement both from a theoretical and a computational points of view, since the set of constraints in eq. (6.1) is linear. The method can be trivially extended to Ellipsoidal Uncertainty and Tail Uncertainty simply changing the constraints concerning the set of probability measure: from a computational point of view, the resulting optimization problems are more complex because the constraints added are not linear anymore.
Thus, instead of assuming that the distribution is known, we suppose that the density function belong to a set $\mathcal{P}$ of distributions, as described by Zhu and Fukushima (2009).

6.2 Formulation

Definition 35. *The* Robust Expectile Value-at-Risk (WEVaR) *for a fixed event $x \in \mathcal{X}$ with respect to $\mathcal{P}$ is defined as*

$$WEVaR_\tau(X) := \sup_{p \in \mathcal{P}} EVaR_\tau(X).$$

Considering the dual representation of the EVaR (3.4) we have that:

$$\begin{aligned} WEVaR_\tau(X) &:= \sup_{p \in \mathcal{P}} EVaR_\tau(X) = \sup_{p \in \mathcal{P}} \max_{\varphi \in \mathcal{M}_\beta} \mathbb{E}_p[-X\psi] \\ &= \sup_{(p,\varphi) \in \mathcal{M}_\beta^{\mathcal{P}}} \mathbb{E}_p[-x\psi], \end{aligned} \tag{6.4}$$ DualWEVaR

where

$$\begin{aligned} \mathcal{M}_\beta &= \left\{ \varphi \in L^\infty(\Omega, \mathcal{F}, \mathbb{P}),\ \varphi \geq 0,\ \mathbb{E}[\varphi] = 1, \frac{\operatorname{ess\,sup}(\varphi)}{\operatorname{ess\,inf}(\varphi)} \leqslant \beta \right\}, \\ \mathcal{M}_\beta^{\mathcal{P}} &= \left\{ (p, \varphi) \in \mathcal{P} \times L^\infty(\Omega, \mathcal{F}, \mathbb{P}),\ \varphi \geq 0,\ \mathbb{E}_p[\varphi] = 1, \frac{\operatorname{ess\,sup}(\varphi)}{\operatorname{ess\,inf}(\varphi)} \leqslant \beta \right\} \end{aligned}$$

with $\beta = \dfrac{1-\tau}{\tau}$.

We now suppose that X follows a discrete distribution, introducing a *box uncertainty* on the distribution (see Halldórsson and Tütüncü (2003)), the set $\mathcal{P}$ becomes

$$\mathcal{P} := \left\{ p \in \mathbb{R}^n,\ \sum_{i=1}^n p_i = 1,\ \underline{p} \leqslant p_i \leqslant \overline{p} \ \ \forall i,\ \text{with } 0 < \underline{p} < \overline{p} < 1 \right\}.$$

Hence we can discretize the equation (6.4) introducing the box uncertainty, so we obtain:

$$\begin{aligned} WEVaR_\tau(X) = \quad & \max_{p_i, f_i} \sum_{i=1}^n -x_i f_i p_i \quad \text{s.t.} \\ & \sum_{i=1}^n f_i p_i = 1 \\ & f_i \leqslant \beta f_j,\ \forall i \neq j \\ & \sum_{i=1}^n p_i = 1, \\ & \underline{p} \leqslant p_i \leqslant \overline{p}, \quad f_i \geqslant 0, \forall i = 1, \dots, n. \end{aligned} \tag{6.5}$$ DWEVaR

This is a quadratic problem with quadratic constraints, which is not easy to handle with standards optimization solvers.

The box uncertainty is given by the constraints $\sum_{i=1}^{n} p_i = 1$ and $\underline{p} \leqslant p_i \leqslant \overline{p}$: in order to adapt the method to ellipsoidal uncertainty and tail uncertainty we can easily substitute these constraints towards complying with sets in eq. 6.2 or 6.3. We now focus on rewriting the problem: adding again the extra decision variable m we substitute the set of constraints $\{f_i - \beta f_j \leqslant 0, \forall i \neq j\}$ with $\{m \leqslant f_i \leqslant \beta m, \forall i\}$ and imposing $\eta_i = f_i p_i$, the equation (6.5) becomes:

$$
\begin{aligned}
WEVaR_\tau(X) = \quad & \max_{\eta_i, m, p_i} \sum_{i=1}^{n} -x_i \eta_i \quad \text{s.t.} \\
& \sum_{i=1}^{n} \eta_i = 1 \\
& p_i m \leqslant \eta_i \leqslant \beta p_i m, \ \forall i = 1, \dots, n \\
& \sum_{i=1}^{n} p_i = 1, \\
& \underline{p} \leqslant p_i \leqslant \overline{p}, \ m \geqslant 0, \ \eta_i \geqslant 0, \forall i = 1, \dots, n.
\end{aligned}
$$

The problem has now a linear objective function and $2n$ quadratic constraints given by $p_i m \leqslant \eta_i \leqslant \beta p_i m, \ \forall i = 1, \dots, n$.

6.3 Portfolio Problem

Suppose now that the set of the scenarios X are portfolio logreturns, so we have that $X = Rw$, where R is the matrix of the historical logreturn and $w = (w_1, \dots, w_m)$ are the weights. Finding the portfolio which minimize the $WEVaR_\tau$ means finding w such that,

$$
\begin{aligned}
w = & \operatorname*{argmin}_w WEVaR_\tau(Rw) \quad \text{s.t} \\
& \sum_{j=1}^{m} w_j = 1, \\
& w_j \geq 0, \quad \forall j = 1, \dots, m.
\end{aligned} \tag{6.6}
$$

eqport

Therefore the global minimum WEVaR is given by

$$
\begin{array}{lll}
WEVaR_\tau(R) = & \min\limits_{w} & \max\limits_{\eta_i,m,p_i} \sum_{i=1}^{n} -(Rw)_i\eta_i \quad \text{s.t.} \\
& \sum_{j=1}^{m} w_j = 1 & \sum_{i=1}^{n} \eta_i = 1 \\
& w_j \geq 0. & p_i m \leqslant \eta_i \leqslant \beta p_i m, \ \forall i = 1, \dots, n \\
& & \sum_{i=1}^{n} p_i = 1, \\
& & \underline{p} \leqslant p_i \leqslant \overline{p}, \ m \geqslant 0, \ \eta_i \geqslant 0, \forall i = 1, \dots, n
\end{array}
\tag{6.7}
$$

DEVaR2

Observing that the objective function $\phi(w, \eta)$ is quadratic but $\phi(w, \cdot)$ is linear in w and $\phi(\cdot, \eta)$ is linear in η, we can apply the minmax theorem in Fan (1953).

⟨minmax⟩ **Theorem 36.** *Suppose that $\mathcal{X}$ and $\mathcal{Y}$ are nonempty compact convex set in $\mathbb{R}^n$ and $\mathbb{R}^m$, respectively, and the function $\phi(x, y)$ is convex in x for any given y, and concave in y for any given x. Then we have*

$$\min_{x \in \mathcal{X}} \max_{y \in \mathcal{Y}} \phi(x, y) = \max_{y \in \mathcal{Y}} \min_{x \in \mathcal{X}} \phi(x, y).$$

Applying the theorem 36 to eq. (6.7) we obtain

$$
\begin{array}{lll}
WEVaR_\tau(R) = \max\limits_{\eta_i,m,p_i} & & \min\limits_{w} \sum_{i=1}^{n} -(Rw)_i\eta_i \quad \text{s.t.} \\
& \sum_{i=1}^{n} \eta_i = 1 & \sum_{j=1}^{m} w_j = 1, \\
& p_i m \leqslant \eta_i \leqslant \beta p_i m, & w_j \geq 0, \quad \forall j = 1, \dots, m. \\
& \sum_{i=1}^{n} p_i = 1, & \\
& \underline{p} \leqslant p_i \leqslant \overline{p}, \ m \geqslant 0, \ \eta_i \geqslant 0 &
\end{array}
\tag{6.8}
$$

DEVaR3

For every fixed η, the problem

$$
\begin{array}{l}
\min\limits_{w} \sum_{i=1}^{n} -(Rw)_i\eta_i \quad \text{s.t.} \\
\sum_{j=1}^{m} w_j = 1 \\
w_j \geq 0, \quad \forall j = 1, \dots, m.
\end{array}
$$

is a linear programming problem, so we can evaluate the dual:

$$
\begin{array}{ll}
\text{PRIMAL} & \text{DUAL} \\
\min\limits_{w} \sum_{i=1}^{n} -(Rw)_i \eta_i \quad \text{s.t.} & \max\limits_{\omega} \omega \quad \text{s.t.} \\
\sum_{j=1}^{m} w_j = 1, & \omega \leqslant -(\eta' R)_j, \quad \forall j = 1, \ldots, m, \\
w_j \geq 0, \quad \forall j = 1, \ldots, m, & \omega \geqslant 0.
\end{array}
\tag{6.9}
$$
$\boxed{\texttt{Dual}}$

Combining equation (6.8) and the dual formulation in (6.9) we obtain a maximization problem of a maximization problem, that can be written as:

$$
\begin{aligned}
WEVaR_\tau(R) = \max_{\eta_i, m, p_i, \omega} \omega \quad & \text{s.t.} \\
& \sum_{i=1}^{n} \eta_i = 1 \\
& p_i m \leqslant \eta_i \leqslant \beta p_i m, \\
& \sum_{i=1}^{n} p_i = 1, \\
& \omega \leqslant -(\eta' R)_j, \quad \forall j = 1, \ldots, m. \\
& \underline{p} \leqslant p_i \leqslant \overline{p}, \ \eta_i \geqslant 0, \quad \forall i = 1, \ldots, n \\
& m \geqslant 0, \ \omega \geqslant 0.
\end{aligned}
\tag{6.10}
$$
$\boxed{\texttt{DEVaR7}}$

6.3.1 Efficient Frontier

In order to evaluate the efficient frontier which minimize the robust expectile value-at-risk it is necessary to add the constraint on the expected logreturn of the portfolio to the equation (6.6):

$$
\begin{aligned}
w = \operatorname*{argmin}_{w} WEVaR_\tau(Rw) \quad & \text{s.t} \\
& \sum_{j=1}^{m} w_j = 1, \\
& \sum_{j=1}^{m} w_j \mu_j = \mu, \\
& w_j \geq 0, \quad \forall j = 1, \ldots, m,
\end{aligned}
$$

where μ_i is the mean of the logreturn of the i-th asset, and μ is an arbitrary constant. Therefore we have

$$
\begin{aligned}
WEVaR_\tau(R,\mu) = \quad & \min_w & & \max_{\eta_i,m,p_i} \sum_{i=1}^n -(Rw)_i\eta_i \quad \text{s.t.} \\
& \textstyle\sum_{j=1}^m w_j = 1 & & \sum_{i=1}^n \eta_i = 1 \\
& \textstyle\sum_{j=1}^m w_j\mu_j = \mu, & & p_i m \leqslant \eta_i \leqslant \beta p_i m,\ \forall i = 1,\dots,n \\
& w_j \geq 0. & & \sum_{i=1}^n p_i = 1, \\
& & & \underline{p} \leqslant p_i \leqslant \overline{p},\ m \geqslant 0,\ \eta_i \geqslant 0,\ \forall i = 1,\dots,n.
\end{aligned}
$$

Applying again theorem 36, it is possible to swap the min function with the max function:

$$
\begin{aligned}
WEVaR_\tau(R,\mu) = \max_{\eta_i,m,p_i} & & & \min_w \sum_{i=1}^n -(Rw)_i\eta_i \quad \text{s.t.} \\
& \sum_{i=1}^n \eta_i = 1 & & \sum_{j=1}^m w_j = 1, \\
& p_i m \leqslant \eta_i \leqslant \beta p_i m, & & \sum_{j=1}^m w_j\mu_j = \mu, \\
& \sum_{i=1}^n p_i = 1, & & w_j \geq 0, \quad \forall j = 1,\dots,m. \\
& \underline{p} \leqslant p_i \leqslant \overline{p},\ m \geqslant 0,\ \eta_i \geqslant 0 & &
\end{aligned}
\tag{6.11}
$$

DEVaR6

Proceeding as in equation (6.8) it is possible to evaluate the dual of min LP problem:

$$
\begin{aligned}
& \text{PRIMAL} & & \text{DUAL} \\
& \min_w \sum_{i=1}^n -(Rw)_i\eta_i \quad \text{s.t.} & & \max_\omega \omega_1 + \mu\omega_2 \quad \text{s.t.} \\
& \sum_{j=1}^m w_j = 1, & & \omega_1 + \mu_j\omega_2 \leqslant -(\eta' R)_j, \quad \forall j = 1,\dots,m, \\
& \sum_{j=1}^m w_j\mu_j = \mu, & & \omega_1 \geqslant 0,\ \omega_2 \geqslant 0. \\
& w_j \geq 0, \quad \forall j = 1,\dots,m, & &
\end{aligned}
\tag{6.12}
$$

Dual2

Combining equation (6.11) and the dual formulation in (6.12) we obtain

$$
\begin{aligned}
WEVaR_\tau(R,\mu) = \max_{\eta_i, m, p_i, \omega_k} \ & \omega_1 + \mu\omega_2 \quad \text{s.t.} \\
& \sum_{i=1}^{n} \eta_i = 1 \\
& p_i m \leqslant \eta_i \leqslant \beta p_i m, \\
& \sum_{i=1}^{n} p_i = 1, \\
& \omega_1 + \mu_j \omega_2 \leqslant -(\eta' R)_j, \quad \forall j = 1, \ldots, m. \\
& \underline{p} \leqslant p_i \leqslant \overline{p}, \ \eta_i \geqslant 0, \quad \forall i = 1, \ldots, n \\
& m \geqslant 0, \ \omega_1 \geqslant 0, \ \omega_2 \geq 0.
\end{aligned}
$$

6.4 Numerical Examples

In the first part of this section we consider optimal portfolios on simulated data: our purpose is to compare the portfolio problem using $WEVaR_\tau$ as risk measure with the portfolio which minimize $EVaR_\tau$. In the second part of the section we test the method on historical data.
All the experiments are performed using MatLab and the optimization problems are solved with the functions `linprog`, `quadprog` and `fmincon`. The solutions to Linear and Quadratic Programming problems are checked with GAMS (General Algebraic Modeling System).

6.4.1 Simulated Data

As toy problem we consider again the pattern of section 4.4.1, that is a portfolio composed of three assets varying the distributions and the correlations.
We perform only 100 simulations because the problem described in equation (6.10) is non linear and consequently has an high computational cost.
In the first example we suppose that the logreturn of three assets A, B and, C are described by uncorrelated normal distributions: in order to assess the variability of the optimal portfolio, we simulate samples of length $T = 20, 50, 100$ and compute the distribution of minimal $WEVaR_\tau$ portfolios for τ ranging from 5% to 45%. The results are reported in Fig. 6.1; the three lines correspond, respectively from top to bottom, to the 10^{th} percentile, the median and the 90^{th} percentile of the portfolio weights.
We perform experiments similar to the ones presented by Bertsimas et al. (2004) in order to study the variability of the portfolio with respect to τ and the length of the historical sample.
In Fig. 6.2 we compare the portfolios originated by the minimization of EVaR_τ and WEVaR_τ, through the scatter plot of the weights of A and B and the relative histogram (the weight of C is given by $1 - w_A - w_B$).

In Fig. 6.3, 6.4 we perform two experiments introducing correlation among the assets: A, B and C are described by normal distribution with correlation matrices

$$\Sigma_+ = \begin{pmatrix} 1 & 0.3 & 0.5 \\ 0.3 & 1 & 0.7 \\ 0.5 & 0.7 & 1 \end{pmatrix}, \quad \Sigma_- = \begin{pmatrix} 1 & -0.3 & 0.5 \\ -0.3 & 1 & -0.7 \\ 0.5 & -0.7 & 1 \end{pmatrix}$$

Again in Fig. 6.5 and 6.6 we consider uncorrelated asymmetric distributions of the following types: asset A and B has a standard normal distribution, asset C has a reflected gamma distribution (Fig. 4.2).
The parameters of the distributions are set as in eq. 4.12, i.e.

$$\begin{aligned} &\text{Asset A} \sim \mathcal{N}(\mu_{\mathcal{N}}, \sigma_{\mathcal{N}}), && \text{with} \;\; \mu_{\mathcal{N}} = 0, \;\; \sigma_{\mathcal{N}} = 2.25 \\ &\text{Asset B} \sim \mathcal{N}(\mu_{\mathcal{N}}, \sigma_{\mathcal{N}}), && \text{with} \;\; \mu_{\mathcal{N}} = 0, \;\; \sigma_{\mathcal{N}} = 2.25 \\ &\text{Asset C} \sim -\Gamma(k, \theta, \mu_\Gamma, \sigma_\Gamma), && \text{with} \;\; k = 5, \;\; \theta = 1, \;\; \mu_\Gamma = -0.912, \sigma_\Gamma = 1.5 \end{aligned} \tag{6.13}$$

parameters4

With this frame the three assets have the same left tail, but significantly different right tail.

Finally in Fig. 6.7 and 6.8 we introduce correlation among the assets: A, B and C have first correlation matrix Σ_+ and then correlation matrix Σ_-.

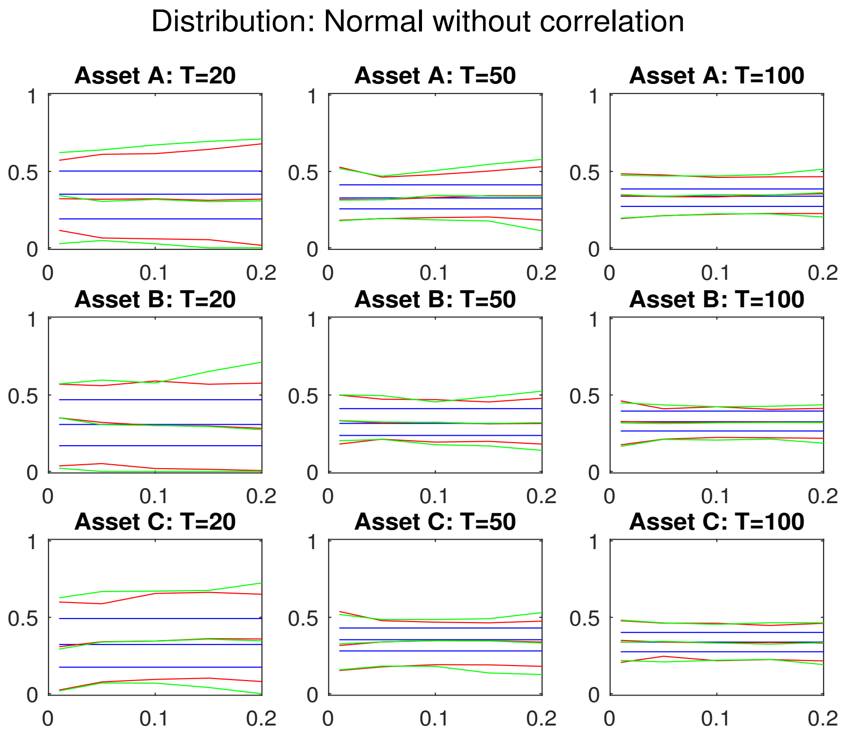


Figure 6.1: Variability of the weights of the portfolios found minimizing $WEVaR_\tau$ for $\tau = 0.01, \ldots, 0.2$. The red lines indicate respectively the 90th percentile, the median and the 10th percentile of the portfolio weights. The blue lines and the green lines represent respectively the same quantities for a minimal variance portfolio and a minimal $EVaR_\tau$ portfolio.

⟨Wnuncorrber⟩

In Fig. 6.1, 6.3, 6.5 and 6.7 we compare the portfolio compositions of the portfolios found minimizing $WEVaR_\tau$ and $EVaR_\tau$ for different values of τ. In most of the cases, the 10th percentile and the 90$^{\text{th}}$ percentile lines (red) of the $WEVaR_\tau$ are closer to the medians than the respectively lines (green) of the $EVaR_\tau$. Moreover in both cases the minimum of the distance between the 10th percentile and the 90th percentile is often reached in a neighborhood of $\tau = 0.05$ and the maximum is reached for $\tau = 0.2$. Thus we study in details how are composed the portfolios found minimizing $WEVaR_{0.05}$ and $WEVaR_{0.2}$ comparing them with the portfolios found minimizing $EVaR_{0.05}$ and $EVaR_{0.2}$ (Fig. 6.2, 6.4, 6.6 and 6.8).
Lookin at the figures from left to right, we can see that the variability decreases for increasing value of T, hence for longer historical dataset we find portfolios more accurate.

Considering the uncorrelated standard normal case (Fig. 6.2) the cloud of points the minimizing $WEVaR_\tau$ has the same shape of the one evaluated minimizing the $EVaR_\tau$. The only difference is that the robust case seems to have slightly less dispersion than the $EVaR_\tau$ case.

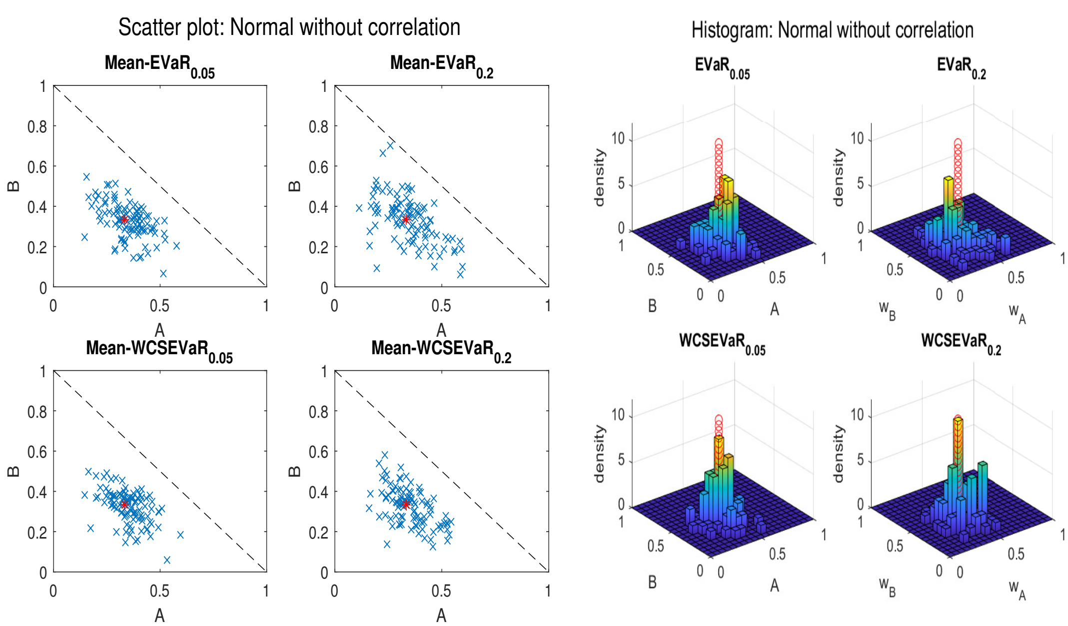


Figure 6.2: Scatter plot and histogram of the weights of the assets A and B, the weight of C is given by the relation $\sum \omega_i = 1$, the pattern is the same of Fig. 6.1. The red points represent the theoretical value of the minimal variance portfolio.
⟨Wnuncorr⟩

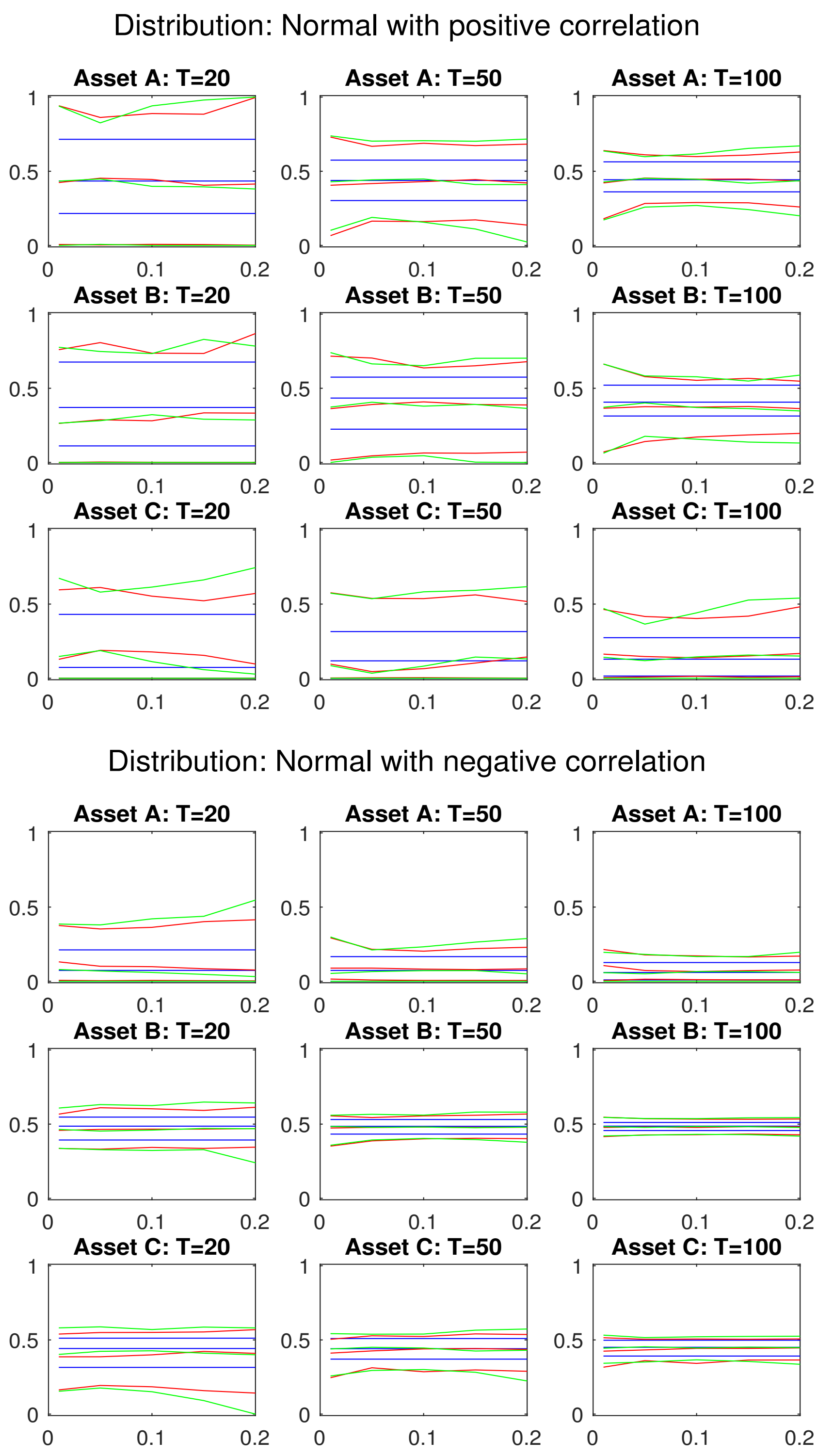


Figure 6.3: Variability of the weights of portfolios found minimizing $WEVaR_\tau$ for $\tau = 0.01, \ldots, 0.2$ on simulated correlated standard normal data of length $T = 20, 50, 100$. The red lines indicate respectively the 90^{th} percentile, the median and the 10^{th} percentile of the portfolio weights. The blue lines and the green lines represent respectively the same quantities for a minimal variance portfolio and a minimal $EVaR_\tau$ portfolio.

⟨Wncorrber⟩

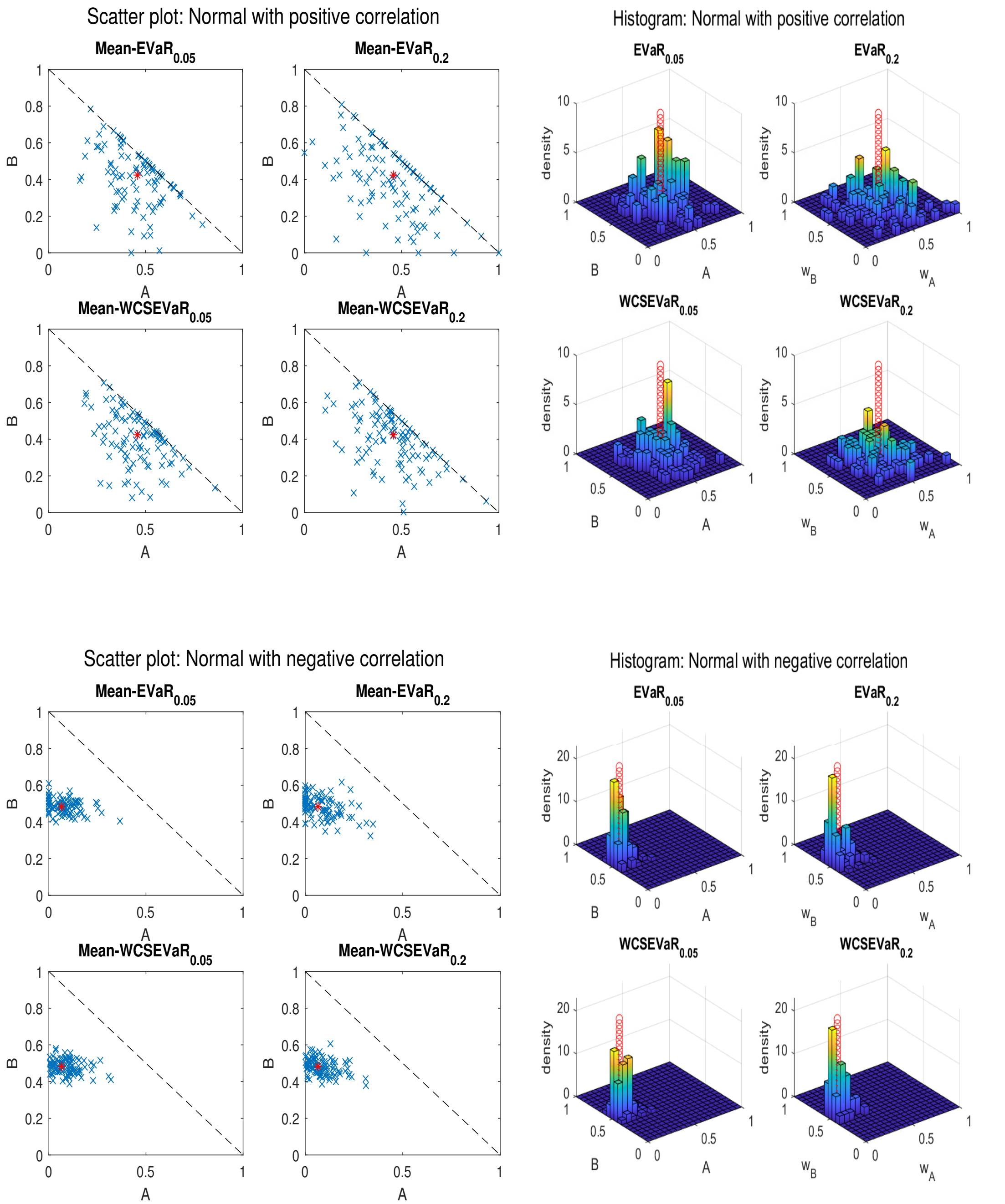


Figure 6.4: Scatter plot and histogram of the weights of the assets A and B, the weight of C is given by the relation $\sum \omega_i = 1$, the pattern is the same of Fig. 6.3 with 100 simulations. The red points represent the theoretical value of the portfolio weights given by the minimal variance portfolio.

⟨Wncorr⟩

Introducing correlation among the asset logreturns (Fig. 6.3 and 6.4) we can observe that in all cases the shapes of the clouds of ponits in the scatter plots again gather on the diagonal when the correlation matrix is Σ_+, on the contrary considering correlation matrix Σ_- the resulting portfolio is composed prevalently by assets B and C. In all cases evaluating the weights minimizing the $WEVaR_\tau$ has no significant differences from evaluating the weights minimizing the $EVaR_\tau$, but again the robust case seems to have slightly less dispersion than the $EVaR_\tau$ case.

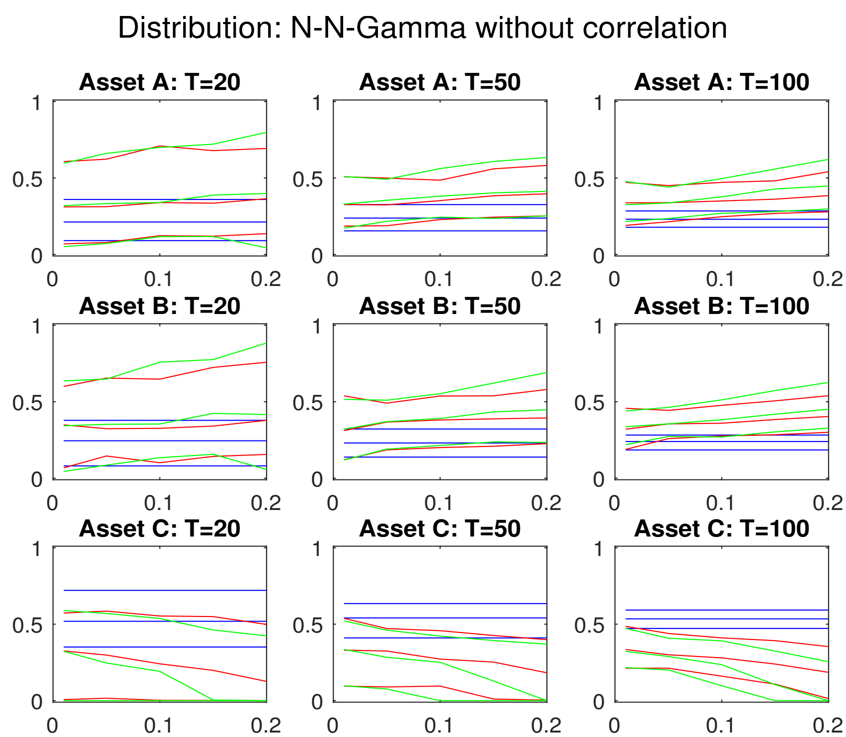


Figure 6.5: Variability of the weights of the portfolios found minimizing $WEVaR_\tau$ for $\tau = 0.01, \ldots, 0.2$ on simulated uncorrelated data of length $T = 20, 50, 100$: the asset A and B come from a normal distribution, C come from a reflected Γ distribution. The red lines indicate respectively the 90^{th} percentile, the median and the 10^{th} percentile of the portfolio weights. The blue lines and the green lines represent respectively the same quantities for a minimal variance portfolio and a minimal $EVaR_\tau$ portfolio.

⟨WNGGuncorrber⟩

Let now consider uncorrelated asymmetric distributions: assets A and B has a normal distribution, asset C has a reflected Γ distribution. The parameters that

identify the distributions are listed in eq. 6.13: the distributions have approximately the same left tail but significantly different right tail. The results are reported in Fig. 6.5 and 6.6.
As in the previous examples evaluating the weights minimizing the $WEVaR_\tau$ has no significant differences from evaluating the weights minimizing the $EVaR_\tau$, but the robust case seems to have slightly less dispersion than the $EVaR_\tau$ case.

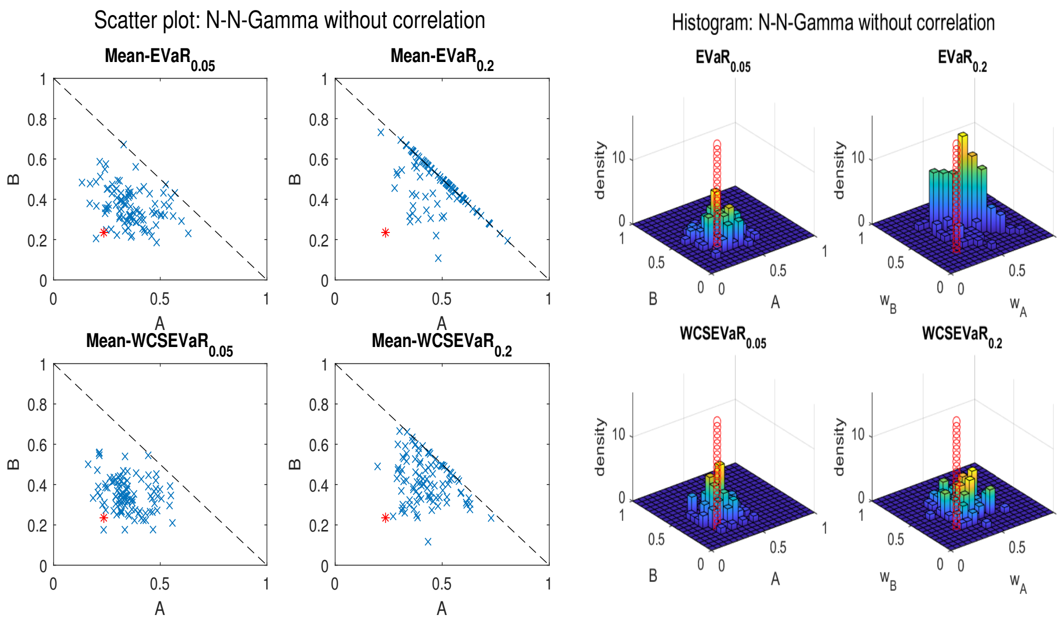


Figure 6.6: Scatter plot and histogram of the weights of the assets A and B, the weight of C is given by the relation $\sum \omega_i = 1$, the pattern is the same of Fig. 6.5 with 100 simulations. The red points represent the theoretical value of the portfolio weights given by the minimal variance portfolio.

⟨WNGGuncorr⟩

In conclusion, introducing correlation among the skewed data, in Fig. 6.7 and 6.8 we can observe the same behavior seen in the previous examples: the robust case seems to have slightly less dispersion than the $EVaR_\tau$ case.
Consequently there is a tradeoff between choosing a radically more complex NLP problem to reduce slightly the variability of the portfolio and choosing a significantly easier LP problem with little more variability.

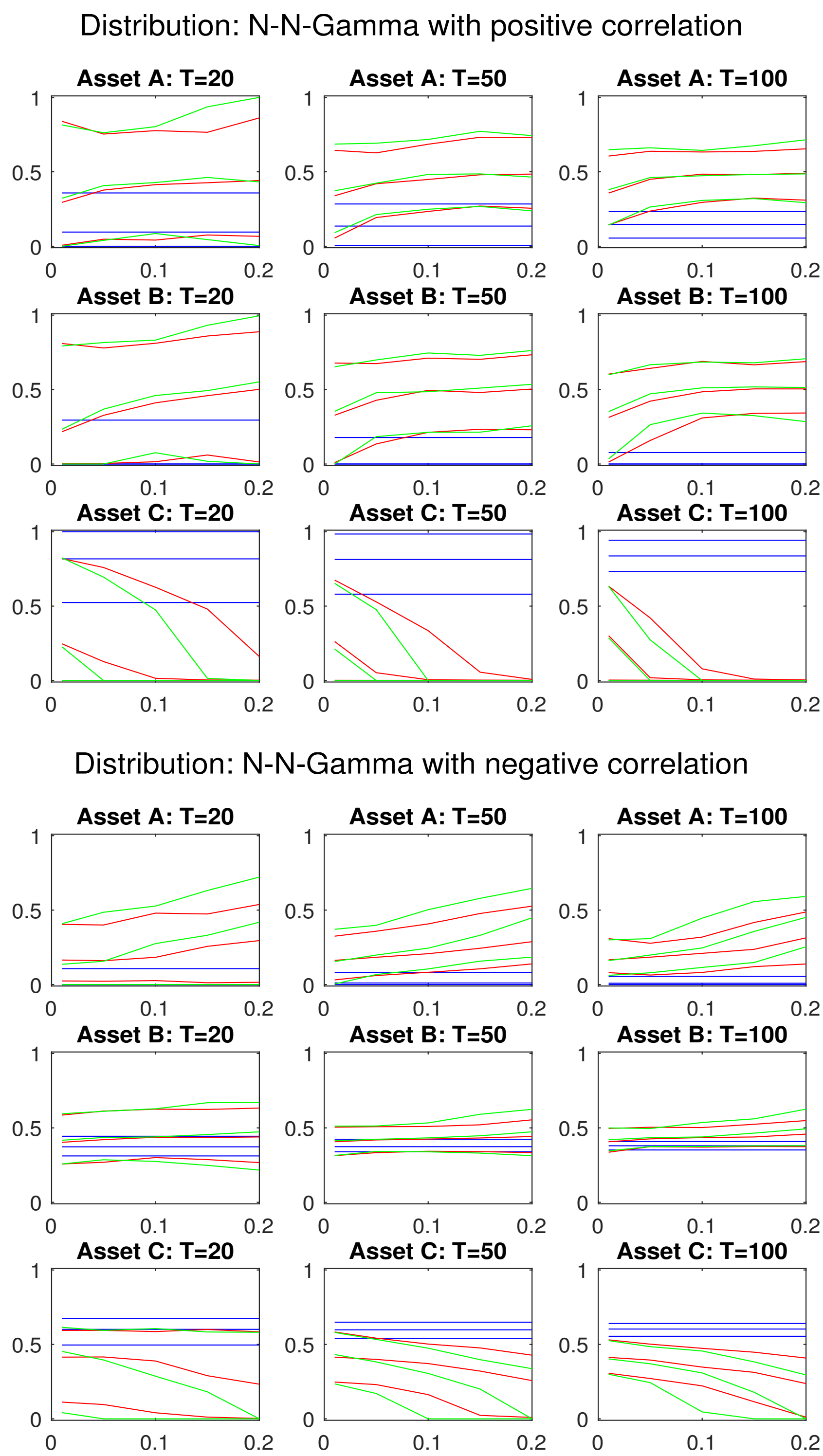


Figure 6.7: Variability of the weights of the portfolios found minimizing $WEVaR_\tau$ for $\tau = 0.01, \ldots, 0.2$ on simulated correlated data with correlation matrices Σ_+ and Σ_-: assets A and B come from normal distributions, asset C comes from a reflected Γ distribution. The red lines indicate respectively the 90^{th} percentile, the median and the 10^{th} percentile of the portfolio weights. The blue lines and the green lines represent respectively the same quantities for a minimal variance portfolio and a minimal $EVaR_\tau$ portfolio.

⟨WNGGcorrber⟩

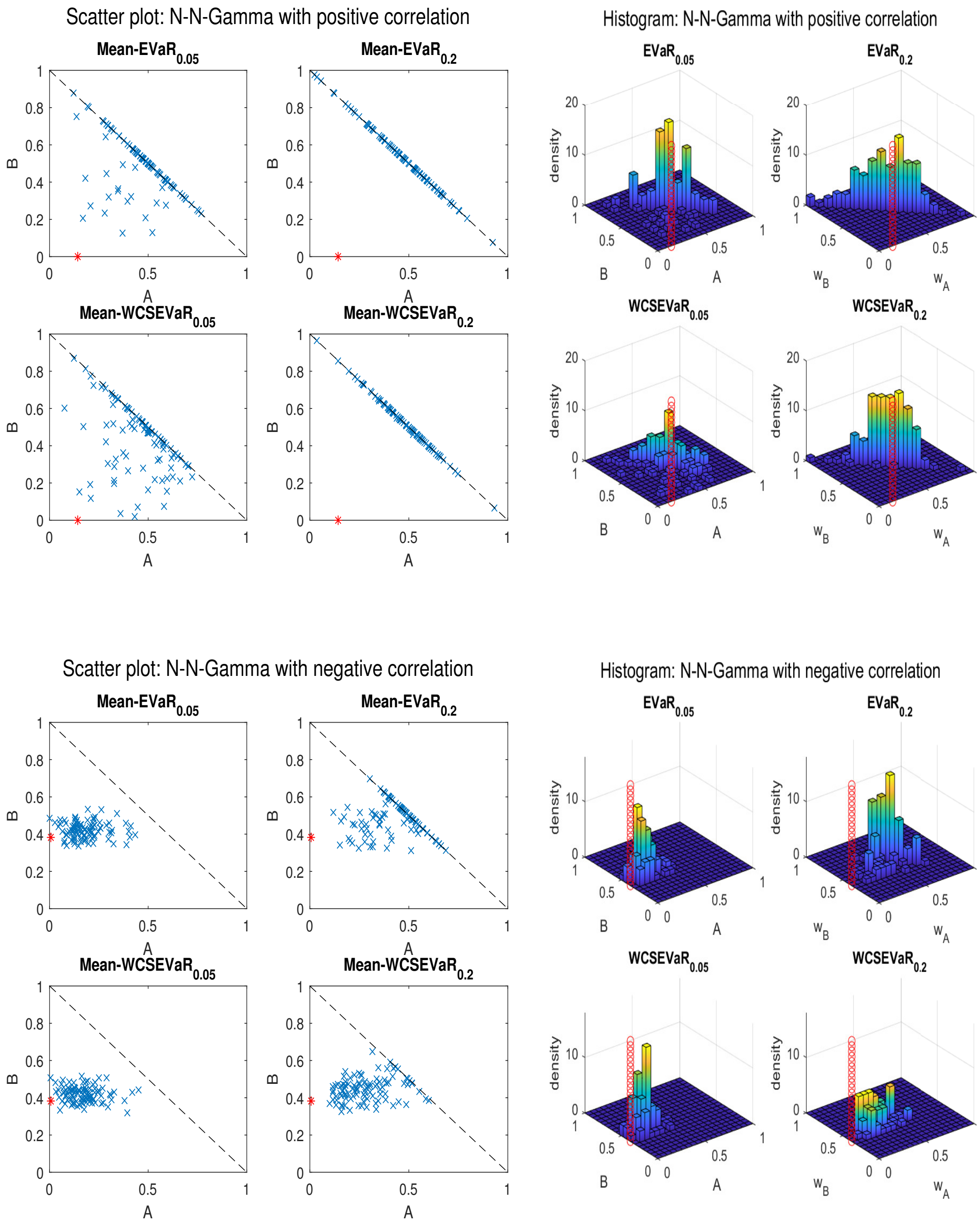


Figure 6.8: Scatter plot and histogram of the weights of the assets A and B, the weight of C is given by the relation $\sum \omega_i = 1$, the pattern is the same of Fig. 6.7 with 100 simulations. The red points represent the theoretical value of the portfolio weights given by the minimal variance portfolio.

⟨WNGGcorr⟩

6.4.2 Historical Data

In this section we perform tests on $WEVaR_\tau$ using the same dataset of section 4.4.2, i.e. using the logreturns of the companies that constitute the $S\&P$ 500 index.
In our analysis we consider two different periods: January 2011 - December 2013 and January 2014 - December 2016.

First we evaluate the optimal portfolios' frontiers evaluated considering the whole dataset over the two periods.
Afterwards, we make an *out-of-sample* analysis considering a rolling window of length 100 days giving equal probability 1/100 to the possible outcomes: every day the portfolio is rebalanced following the rules of Markowitz portfolio theory, mean-$EVaR_{0.05}$ analysis, mean-$WEVaR_{0.05}$ analysis. For the sake of completeness we compare the results with the equally weighted portfolio.
In particular we evaluate the empirical distribution of the realized logreturns and the relative statistics.
Similar experiments are performed for a portfolio composed of the first 250 assets of the $S\&P$ 500 in order to asses the stability of the algorithm with large portfolios.

In Fig. 6.9 and 6.10 we present the comparison between the optimal portfolios' frontiers of the different methods varying τ: the portfolio is composed again of the first five assets of $S\&P$ 500 ordered by their weight in the index (Apple Inc., Microsoft Corporation, Exxon Mobil Corporation, Amazon.com Inc. and JPMorgan Chase & Co). We evaluate the frontiers considering the historical logreturns over the two different periods. As we expect, the optimal portfolios' frontiers are convex in all the cases and there are not substantial differences between the $EVaR_\tau$ case and his robust version.

In Fig. 6.11 and 6.12 the portfolio weights of the optimal portfolios' frontiers are represented, τ is set to 0.025: as we expect from theory, the graphic of the mean-$EVaR_{0.025}$ portfolio weights is quite similar to graphic of the relative robust version, mean-$WEVaR_\tau$ portfolio seems to diversify more buying some asset of JPMorgan Chase & Co and Microsoft.

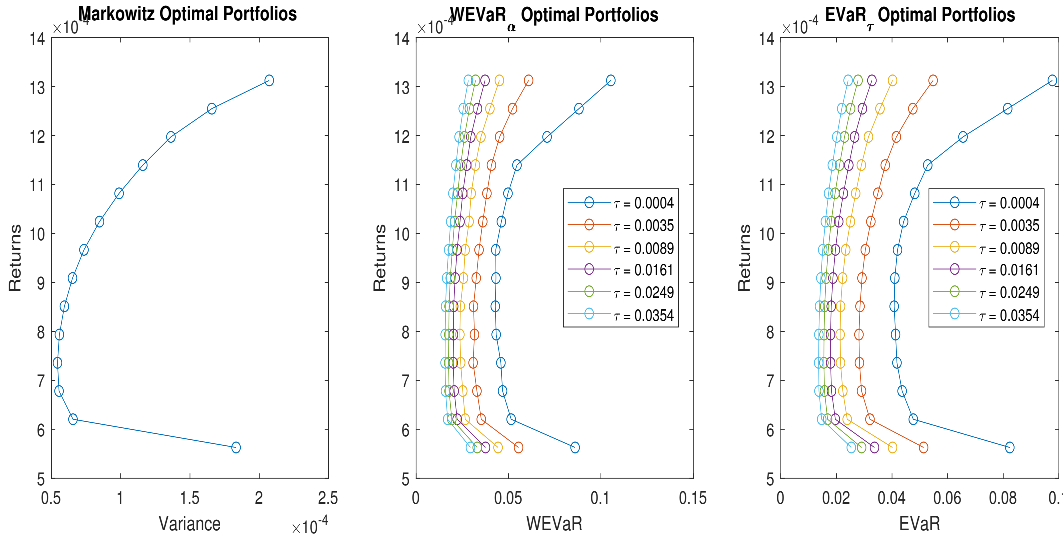


Figure 6.9: Optimal portfolios' frontiers evaluated with the Markowitz portfolio theory, the mean-$WEVaR_\tau$ analysis and the mean-$EVaR_\tau$ analysis for different values of α and τ during the period January 2011 December 2013.

⟨45ef⟩

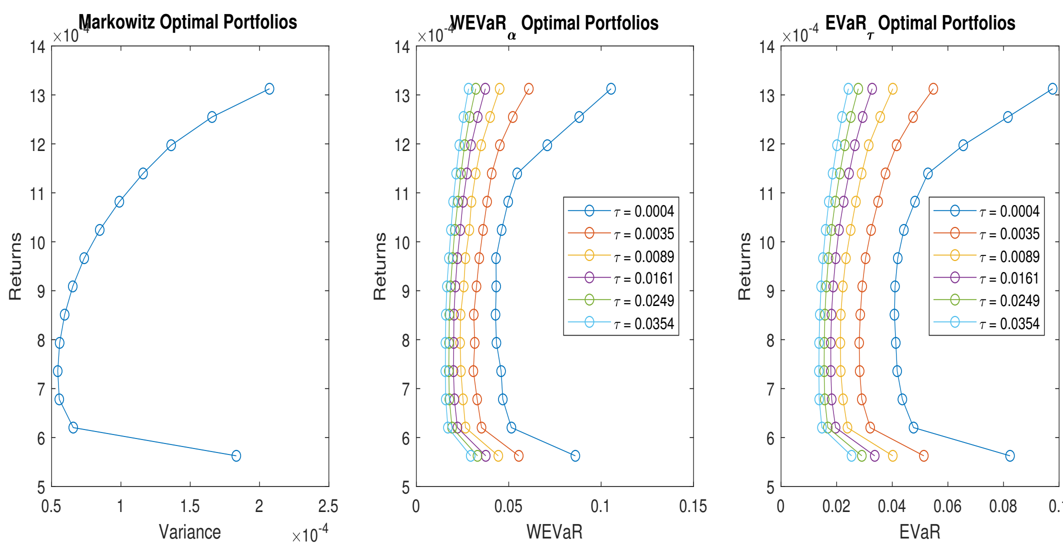


Figure 6.10: Optimal portfolios' frontiers evaluated with the Markowitz portfolio theory, the mean-$WEVaR_\tau$ analysis and the mean-$EVaR_\tau$ analysis for different values of α and τ during the period January 2014 - December 2016.

⟨45efa⟩

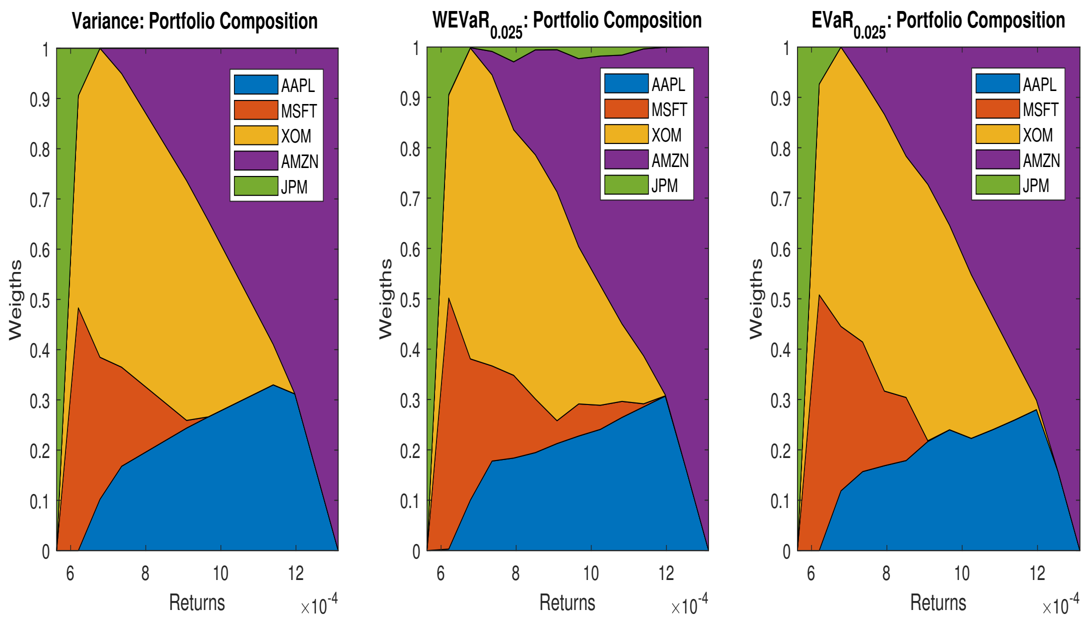


Figure 6.11: Portfolio weights of the optimal portfolios evaluated in the different methods during the period January 2011 - December 2013.

⟨45efpw⟩

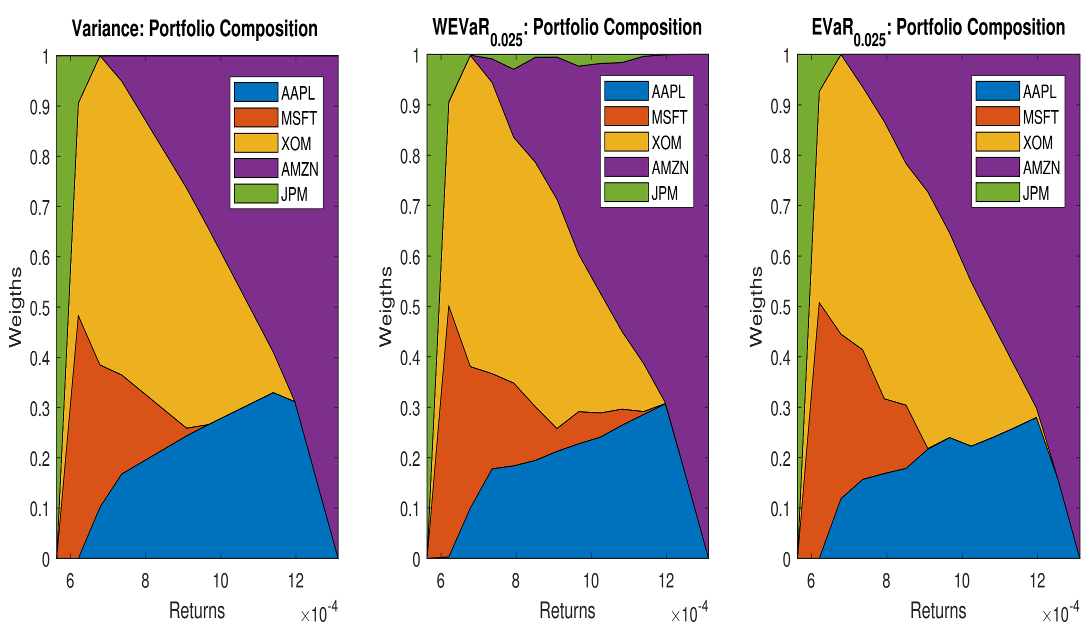


Figure 6.12: Portfolio weights of the optimal portfolios evaluated in the different methods during the period January 2014 - December 2016.

⟨45efpwa⟩

In Fig. 6.13 and 6.14 the empirical distribution of the out-of-sample portfolios' logreturns is represented, and in Tab. 6.1 and 6.2 the extrema and the moments of the distributions are evaluated.

Each day we compute the realized logreturn of the mean-risk portfolio evaluated the day before. We consider variance, $EVaR_{0.05}$, $WEVaR_{0.05}$ as risk measures and we compare them with the realization of the equally weighted portfolio. In order to focus on the upper part of the efficient frontier, in addition the non negativity constraint on the portfolio weights, we added again the same lower bound on expected logreturn as in Sections 4.4.2 and 5.2.2.

1^{st} January 2011 - 31^{th} December 2013

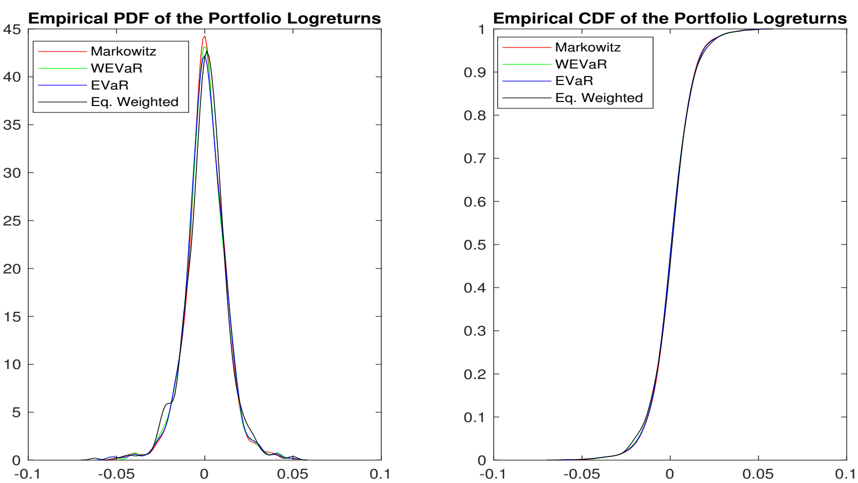


Figure 6.13: Empirical distribution of the 5-asset-portfolio logreturns computed out-of-sample for the period Jan11-Dec13.

⟨4511-13⟩

	Markowitz	**WEVaR**	**EVaR**	**Eq. Weighted**
Min	-5.08e-02	-5.39e-02	-5.36e-02	-6.24e-02
Max	4.76e-02	4.90e-02	5.00e-02	5.04e-02
Mean	6.77e-04	5.85e-04	6.26e-04	6.78e-04
Std	1.12e-02	1.15e-02	1.15e-02	1.19e-02
Skew	-1.48e-01	-1.48e-01	-1.68e-01	-2.49e-01
Kurt	5.38	5.52	5.60	5.68

Table 6.1: Statistics of the 5-asset portfolio logreturn distributions over the period January 2011 - December 2013

⟨tablestat41⟩

1^{st} January 2014 - 31^{th} December 2016

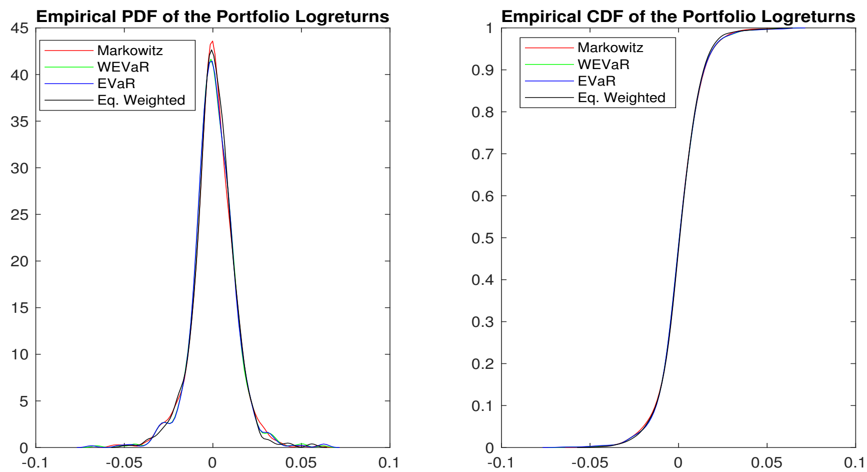


Figure 6.14: Empirical distribution of the 5-asset-portfolio logreturns computed out-of-sample for the period Jan14-Dec16.

⟨4514-16⟩

	Markowitz	**WEVaR**	**EVaR**	**Eq. Weighted**
Min	-5.63e-02	-6.56e-02	-6.86e-02	-4.95e-02
Max	6.01e-02	6.19e-02	6.33e-02	5.67e-02
Mean	5.55e-04	6.05e-04	6.09e-04	6.28e-04
Std	1.19e-02	1.20e-02	1.21e-02	1.15e-02
Skew	-1.68e-01	-9.57e-02	-5.86e-02	-3.00e-03
Kurt	5.79	6.72	7.40	5.71

Table 6.2: Statistics of the 5-asset portfolio logreturn distributions over the period January 2014 - December 2016

⟨tablestat42⟩

Looking at the empirical distributions we do not notice substantial differences among the different cases, the portfolio which minimize $WEVaR_{0.05}$ has performances comparable with the Markowitz portfolio and the mean-$EVaR_{0.05}$ portfolio: the statistics reported in Tab. 6.1 and 6.2 are comparable.

Let now consider large portfolio. In Fig. 6.15 we present the comparison between the optimal portfolios' frontiers of the different methods varying τ: the portfolio is composed of the first 250 assets considering the historical distribution

of the logreturns from 1^{st} January 2012 to 31^{th} December 2016.
The efficient frontier in the $WEVaR_\tau$ is not convex for all τ due to numerical issue: the problem is NLP with 751 variables and 501 constraints and the optimizer may find local minima especially when τ is small.
In Fig. 6.17 and 6.18 the empirical distribution of the out-of-sample portfolios' logreturns are represented and in the Tab. 6.3 and 6.4, the extrema and the moments of the distributions are evaluated.

Looking at Fig. 6.17 and 6.18 we notice that minimizing $WEVaR_\tau$ we find empirical distribution with higher variance, and considering the moments of the distributions of the realized logreturns of large portfolios we can observe that minimizing $WEVaR_\tau$ we obtain discordant results.

We conclude that the proposed method applied to large portfolio is not stable. The problem should be rewritten in an easier way reducing the number of variables and constraints and trying to avoid as more as possible the number of nonlinear constraints. In future studies we will try to analyze more in detail this issue.

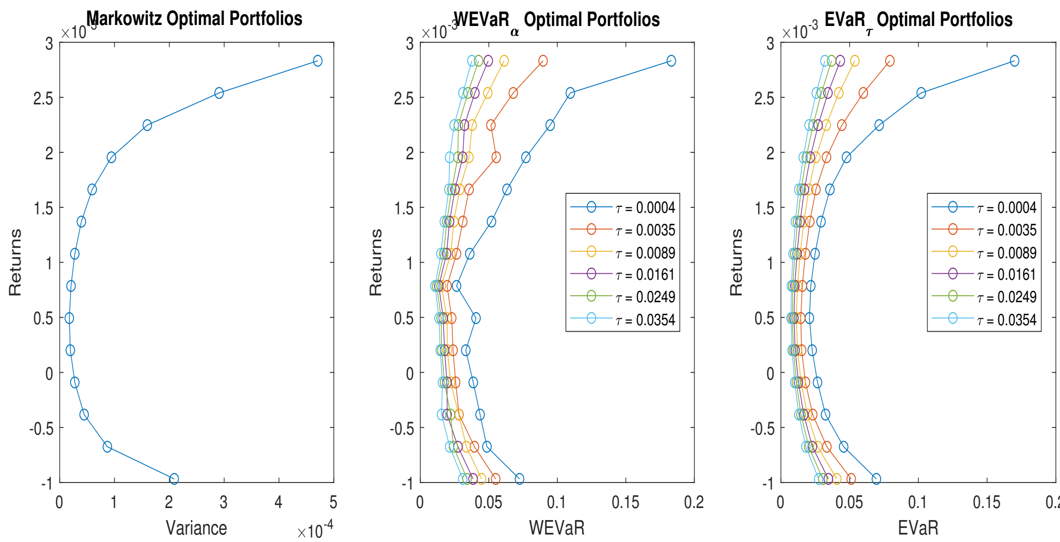


Figure 6.15: Optimal portfolios' frontiers evaluated with the Markowitz portfolio theory, the mean-$WEVaR_\tau$ analysis and the mean-$EVaR_\tau$ analysis for different values of α and τ during the period January 2011 - December 2013.

⟨4250ef⟩

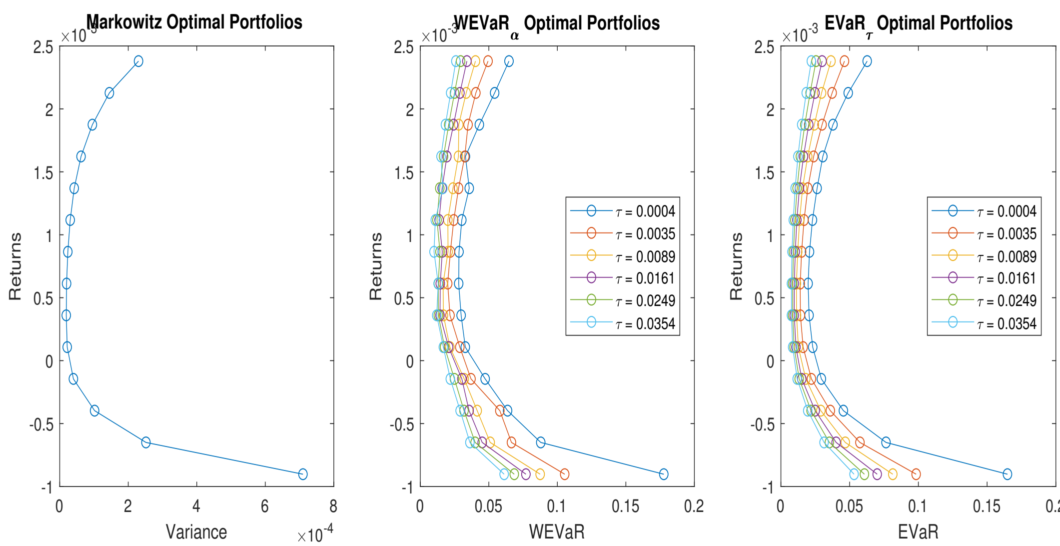


Figure 6.16: Optimal portfolios' frontiers evaluated with the Markowitz portfolio theory, the mean-$WEVaR_\tau$ analysis and the mean-$EVaR_\tau$ analysis for different values of α and τ during the period January 2014 - December 2016.

?⟨4250efa⟩?

1^{st} January 2011 - 31^{th} December 2013

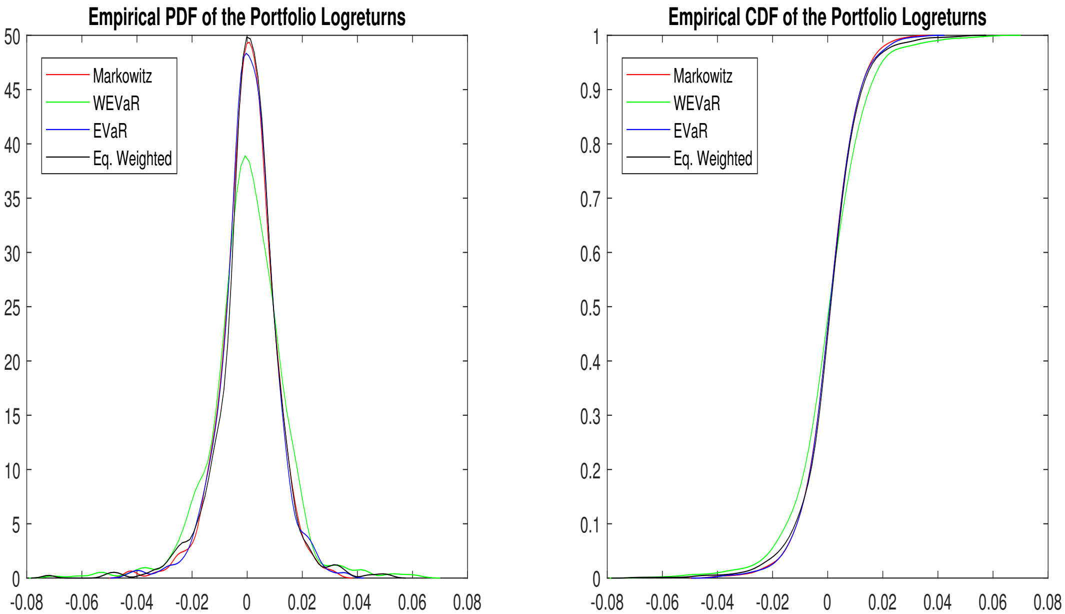


Figure 6.17: Empirical distribution of the 250-asset-portfolio logreturns computed out-of-sample for the period Jan11-Dec13.

⟨425011-13⟩

	Markowitz	**WEVaR**	**EVaR**	**Eq. Weighted**
Min	-4.34e-02	-7.04e-02	-4.31e-02	-7.20e-02
Max	3.16e-02	6.13196e-02	3.59116e-02	5.10979e-02
Mean	4.64e-04	3.78e-04	5.64e-04	6.83e-04
Std	9.53e-03	1.32e-02	9.69e-03	1.10e-02
Skew	-4.50e-01	-2.53e-01	-3.55e-01	-5.83e-01
Kurt	5.00	7.08	5.14	8.41

Table 6.3: Statistics of the 250-asset portfolio logreturn distributions over the period January 2011 - December 2013.

⟨tablestat43⟩

1^{st} January 2014 - 31^{th} December 2016

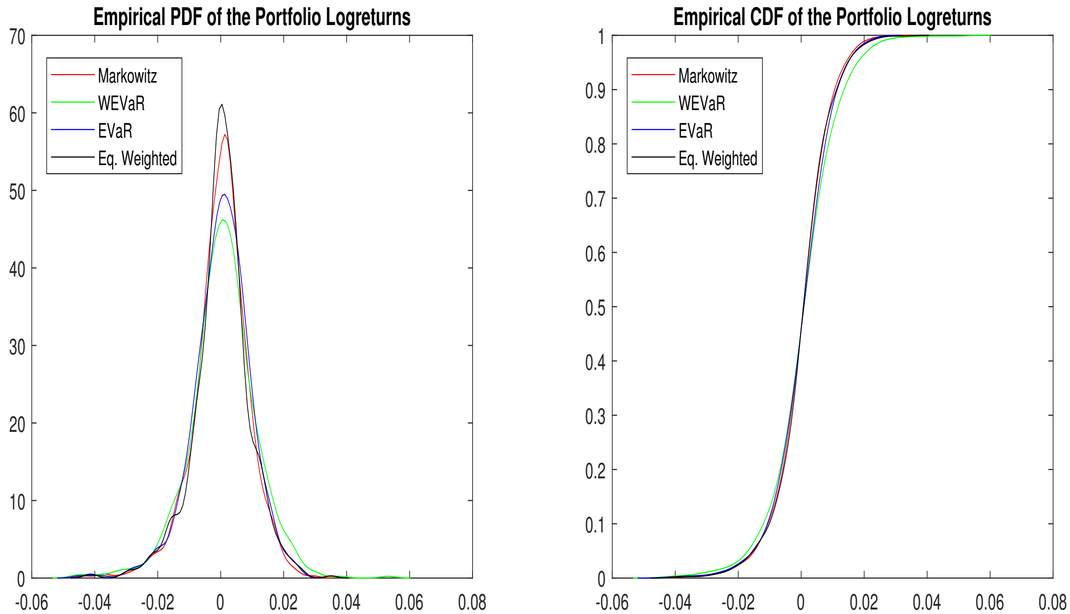


Figure 6.18: Empirical distribution of the 250-asset-portfolio logreturns computed out-of-sample for the period Jan14-Dec16.

⟨425014-16⟩

	Markowitz	**WEVaR**	**EVaR**	**Eq. Weighted**
Min	-4.30e-02	-4.63e-02	-4.52e-02	-4.21e-02
Max	2.94e-02	5.31e-02	2.50e-02	3.50e-02
Mean	2.24e-04	7.39e-04	4.06e-04	4.11e-04
Std	8.65e-03	1.05e-02	9.05e-03	8.81e-03
Skew	-5.97e-01	-2.61e-01	-6.13e-01	-4.35e-01
Kurt	5.10	5.33	4.93	5.13

Table 6.4: Statistics of the 250-asset portfolio logreturn distributions over the period January 2014 - December 2016.

⟨tablestat44⟩

Chapter 7

Risk Parity Portfolios with Expectiles

⟨capriskparity⟩

7.1 Risk Parity

In recent years, the theory of risk parity has prompted interest due to the necessity for the institutional investors to dealing with "risk budgeting", that is the analysis in term of risk contributions instead of the usual mean-risk analysis. The basic idea is to build a portfolio in which every title contributes with the same amount of risk to the global risk of the portfolio. In the work of Maillard et al. (2010) it is presented the construction of a portfolio in which every component gives the same marginal volatility contribution to the total risk.

Definition 37. *Let $J = \{1, \ldots, m\}$ be the set of assets taken into account for a possible investment and r_j the random variable which describes the logreturn of the asset $j \in J$, the expected logreturn is given by $\mu_j = \mathbb{E}[r_j]$ and the related covariance matrix is given by $\Sigma = [\sigma_{i,j}]_{i,j=1,\ldots,m} = [\mathbb{E}[(r_i - \mu_i)(r_j - \mu_j)]]_{i,j=1,\ldots,m}$.*
Let $\sigma(w) = \sqrt{w^T \Sigma w} = \sqrt{\sum_{i,j} w_i w_j \sigma_{i,j}}$ be the standard deviation of the portfolio with weights $w = w_1, \ldots, w_m$.
Marginal risk contributions, are defined as

$$\partial_{w_j} \sigma(w) = \frac{\partial \sigma(w)}{\partial w_j} = \frac{w_j (\Sigma w)_j}{\sigma(w)}. \tag{7.1}$$ marginalP

Thus, the equally-weighted risk contribution (ERC) portfolio must satisfy:

$$w_i \partial_{w_i} \sigma(w) = w_j \partial_{w_j} \sigma(w), \quad \forall i, j = 1, \ldots, m. \tag{7.2}$$ propERC

that is the marginal risk contribution of an asset times his weight must be equal to a constant. Thus the ERC portfolio can be found solving

$$\begin{aligned} w^\star = &\operatorname{argmin} \sum_{i=1}^m \sum_{j=1}^m (\partial_{w_i}\sigma(w) - \partial_{w_j}\sigma(w))^2 \quad s.t. \\ &\sum_{i=1}^m w_i = 1 \\ &w_i \geq 0. \end{aligned}$$

Considering eq. (7.1), this problem is equivalent to solve

$$\begin{aligned} w^\star = &\operatorname{argmin} \sum_{i=1}^m \sum_{j=1}^m (w_i(\Sigma w)_i - w_j(\Sigma w)_j)^2 \quad s.t. \\ &\sum_{i=1}^m w_i = 1 \\ &w_i \geq 0. \end{aligned} \tag{7.3}$$ RP

The existence of the ERC portfolio is ensured only when the value of the objective function in (7.3) evaluated in $w^\star$ is equal to 0, which implies that $w_i(\Sigma w)_i = w_j(\Sigma w)_j$ for all i, j.

An alternative approach of finding the ERC portfolio consists in solving the following optimization problem:

$$\begin{aligned} y^\star = &\operatorname{argmin} \sqrt{y^t \Sigma y} \quad s.t. \\ &\sum_{i=1}^m \ln(y_i) \geqslant c \\ &y_i \geqslant 0 \quad \forall i = 1, \dots, m, \end{aligned} \tag{7.4}$$ RPlog

where c is an arbitrary constant. The ERC portfolio is given by the normalization of the solution $y^\star$:

$$w_i^\star = \frac{y_i^\star}{\sum_{j=1}^m y_j^\star}.$$

In fact if we compute the Lagrangean of the problem (7.4), we obtain:

$$l(y; \lambda, \lambda_c) = \sqrt{y^T \Sigma y} - \lambda^T y - \lambda_c \left(\sum_{i=1}^m \ln(y_i) - c \right).$$

The solution $y^\star$ must satisfy the first-order condition and the Kuhn-Tucker conditions, i.e.

$$\begin{cases} \partial_{y_i} l(y;\lambda,\lambda_c) = \partial_{y_i}\sigma(y) - \lambda_i - \frac{\lambda_c}{y_i} = 0, \quad \forall i = 1,\dots,m \\ \lambda_i y_i = 0, \quad \forall i = 1,\dots,m \\ \lambda_c \left(\sum_{i=1}^m \ln(y_i) - c\right) = 0 \end{cases} \tag{7.5}$$ eqnlag

Since y_i must be positive, we have that $\lambda_i = 0$. Moreover the constraint $\sum_{i=1}^m \ln(y_i) = c$ is reached, in fact otherwise $\lambda_c = 0$, that cannot be true because it should imply:

$$\partial_{y_i} l(y;0,0) = \partial_{y_i}\sigma(y) = 0 \quad \Rightarrow \quad y_i(\Sigma y)_i = 0 \quad \Rightarrow \quad y_i = 0,$$

which is impossible.
Hence, from (7.5) we have that

$$\partial_{y_i} l(y;0,\lambda_c) = \partial_{y_i}\sigma(y) - \frac{\lambda_c}{y_i} = 0 \quad \Rightarrow \quad y_i\partial_{y_i}\sigma(y) = \lambda_c \quad \forall i = 1,\dots,m.$$

which is the the property (7.2) of the ERC portfolio.

7.1.1 Risk Parity for differentiable Risk Measures

Immediately after the publication of Maillard et al. (2010), many authors extended it using other risks measure: Stefanovits (2010) introduced the ERC portfolio using Value-at-Risk and Conditional Value-at-Risk as risk measures, Cesarone and Colucci (2017) rewrote the problem of the ERC portfolio using CVaR using the formulation of Rockafellar and Uryasev (2000).
More generally, the risk parity pattern can be extended to any differentiable risk measure:

Definition 38. *Let $J = \{1,\dots,m\}$ be the set of assets taken into account for a possible investment and r_j the random variable which describes the logreturn of the asset $j \in J$, $w = w_1,\dots,w_m$ the weights of the portfolio and $\rho(w)$ a differentiable risk measure.*
Marginal risk contributions, $\partial_{w_j}\rho(w)$, are defined as

$$\partial_{w_j}\rho(w) = \frac{\partial\rho(w)}{\partial w_j}.$$

Hence the ERC portfolio must satisfy the property

$$w_i\partial_{w_i}\rho(w) = w_j\partial_{w_j}\rho(w) \quad \forall i,j = 1,\dots,m \quad \Rightarrow \quad w_j\partial_{w_j}\rho(w) = \lambda.$$

Generalizing 7.4 we have that the ERC portfolio related to the risk measure ρ is given by

$$w_i^\star = \frac{y_i^\star}{\sum_{j=1}^m y_j^\star}, \tag{7.6}$$ wstar

where $y^\star$ is the solution of the problem

$$\begin{aligned} y^\star = &\operatorname{argmin} \rho(y) \quad s.t. \\ &\sum_{i=1}^m \ln(y_i) \geqslant c \\ &y_i \geqslant 0 \quad \forall i = 1, \ldots, m. \end{aligned} \tag{7.7}$$ RPlog2

In fact if we compute the Lagrangean of the problem (7.7) we have that:

$$l(y; \lambda, \lambda_c) = \rho(y) - \lambda^T y - \lambda_c \left(\sum_{i=1}^m \ln(y_i) - c \right).$$

The solution $y^\star$ must satisfy the first-order condition and the Kuhn-Tucker conditions, i.e.

$$\begin{cases} \partial_{y_i} l(y; \lambda, \lambda_c) = \partial_{y_i} \rho(y) - \lambda_i - \frac{\lambda_c}{y_i} = 0, \quad \forall i = 1, \ldots, m \\ \lambda_i y_i = 0, \quad \forall i = 1, \ldots, m \\ \lambda_c \left(\sum_{i=1}^m \ln(y_i) - c \right) = 0 \end{cases}$$

As for the case of the variance we have that $y_i > 0$ and $\lambda_i = 0$ for all i, moreover the constraint $\sum_{i=1}^m \ln(y_i) = c$ is reached, in fact by contradiction if we consider the second equation of the Kuhn-Tucker conditions we have that

$$\sum_{i=1}^m \ln(y_i) > c \quad \Rightarrow \quad \lambda_c = 0.$$

Thus

$$\partial_{y_i} l(y; 0, 0) = \partial_{y_i} \rho(y) = 0, \quad \forall i = 1, \ldots, m \quad \Rightarrow \quad \rho(y) = c, \quad c \in \mathbb{R}, \quad \forall y \in \mathbb{R}^m,$$

which is absurd by the definition of risk measure.
So we have that

$$\partial_{y_i} l(y; 0, \lambda_c) = \partial_{y_i} \rho(y) - \frac{\lambda_c}{y_i} = 0, \quad \Rightarrow \quad y_i \partial_{y_i} \rho(y) = \lambda_c \quad \forall i = 1, \ldots, m.$$

7.1.2 Risk Parity for Conditional Value-at-Risk

In Cesarone and Colucci (2017) it is studied the ERC portfolio using CVaR. If we consider formulation in (7.6) and (7.7), we have that the ERC portfolio is found by solving

$$w_i^\star = \frac{y_i^\star}{\sum_{j=1}^m y_j^\star},$$

where $y^\star$ is the solution of the problem

$$\begin{aligned} y^\star = &\operatorname{argmin} CVaR_\alpha(y) \quad s.t. \\ &\sum_{i=1}^m \ln(y_i) \geqslant c \\ &y_i \geqslant 0 \quad \forall i = 1, \dots, m. \end{aligned} \tag{7.8}$$

RPCVaR

Considering the formulation of the mean-CVaR problem in Rockafellar and Uryasev (2000) we have that the CVaR_α of a portfolio $w = (w_1, \dots, w_m)$, given the logreturns r_{jt} of the assets $j = 1, \dots, m$ in the time $t = 1, \dots, T$, is given by :

$$\begin{aligned} CVaR_\alpha(w) = &\min \zeta + \frac{1}{\alpha}\frac{1}{T}\sum_{t=1}^T d_t \quad \text{s.t.} \\ &d_t \geqslant \sum_{j=1}^m -r_{jt}w_j - \zeta, \quad t = 1, \dots, T \\ &d_t \geqslant 0, \quad t = 1, \dots, T, \\ &\zeta \in \mathbb{R} \end{aligned} \tag{7.9}$$

CVaRrocka

Hence substituting (7.9) in (7.8) we obtain:

$$\begin{aligned} y^\star = &\operatorname{argmin} \zeta + \frac{1}{\alpha}\frac{1}{T}\sum_{t=1}^T d_t \quad \text{s.t.} \\ &d_t \geqslant \sum_{j=1}^m -r_{jt}y_j - \zeta, \quad t = 1, \dots, T \\ &\sum_{i=1}^m \ln(y_i) \geqslant c \\ &d_t \geqslant 0, \quad t = 1, \dots, T, \\ &\zeta \in \mathbb{R} \\ &y_i \geqslant 0 \quad \forall i = 1, \dots, m. \end{aligned}$$

In the next section we are going to study the *equally risk contribution portfolio* using the *Expectile Value-at-Risk* as risk measure: following the idea of Maillard et al. (2010), the aim consists in finding the portfolio in which all the components give the same marginal EVaR_τ contribution to the global risk of the portfolio.

7.2 Formulation

Considering the formulation in (7.6) and (7.7), the ERC portfolio using the EVaR_τ as risk measure is given by:

$$w_i^\star = \frac{y_i^\star}{\sum_{j=1}^m y_j^\star},$$

where $y^\star$ is the solution of the problem

$$\begin{aligned} y^\star = &\operatorname{argmin} EVaR_\tau(y) \quad s.t. \\ &\sum_{i=1}^m \ln(y_i) \geqslant c \\ &y_i \geqslant 0 \quad \forall i = 1, \dots, m. \end{aligned}$$

Considering the dual formulation of EVaR in (3.4), we have that:

$$\begin{aligned} y^\star = &\operatorname{argmin} && \max \mathbb{E}[-y\varphi] \quad s.t. \\ &\sum_{i=1}^m \ln(y_i) \geqslant c && \varphi \in \mathcal{M}_\tau \\ &y_i \geqslant 0 \quad \forall i = 1, \dots, m. && \end{aligned}$$

where

$$\mathcal{M}_\tau = \left\{ \varphi \in L^\infty,\ \varphi \geqslant 0 \text{ a.s.},\ \mathbb{E}[\varphi] = 1,\ \frac{\operatorname{ess\,sup} \varphi}{\operatorname{ess\,inf} \varphi} \leqslant \frac{1-\tau}{\tau} \right\}$$

Hence in the discrete case:

$$\begin{aligned} y^\star = &\operatorname{argmin} && \max \sum_{t=1}^T -f_t p_t (Ry)_t \quad s.t. \\ &\sum_{i=1}^m \ln(y_i) \geqslant c && \sum_{t=1}^T f_t p_t = 1 \\ &y_i \geqslant 0 \quad \forall i = 1, \dots, m && m \leqslant f_t \leqslant \beta m,\ \forall t \\ & && m \geq 0,\ f_t \geq 0\ \forall t \end{aligned} \tag{7.10}$$

RPEVaR3

where $R \in \mathbb{R}^{T \times m}$ is the matrix of the historical logreturns of the m assets, $\beta = (1-\tau)/\tau$ and p_t is the probability associated to the t^{th} scenario.

We can observe that

$$\begin{aligned}
&\max \sum_{t=1}^{T} -f_t p_t (Ry)_t \quad s.t. \\
&\sum_{t=1}^{T} f_t p_t = 1 \\
&m \leqslant f_t \leqslant \beta m,\ \forall t \\
&m \geq 0,\ f_t \geq 0 \forall t
\end{aligned}$$

is a linear programming problem with parameter y, hence is dual is:

$$\begin{aligned}
&\min_{u_t, v_t} \ \xi \quad \text{s.t.} \\
&p_t \xi - u_t + v_t \geqslant -p_t (Ry)_t \ \forall t \\
&\sum_{t=1}^{T} u_t - \beta \sum_{t=1}^{n} v_t \geqslant 0 \\
&u_t, v_t, \geqslant 0,\ \forall t = 1, \ldots, T
\end{aligned}$$

Hence the problem (7.10) becomes

$$\begin{aligned}
y^\star = \operatorname{argmin} \xi & \\
& p_t \xi - u_t + v_t \geqslant -p_t (Ry)_t \ \forall t \\
& \sum_{t=1}^{T} u_t - \beta \sum_{t=1}^{n} v_t \geqslant 0 \\
& \sum_{i=1}^{m} \ln(y_i) \geqslant c \\
& y_i \geq 0,\ \forall i = 1, \ldots, m \\
& u_t, v_t, \geqslant 0,\ \forall t = 1, \ldots, T
\end{aligned}$$

7.3 Numerical Examples

In the first part of this section we consider optimal portfolios on simulated data: our purpose is to compare the equally-weighted risk contribution portfolio using different risk measures including $EVaR_\tau$. In the second part of the section we test the method on historical data.
All the experiments are performed using MatLab and the optimization problems are solved with the functions `linprog`, `quadprog` and `fmincon`. The solutions to Linear and Quadratic Programming problems are checked with GAMS (General Algebraic Modeling System).

7.3.1 Simulated Data

As in the sections 4.4.1, we first consider three assets A, B and C with a standard normal distributions without correlation. In order to assess the variability of the optimal portfolio, we simulate samples of lenght $T = 20, 50, 100$, from the aforementioned normal distribution, and compute the portfolio in which every component give the same marginal $EVaR_\tau$ contribution to the risk for τ ranging from 1% to 20%. The results are reported in Fig. 7.1, 7.3, 7.5 and 7.7; the three lines in each subpanel correspond, respectively from top to bottom, to the 10^{th} percentile, the median and the 90^{th} percentile of the portfolio weights. In Fig. 7.2 we compare the distribution of the optimal portfolios originated by the minimization of the variance and the ERC portfolios based on variance, $EVaR_{0.05}$ and $EVaR_{0.20}$. In Fig. 7.3 and 7.4 we show two examples of correlated standard normal marginals, with correlation matrices

$$\Sigma_+ = \begin{pmatrix} 1 & 0.3 & 0.5 \\ 0.3 & 1 & 0.7 \\ 0.5 & 0.7 & 1 \end{pmatrix} \quad \text{and} \quad \Sigma_- = \begin{pmatrix} 1 & -0.3 & 0.5 \\ -0.3 & 1 & -0.7 \\ 0.5 & -0.7 & 1 \end{pmatrix}.$$

In Fig. 7.5 and 7.6 we consider uncorrelated asymmetric distributions of the following types: assets A and B have normal distributions, asset C has a reflected Γ distribution.

The parameters of the distributions are set as in the previous chapters, i.e.

$$\begin{aligned}
&\text{Asset A} \sim \mathcal{N}(\mu_{\mathcal{N}}, \sigma_{\mathcal{N}}), && \text{with} \;\; \mu_{\mathcal{N}} = 0, \;\; \sigma_{\mathcal{N}} = 2.25 \\
&\text{Asset B} \sim \mathcal{N}(\mu_{\mathcal{N}}, \sigma_{\mathcal{N}}), && \text{with} \;\; \mu_{\mathcal{N}} = 0, \;\; \sigma_{\mathcal{N}} = 2.25 \\
&\text{Asset C} \sim -\Gamma(k, \theta, \mu_{\Gamma}, \sigma_{\Gamma}), && \text{with} \;\; k = 5, \;\; \theta = 1, \;\; \mu_{\Gamma} = -0.912, \sigma_{\Gamma} = 1.5
\end{aligned} \tag{7.11}$$

parameters3

In this way we have similar left tails but significantly different right tails (we refer to Fig. 4.2).

Finally in Fig. 7.5 and 7.6 we introduce correlation among the assets: we consider two cases in which A, B and C have correlation matrices Σ_+ and Σ_- respectively.

In Fig. 7.1, 7.3, 7.5 and 7.7 we compare the composition of the ERC portfolio using the $EVaR_\tau$ as risk measure for different values of τ with the ERC using the variance as risk measure.
The distance between the 10^{th} percentile and the 90^{th} percentile is approximately constant in all cases for all values of τ and it is smaller than the Markowitz case (see Fig. 4.3 for a comparison). The distance between the 10^{th} percentile and the 90^{th} percentile in the risk parity portfolio with variance is the smallest in all cases. Moreover in all cases for greater values of the sample size T we have smaller variability, assuring the stability of the algorithm.

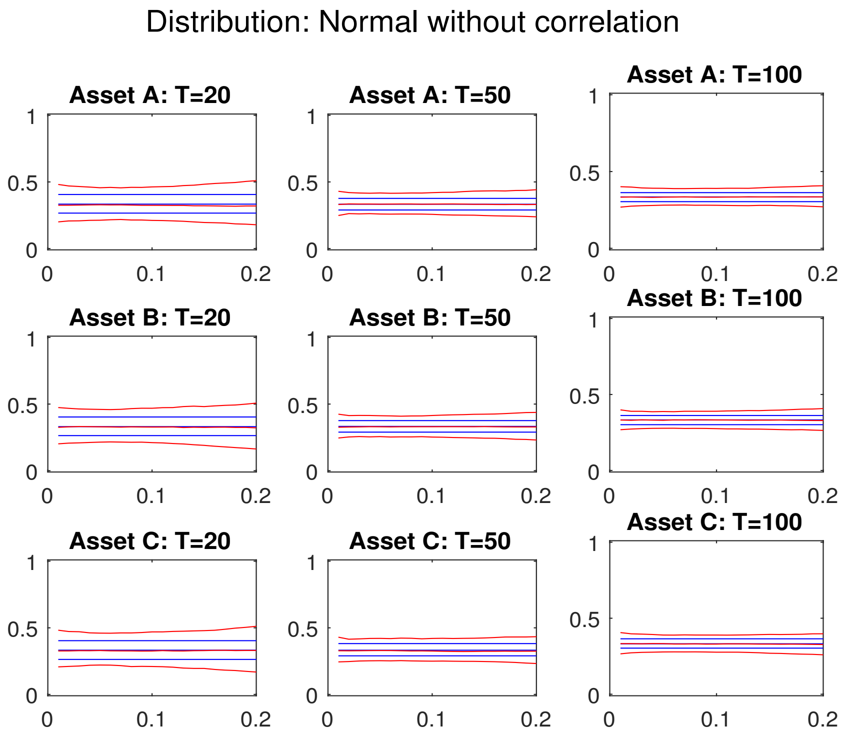


Figure 7.1: Variability of the weights of the ERC portfolios based on $EVaR_\tau$, for $\tau = 0.01, \ldots, 0.2$. The data are simulated and come from correlated standard normal distribution of length $T = 20, 50, 100$. The red lines indicate respectively the 90$^{\text{th}}$ percentile, the median and the 10$^{\text{th}}$ percentile of the portfolio weights. The blue lines represent the same quantities for a minimal variance portfolio.

⟨nuncorrbertzimas⟩

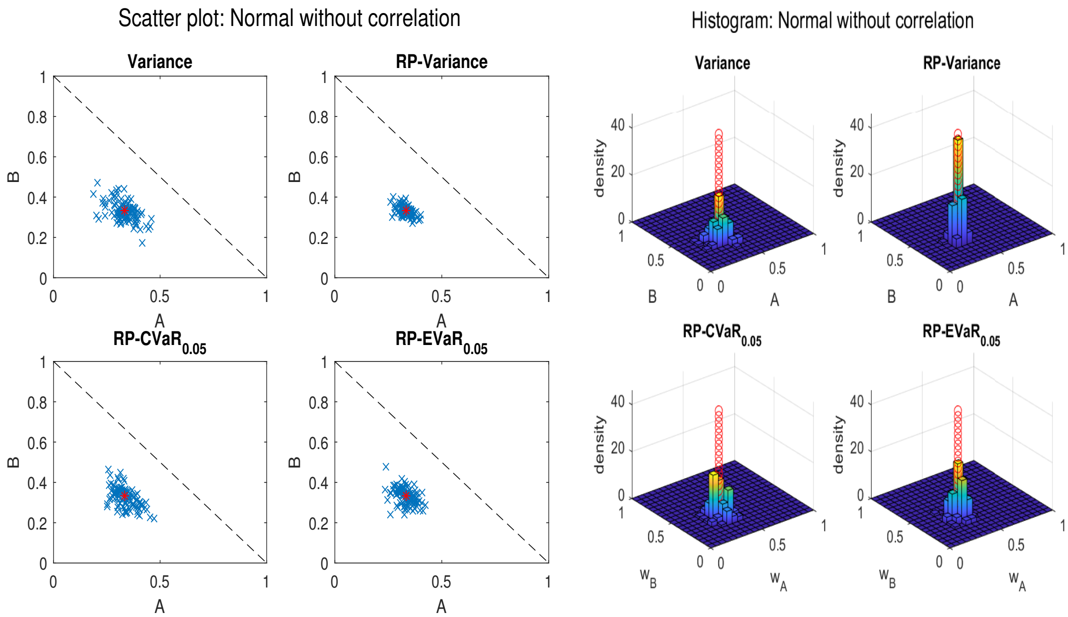


Figure 7.2: Scatter plot and histogram of the weights of the assets A and B, the weight of C is given by the relation $\sum \omega_i = 1$, the pattern is the same of Fig. 7.1 with 100 simulations. The red points represent the theoretical value of the portfolio weights given by the minimal variance portfolio.

⟨RP_nuncorrPW⟩

In Fig. 7.2, 7.4, 7.6, and 7.8 we compare the compositions of the Markowitz portfolio and the ERC portfolio using variance, $CVaR_{0.05}$ and $EVaR_{0.05}$ as risk measures.

Considering the uncorrelated standard normal case (Fig. 7.2) the cloud of points has the same shape different dispersion: the ERC portfolio using variance is the closest to the center of the distribution, while in the other cases we cannot notice substantial differences. Considering the RP-EVaR$_{0.05}$, it seems to have less dispersion than RP-CVaR$_{0.05}$.

Introducing correlation between the data (Fig. 7.3 and 7.4) we can observe that the risk parity approach has less dispersion than the Markowitz approach and the portfolio composition is closer to the center of the graphs, giving approximately the same weight to the three assets.
Considering the RP-EVaR$_{0.05}$, it seems to have again less dispersion than RP-CVaR$_{0.05}$ both with correlation matrices Σ_+ and Σ_-.

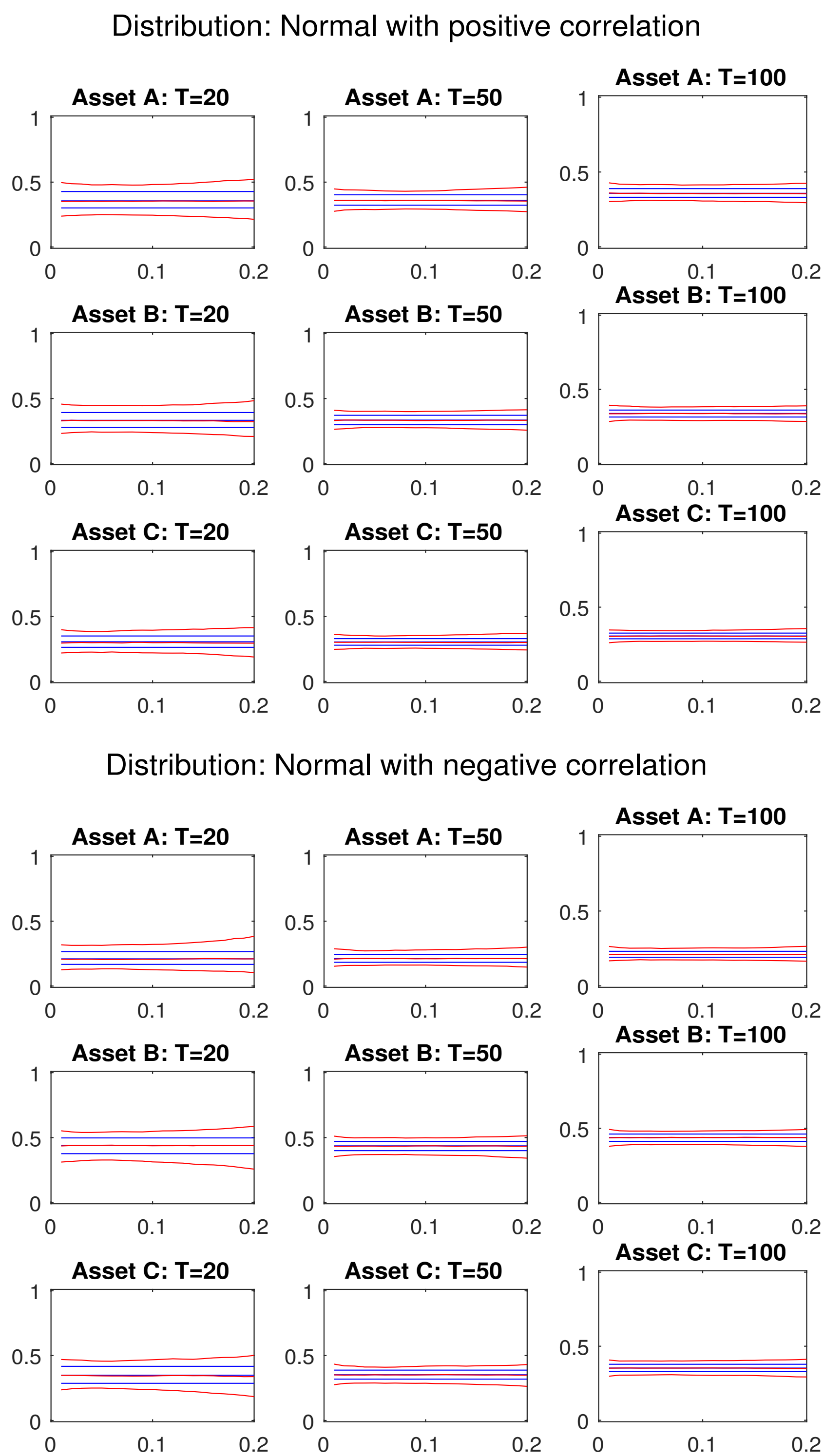


Figure 7.3: Variability of the weights of the ERC portfolios based on $EVaR_\tau$, for $\tau = 0.01, \dots, 0.2$. The data are simulated and come from uncorrelated standard normal distribution of length $T = 20, 50, 100$. The red lines indicate respectively the 90$^{\text{th}}$ percentile, the median and the 10$^{\text{th}}$ percentile of the portfolio weights. The blue lines represent the same quantities for a the risk parity based on variance portfolio.

P_ncorrbertzimas⟩

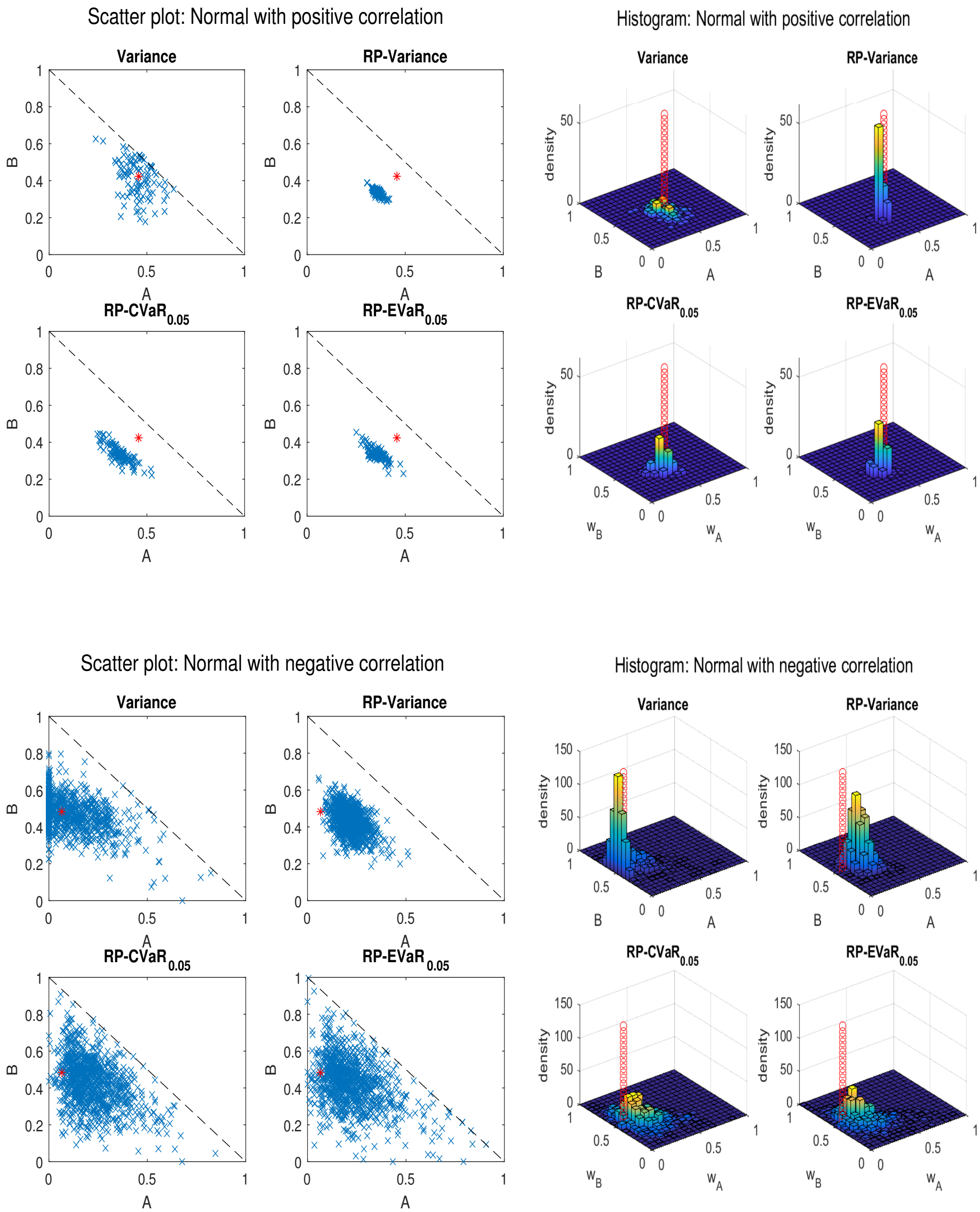


Figure 7.4: Scatter plot and histogram of the weights of the assets A and B, the weight of C is given by the relation $\sum \omega_i = 1$, the pattern is the same of Fig. 7.3 with 100 simulations. The red points represent the theoretical value of the portfolio weights given by the minimal variance portfolio.

⟨RP_ncorrPW⟩

Let now consider uncorrelated asymmetric distributions of the following types: assets A and B have normal distributions, asset C has a reflected Γ distribution, the parameters that describe the distribution are listed in eq. 7.11.

The results are reported in Fig. 7.5 and 7.6: first, in Fig. 7.5 we notice that the greater is the value of τ, the more asset C is penalized.
While in the Markowitz and in the risk parity with variance cases the clouds of points in the scatter plot are centered in the theoretical value. Considering $CVaR_{0.05}$ as risk measure, the clouds of points gather on the diagonal, to the detriment of the asset C: considering and $EVaR_{0.05}$ as risk measure this phenomenon is more pronounced, with this method assets which are described by distribution with thin right tail are penalized.
The dispersion in the $CVaR_{0.05}$ and $EVaR_{0.05}$ cases is comparable, smaller than the Markowitz case and greater than the ERC portfolio based on variance.

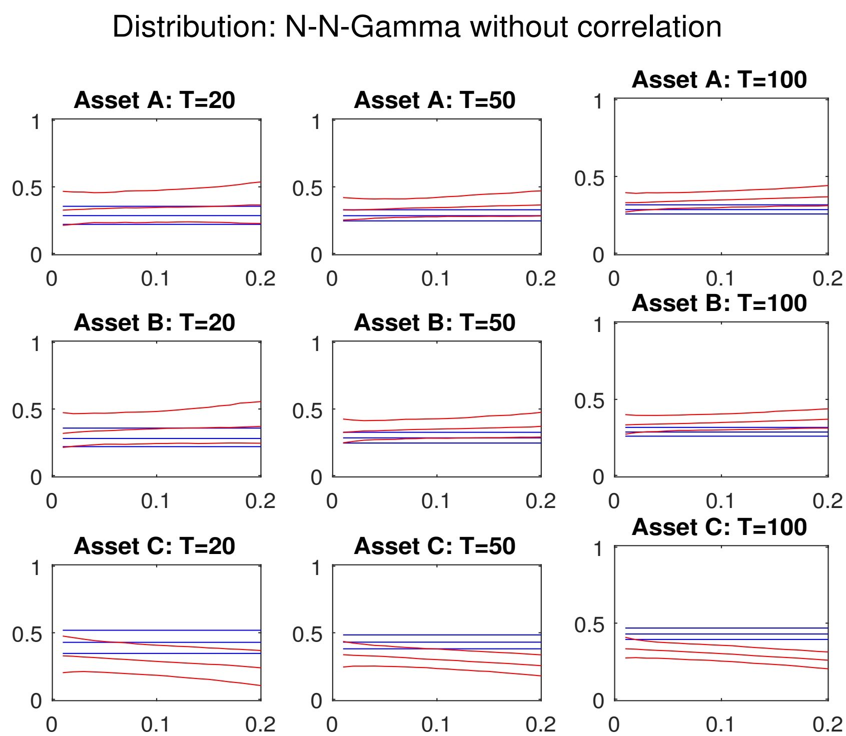


Figure 7.5: Variability of the weights of the ERC portfolios based on $EVaR_{\tau}$, for $\tau = 0.01, \ldots, 0.2$. Asset A and B have normal distributions, asset C has reflected a Γ distribution. The red lines indicate the 90$^{\text{th}}$ percentile, the median and the 10$^{\text{th}}$ percentile of the portfolio weights.The blue lines represent the same quantities for a the risk parity based on variance portfolio.

guncorrbertzimas⟩

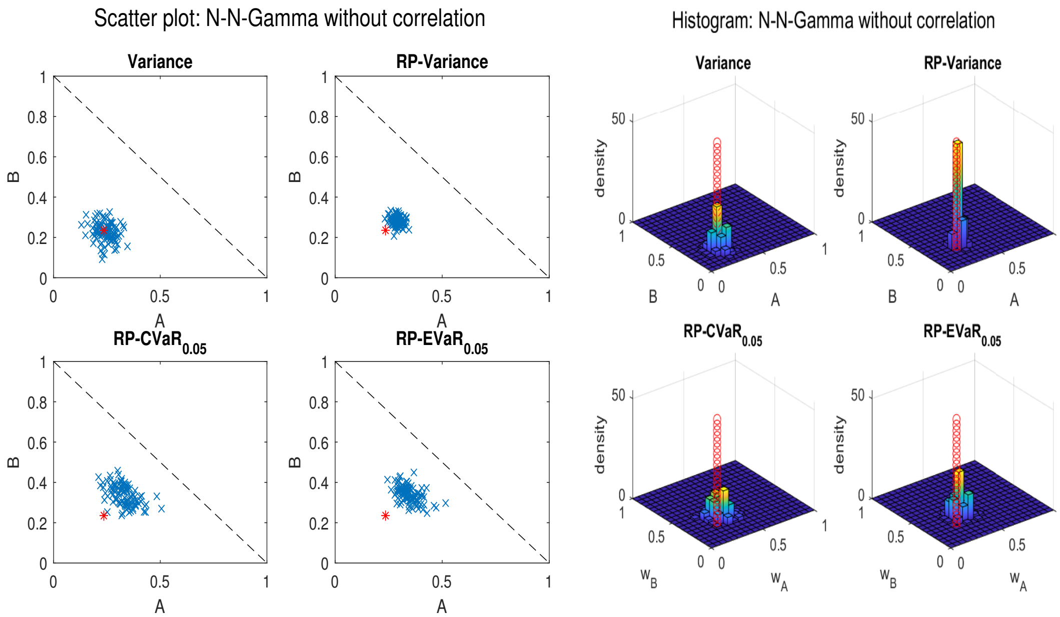


Figure 7.6: Scatter plot and histogram of the weights of the assets A and B, the weight of C is given by the relation $\sum \omega_i = 1$, the pattern is the same of Fig. 7.5 with 1000 simulations. The red points represent the theoretical value of the Markowitz portfolio weights

⟨RP_ngguncorrPW⟩

In conclusion, introducing correlation among the asymmetric data, in Fig. 7.7 and 7.8 we can observe that the results analogous to the uncorrelated case both considering correlation matrices Σ_+ and Σ_-: first, in Fig. 7.7 we notice that the greater is the value of τ, the more asset C is penalized.

Moreover looking at Fig. 7.8, the portfolio composition in the risk parity approach is closer to the center of the graph, giving approximatively equal weight to the three assets. Considering $CVaR_{0.05}$ as risk measure, the clouds of points gather on the diagonal, to the detriment of the asset C: considering and $EVaR_{0.05}$ as risk measure this phenomenon is more pronounced, with this method assets which are described by distribution with thin right tail are penalized.

As we expected from theory, with symmetric distribution we do not notice substantial differences among the proposed methods, while considering asymmetric distributions the EVaR approach allocates more risk to asset with thin right tail.

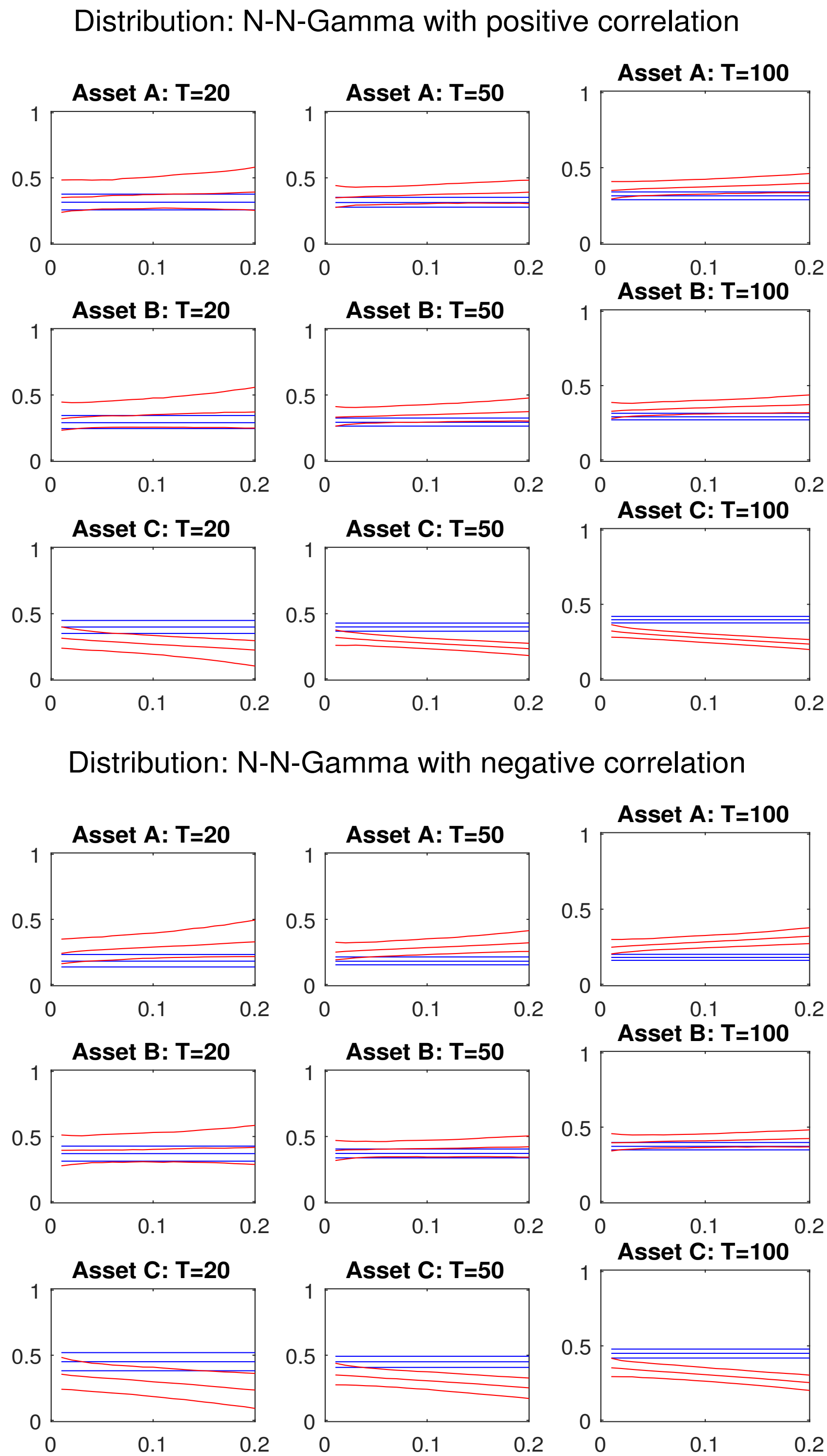


Figure 7.7: Variability of the weights of the ERC portfolios based on $EVaR_\tau$, for $\tau = 0.01, \ldots, 0.2$. Assets A and B have normal distributions, asset C has a reflected Γ distribution. The red lines indicate the 90^{th} percentile, the median and the 10^{th} percentile of the portfolio weights. The blue lines represent the same quantities for a the risk parity based on variance portfolio.

nggcorrbertzimas⟩

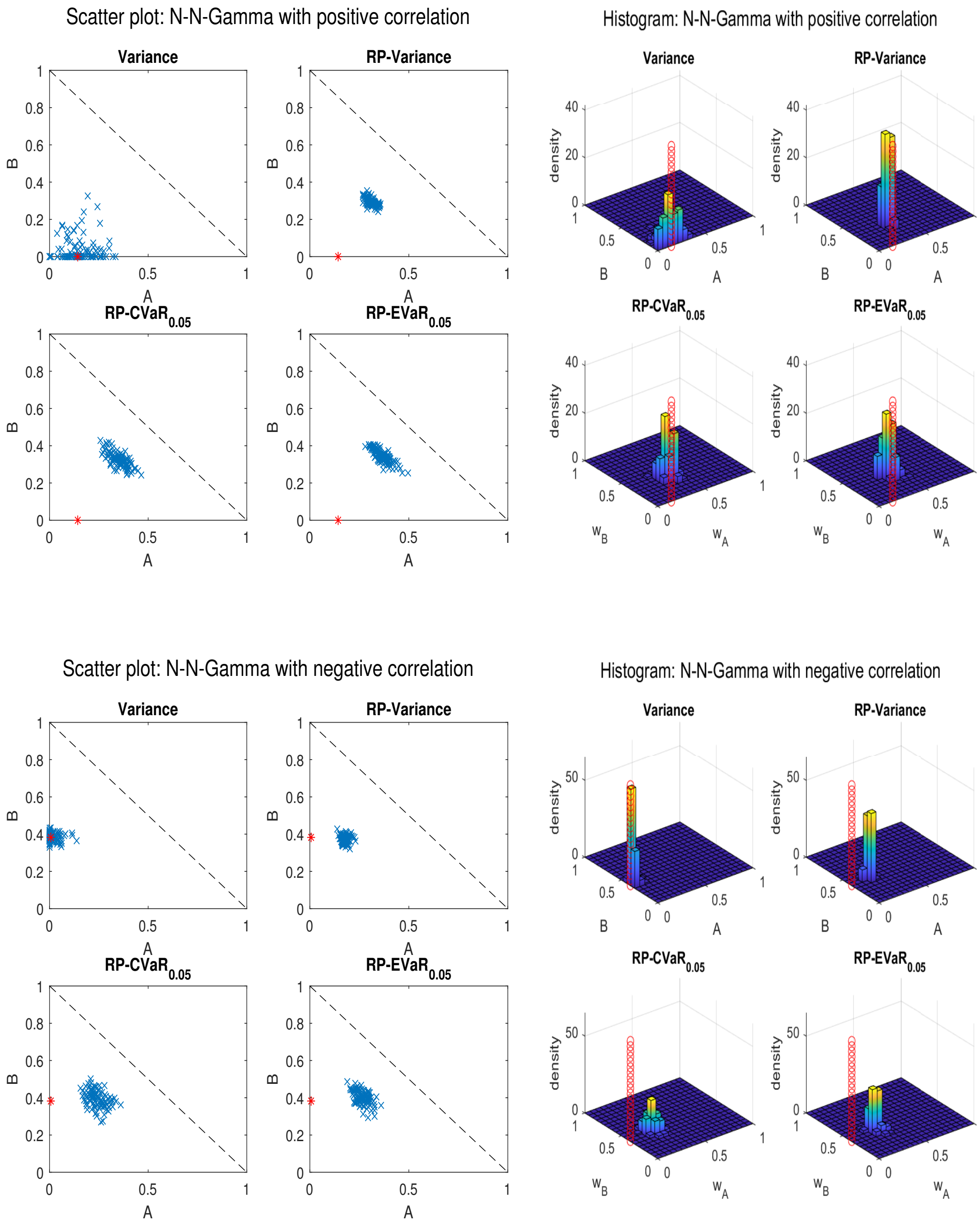


Figure 7.8: Scatter plot and histogram of the weights of the assets A and B, the weight of C is given by the relation $\sum \omega_i = 1$, the pattern is the same of Fig. 7.5 with 1000 simulations. The red points represent the theoretical value of the Markowitz portfolio weights

⟨RP_nggcorrPW⟩

7.3.2 Historical Data

In this section we perform tests on the ERC portfolios using the same dataset of section 4.4.2, i.e. using the logreturns of the companies that constitute the $S\&P$ 500 index.In our analysis we consider two different periods: January 2011 - December 2013 and January 2014 - December 2016.

As before, we make an *out-of-sample* analysis considering a rolling window of length 100 days with equal weights: every day the portfolio is rebalanced following the risk parity approach based on variance, $CVaR_{0.2}$, $EVaR_{0.05}$. Then empirical distribution of the portfolios' logreturns is evaluated and compared to the equally weighted portfolio. Similar experiments are performed for a portfolio composed of 250 assets of the $S\&P$ 500.

In Fig. 7.9 and 7.10 the empirical distribution of the out-of-sample portfolios' logreturns is represented, and in the Tab. 7.1 and 7.2 the extrema and the moments of the distributions are evaluated.
The portfolio is composed again of the first five assets of $S\&P$ 500 ordered by their weight in the index (Apple Inc., Microsoft Corporation, Exxon Mobil Corporation, Amazon.com Inc. and JPMorgan Chase & Co).
Each day we compute the realized logreturn of the ERC portfolios evaluated the day before.

As we expect from theory, since the distribution of the assets is almost symmetric, in all cases the methods are comparable and there are not substantial differences on the choice of the risk measure to use in the equally-weighted risk contribution portfolio.
Let now consider a portfolio composed the first 250 assets of $S\&P$ 500 ordered by their weight in the index. In Fig. 7.11 and 7.12 the empirical distribution of the out-of-sample portfolios' logreturns are represented and in Tab. 7.3 and 7.4, the extrema and the moments of the distributions are evaluated.

Considering the moments of the distributions of the logreturns of large portfolios we can observe there are not substantial differences in the performances of the portfolios. Again, when considering symmetric distributions, the methods are comparable and there are not substantial differences on the choice of the risk measure to use in the equally-weighted risk contribution portfolio.
From a computational point of view, the algorithm proposed to evaluate the ERC portfolio based on EVaR_τ works also with large portfolios.

1^{st} January 2011 - 31^{th} December 2013

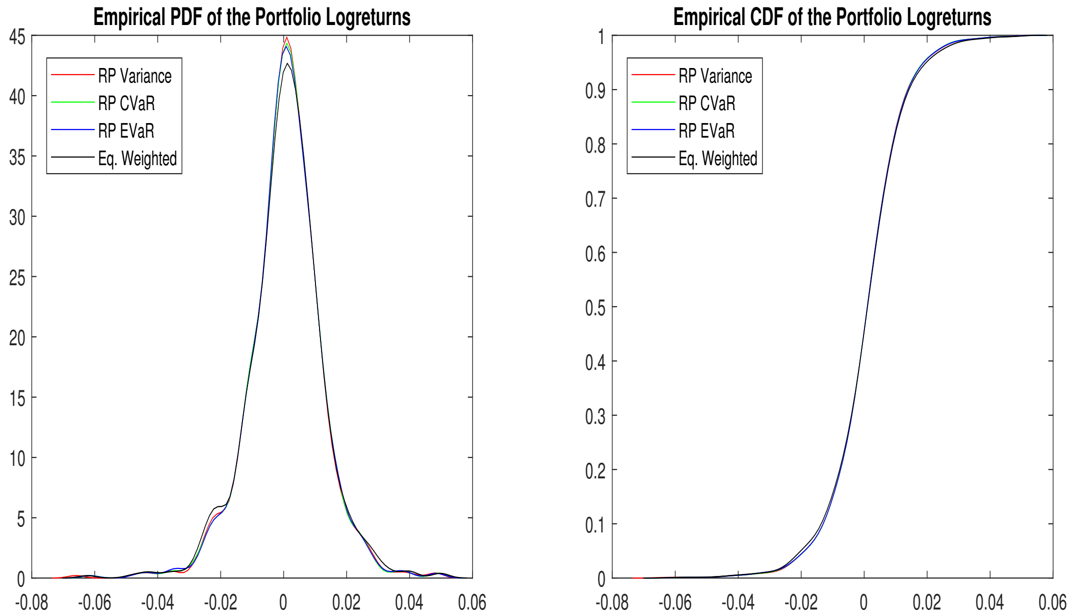


Figure 7.9: Empirical distribution of the 5-asset-portfolio logreturns computed out-of-sample with the rolling window approach over the period Jan11-Dec13: the probability and cumulative density function are smoothed by normal kernels.
⟨3511-13⟩

	RP Variance	**RP CVaR**	**RP EVaR**	**Eq. Weighted**
Min	-6.59e-02	-6.14e-02	-6.16e-02	-6.24e-02
Max	4.82e-02	4.94e-02	4.92e-02	5.04e-02
Mean	6.71e-04	6.80e-04	6.64e-04	6.78e-04
Std	1.14e-02	1.14e-02	1.15e-02	1.19e-02
Skew	-3.29e-01	-2.90e-01	-2.93e-01	-2.49e-01
Kurt	6.19	5.90	5.86	5.68

Table 7.1: Statistics of the 5-asset portfolio logreturn distributions over the period January 2011 - December 2013
⟨tablestat31⟩

1^{st} January 2014 - 31^{th} December 2016

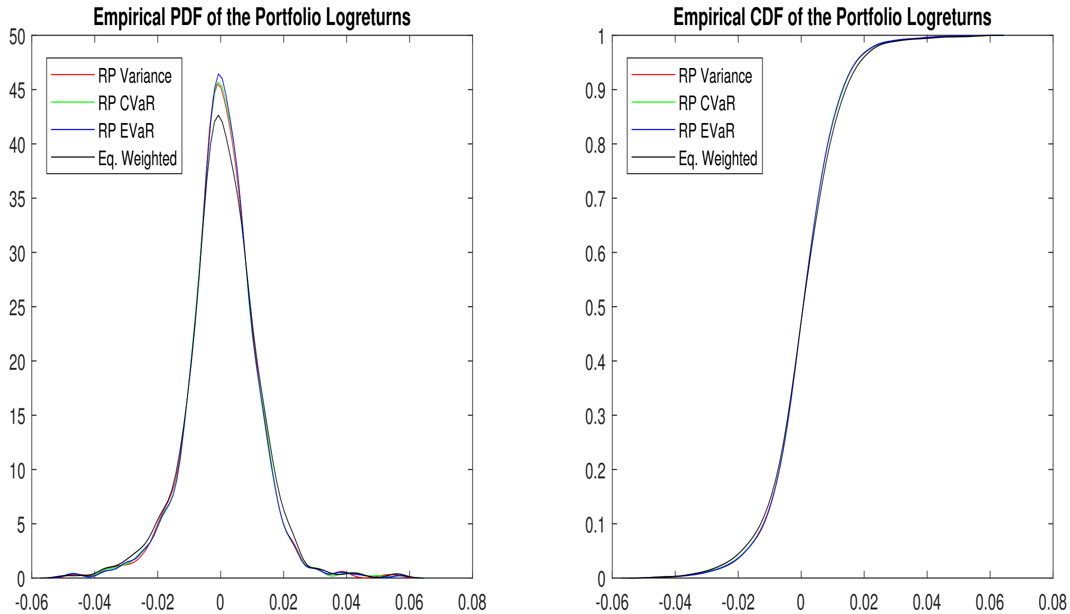


Figure 7.10: Empirical distribution of the 5-asset-portfolio logreturns computed out-of-sample with the rolling window approach over the period Jan14-Dec16: the probability and cumulative density function are smoothed by normal kernels.
⟨3514-16⟩

	RP Variance	**RP CVaR**	**RP EVaR**	**Eq. Weighted**
Min	-4.76e-02	-4.79e-02	-4.76e-02	-4.95e-02
Max	5.59e-02	5.73e-02	5.70e-02	5.67e-02
Mean	5.39e-04	5.64e-04	5.51e-04	6.28e-04
Std	1.08e-02	1.09e-02	1.07e-02	1.15e-02
Skew	-9.21e-03	-1.54e-02	-4.47e-02	-3.00e-03
Kurt	5.99	6.18	6.04	5.71

Table 7.2: Statistics of the 5-asset portfolio logreturn distributions over the period January 2011 - December 2013
⟨tablestat32⟩

1^{st} January 2011 - 31^{th} December 2013

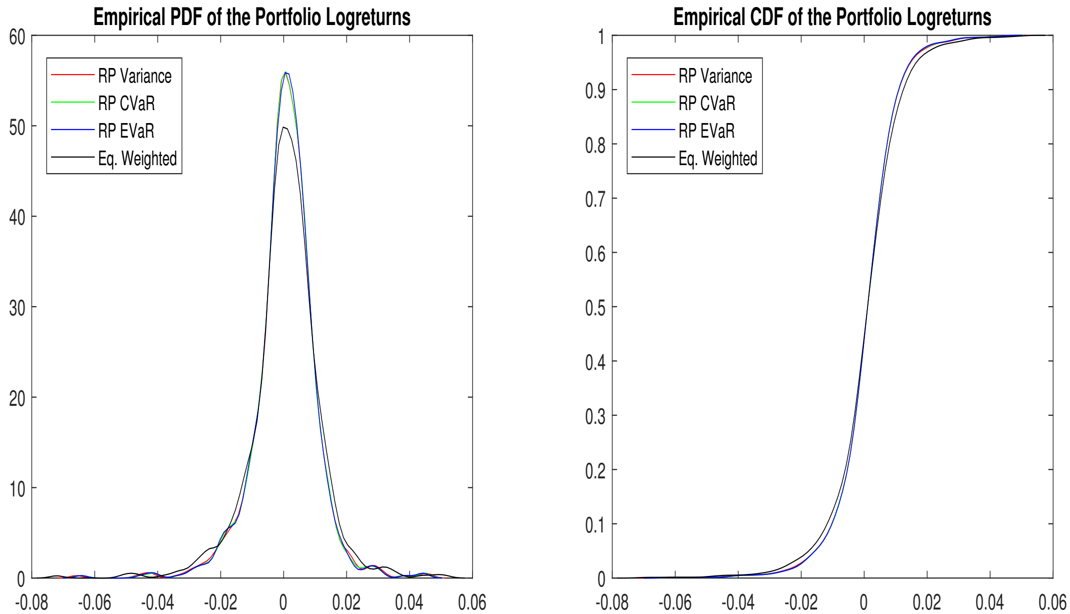


Figure 7.11: Empirical distribution of the 250-asset-portfolio logreturns computed out-of-sample with the rolling window approach over the period Jan11-Dec13: the probability and cumulative density function are smoothed by normal kernels.
⟨325011-13⟩

	RP Variance	RP CVaR	RP EVaR	Eq. Weighted
Min	-6.58e-02	-6.42e-02	-6.38e-02	-7.20e-02
Max	4.56e-02	4.48e-02	4.43e-02	5.10e-02
Mean	6.97e-04	6.86e-04	6.67e-04	6.83e-04
Std	9.76e-03	9.61e-03	9.57e-03	1.10e-02
Skew	-6.31e-01	-6.15e-01	-6.06e-01	-5.83e-01
Kurt	8.86	8.69	8.57	8.41

Table 7.3: Statistics of the 250-asset portfolio logreturn distributions over the period January 2011 - December 2013
⟨tablestat33⟩

1^{st} January 2014 - 31^{th} December 2016

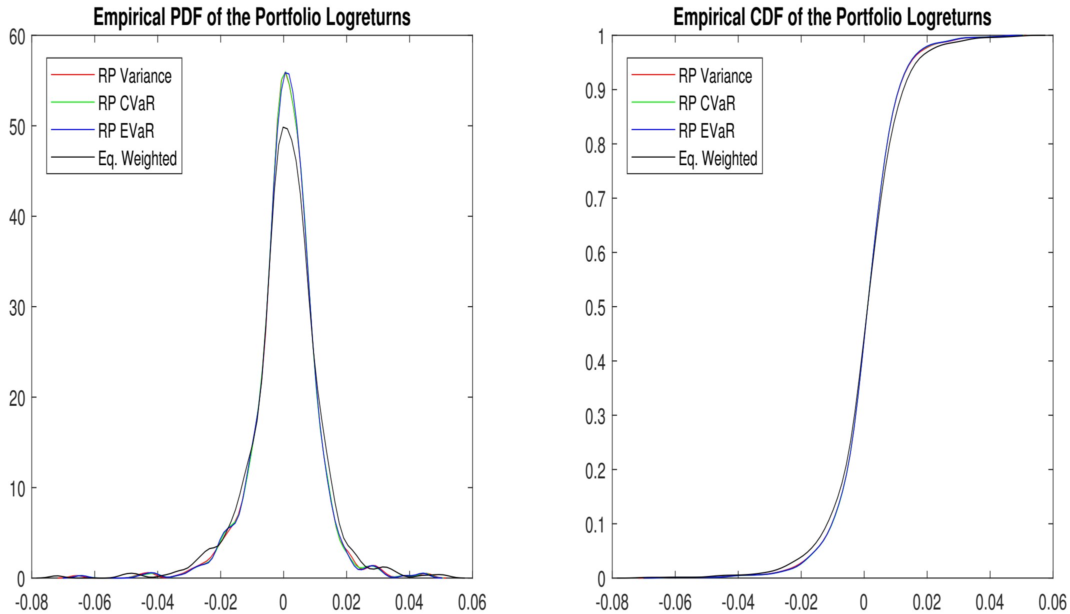


Figure 7.12: Empirical distribution of the 250-asset-portfolio logreturns computed out-of-sample with the rolling window approach over the period Jan14-Dec16: the probability and cumulative density function are smoothed by normal kernels.
⟨325014-16⟩

	RP Variance	RP CVaR	RP EVaR	Eq. Weighted
Min	-4.22e-02	-4.21e-02	-4.21e-02	-4.21e-02
Max	3.36e-02	3.38e-02	3.38e-02	3.50e-02
Mean	3.81e-04	3.83e-04	3.78e-04	4.11e-04
Std	8.09e-03	7.98e-03	7.99e-03	8.81e-03
Skew	-3.93e-01	-4.16e-01	-4.58e-01	-4.35e-01
Kurt	5.18	5.19	5.22	5.13

Table 7.4: Statistics of the 250-asset portfolio logreturn distributions over the period January 2011 - December 2013
⟨tablestat34⟩

Chapter 8

Conclusions

We introduced the mean-risk problem using EVaR and interexpectile differences, and we found that they can be written as linear programming problems. Moreover, considering the theories on portfolio selection based on worst-case scenarios and risk parity, we studied the corresponding EVaR formulations.

Performing tests on simulated data we evinced what we expected from theory: EVaR_τ portfolio optimization gives analogous results as classical risk measures when considering symmetric distributions. On the contrary, considering asymmetric distributions, EVaR_τ portfolio optimization takes into account the whole distribution preferring the one with fatter right tail.

In Tab. 8.1 we summarize all the experiments done with historical data and compare the methods used. We can observe that considering small portfolios, we have not substantial differences using the various methods: the portfolio found minimizing the iqr seems to be the one which performs better.

On the contrary, if we consider large portfolios, the mean-EVaR portfolio optimization seems to be preferable to the mean-CVaR portfolio optimization and the Markowitz portfolio: using EVaR as risk measure, the mean of the distribution of portfolio logreturns is greater and the variance is lower comparing with CVaR and Markowitz.

The equally-weighted risk contributions portfolios have not evident differences either considering the outcomes and the variances.

When considering the case of WEVaR optimization, we run into computational problems for a large investment universe, in fact we found that the efficient frontier that theoretically should be convex due to the coherency of WEVaR, in practice it is not. In future researches we are going to find if there is the possibility to reformulate the problem in a simpler way in order to avoid computational problems.

The portfolio which minimize $iqr_{0.25}$ is again the one which perform best, but on the other hand it has also the largest variance. In literature, this risk measure has not been discussed in detail as VaR or CVaR: considering our experiments on $iqr_{0.25}$, we think that this risk measure, or more in general the interquantile difference, deserve more attention.

		5-asset portfolio		250-asset portfolio	
		mean	std	mean	std
Markowitz	2011-2013	6.32e-04	1.21e-02	4.64e-04	9.53e-03
	2014-2016	5.55e-04	1.19e-02	2.24e-04	8.65e-03
Eq. Weightd	2011-2013	6.78e-04	1.19e-02	6.83e-04	1.10e-02
	2014-2016	6.28e-04	1.15e-02	4.11e-04	8.81e-03
$CVaR_{0.2}$	2011-2013	6.20e-04	1.24e-02	4.39e-04	1.26e-02
	2014-2016	5.60e-04	1.35e-02	3.62e-04	1.15e-02
$EVaR_{0.05}$	2011-2013	6.03e-04	1.25e-02	5.64e-04	9.69e-03
	2014-2016	6.09e-04	1.21e-02	4.06e-04	9.05e-03
$WEVaR_{0.05}$	2011-2013	5.85e-04	1.15e-02	3.78e-04	1.32e-02
	2014-2016	6.05e-04	1.20e-02	7.39e-04	1.05e-02
RP Variance	2011-2013	6.71e-04	1.14e-02	6.97e-04	9.7e-03
	2014-2016	5.39e-04	1.08e-02	3.81e-04	8.094e-03
RP $CVaR_{0.2}$	2011-2013	6.80e-04	1.14e-02	6.86e-04	9.61e-03
	2014-2016	5.64e-04	1.09e-02	3.83e-04	7.98e-03
RP $EVaR_{0.05}$	2011-2013	6.64e-04	1.15e-02	6.67e-04	9.57e-03
	2014-2016	5.51e-04	1.07e-02	3.78e-04	7.99e-03
$\Delta_{0.25}$	2011-2013	6.62e-04	1.13e-02	6.07e-04	9.45e-03
	2014-2016	5.01e-04	1.20e-02	2.61e-04	8.79e-03
$iqr_{0.25}$	2011-2013	7.50e-04	1.25e-02	8.98e-04	1.53e-02
	2014-2016	7.52e-04	1.26e-02	5.09e-04	1.23e-02

Table 8.1: Statistics the logreturns of the portfolios evaluated in the different methods

⟨tableconclusion⟩

List of Figures

List of Tables

Bibliography

Acerbi2001 [1] Acerbi, C. and D. Tasche (2002a). Expected shortfall: A natural coherent alternative to value at risk. *Economic Notes 31*(2), 379–388.

Acerbi2002 [2] Acerbi, C. and D. Tasche (2002b). On the coherence of expected shortfall. *Journal of Banking & Finance 26*(7), 1487–1503.

ner99Coherent [3] Artzner, P., F. Delbaen, J.-M. Eber, and D. Heath (1999). Coherent measures of risk. *Mathematical Finance 9*(3), 203–228.

12fundamental [4] Basel Committee on Banking Supervision (2012). Fundamental review of the trading book. Bank of International Settlements.

015elicitable [5] Bellini, F. and V. Bignozzi (2015). On elicitable risk measures. *Quantitative Finance 15*(5), 725–733.

DB15 [6] Bellini, F. and E. Di Bernardino (2017). Risk management with expectiles. *The European Journal of Finance 23*(6), 487–506.

14generalized [7] Bellini, F., B. Klar, A. Müller, and E. Rosazza Gianin (2014). Generalized quantiles as risk measures. *Insurance: Mathematics and Economics 54*, 41–48.

BMR2017 [8] Bellini, F., L. Mercuri, and E. Rroji (2018). Implicit expectiles and measures of implied volatility. *Quantitative Finance*, 1–14.

BNP2017 [9] Bellini, F., I. Negri, and M. Pyatkova (2017). Backtesting var and expectiles with realized scores. *Forthcoming in Statistical Methods and Applications*.

robust5 [10] Ben-Tal, A., T. Margalit, and A. Nemirovski (2000). *High Performance Optimization*, Chapter Robust Modeling of Multi-Stage Portfolio Problems, pp. 303–328. Springer US.

ertsimas2011 [11] Bertsimas, D., D. B. Brown, and C. Caramanis (2011). Theory and applications of robust optimization. *SIAM Rev. 53*(3), 464–501.

ERTSIMAS2004 [12] Bertsimas, D., G. J. Lauprete, and A. Samarov (2004). Shortfall as a risk measure: properties, optimization and applications. *Journal of Economic Dynamics and Control 28*(7), 1353 – 1381.

Fenchel [13] Borwein, J. and A. Lewis (2006). *Convex Analysis and Nonlinear Optimization: Theory and Examples (2 ed.)*, pp. 76–77. Springer.

reckling1988 [14] Breckling, J. and R. Chambers (1988). M-quantiles. *Biometrika 75*(4), 761–771.

Cesarone2017 [15] Cesarone, F. and S. Colucci (2017). Minimum risk versus capital and risk diversification strategies for portfolio construction. *Journal of the Operational Research Society*.

CST13 [16] Cesarone, F., A. Scozzari, and F. Tardella (2013). A new method for mean-variance portfolio optimization with cardinality constraints. *Annals OR 205*(1), 213–234.

CT2017 [17] Cesarone, F. and F. Tardella (2017). Equal risk bounding is better than risk parity for portfolio selection. *J. Global Optimization 68*(2), 439–461.

CHANG00 [18] Chang, T.-J., N. Meade, J. Beasley, and Y. Sharaiha (2000). Heuristics for cardinality constrained portfolio optimisation. *Computers & Operations Research 27*(13), 1271 – 1302.

robust3 [19] Costa, O. L. V. and A. C. Paiva (2002). Robust portfolio selection using linear-matrix inequalities. *Journal of Economic Dynamics and Control 26*(6), 889–909.

Delbaen13 [20] Delbaen, F. (2013). A Remark on the Structure of Expectiles. Technical report, arXiv.org, https://ideas.repec.org/p/arx/papers/1307.5881.html.

baen2015risk [21] Delbaen, F., F. Bellini, V. Bignozzi, and J. F. Ziegel (2015). Risk measures with the cxls property. *Finance and Stochastics*, 1–21.

ElG2003 [22] El Ghaoui, L., M. Oks, and F. Oustry (2003). Worst-case value-at-risk and robust portfolio optimization: A conic programming approach. *Operations Research 51*(4), 543–556.

Fan1953 [23] Fan, K. (1953). Minimax theorems. *Proceedings of the National Academy of Sciences of the United States of America 39*(1), 42–47.

04stochastic [24] Föllmer, H. and A. Schied (2004). *Stochastic Finance: An Introduction in Discrete Time*. De Gruyter studies in mathematics. Walter de Gruyter.

r08convexand [25] Föllmer, H. and A. Schied (2010, 01). Convex and coherent risk measures. *Encyclopedia of Quantitative Finance*.

FW2015 [26] Föllmer, H. and S. Weber (2015). The axiomatic approach to risk measures for capital determination. *Annual Review of Financial Economics 7*(1), 301–337.

Gneiting2011 [27] Gneiting, T. (2011). Making and evaluating point forecasts. *Journal of the American Statistical Association 106*(494), 746–762.

Gold2003 [28] Goldfarb, D. and G. Iyengar (2003). Robust portfolio selection problems. *Mathematics of Operations Research 28*(1), 1–38.

Hall2003 [29] Halldórsson, B. and R. Tütüncü (2003). An interior-point method for a class of saddle-point problems. *Journal of Optimization Theory and Applications 116*(3), 559–590.

jak16 [30] Jakobsons, E. (2016a). Scenario aggregation method for portfolio expectile optimization. *Statistics & Risk Modeling 33*(1-2), 51–65.

jak15 [31] Jakobsons, E. (2016b). Suboptimality in portfolio conditional value-at-risk optimization. *Journal of Risk 18*(4), 1–23.

Kempf2006 [32] Kempf, A. and C. Memmel (2006). Estimating the global minimum variance portfolio. *Schmalenbach Business Review (sbr) 58*(4), 332–348.

Konno1991 [33] Konno, H. and H. Yamazaki (1991). Mean-absolute deviation portfolio optimization model and its applications to tokyo stock market. *Manage. Sci. 37*(5), 519–531.

Leung2011 [34] Lai, T. L., H. Xing, and Z. Chen (2011, August). Mean–variance portfolio optimization when means and covariances are unknown. Technical Report 1108.0996, arXiv.org, https://ideas.repec.org/p/arx/papers/1108.0996.html.

Imre2016 [35] Ledoit, O. and M. Wolf (2003). Improved estimation of the covariance matrix of stock returns with an application to portfolio selection. *Journal of Empirical Finance 10*(5), 603 – 621.

LinsVaR [36] Linsmeier, T. J. and N. D. Pearson (2000). Value at risk. *Financial Analysts Journal 56*(2), 47–67.

robust2 [37] Lobo, M. S. and S. Boyd (2000). The worst-case risk of a portfolio. *Technical Report, Duke University*.

Maillard2010 [38] Maillard, S., T. Roncalli, and J. Teiletche (2010). The properties of equally weigthed risk contributions portfolios. *The Journal of Portfolio Management 36*(4), 60–70.

Mansini2003a [39] Mansini, R., W. Ogryczak, and M. G. Speranza (2003a). LP solvable models for portfolio optimization: a classification and computational comparison. *IMA Journal of Management Mathematics 14*(3), 187–220.

Mansini2003b [40] Mansini, R., W. Ogryczak, and M. G. Speranza (2003b). On LP solvable models for portfolio selection. *Informatica, Lith. Acad. Sci. 14*(1), 37–62.

ni2014twenty [41] Mansini, R., W. Ogryczak, and M. G. Speranza (2014). Twenty years of linear programming based portfolio optimization. *European Journal of Operational Research 234*(2), 518–535.

arkowitz1952 [42] Markowitz, H. M. (1952). Portfolio selection. *The Journal of Finance 7*(1), 77–91.

eyPowell1987 [43] Newey, W. K. and J. L. Powell (1987). Asymmetric least squares estimation and testing. *Econometrica 55*(4), 819–847.

Ogryczak1999 [44] Ogryczak, W. and A. Ruszczynski (1999). From stochastic dominance to mean-risk models: Semideviations as risk measures. *European Journal of Operational Research 116*(1), 33–50.

k2011solving [45] Ogryczak, W. and T. Śliwiński (2011). On solving the dual for portfolio selection by optimizing conditional value at risk. *Computational Optimization and Applications 50*(3), 591–595.

flug2004risk [46] Pflug, G. and A. Ruszczynski (2004). A risk measure for income processes. In G. Szegö (Ed.), *Risk Measures for the 21st Century*, pp. 249–269. J. Wiley and Sons.

Pflug2000 [47] Pflug, G. C. (2000). *Probabilistic Constrained Optimization: Methodology and Applications*, Chapter Some Remarks on the Value-at-Risk and the Conditional Value-at-Risk, pp. 272–281. Springer US.

kafellarConj [48] Rockafellar, R. T. (1974). Conjugate duality and optimization. *Regional Conference Series in Applied Mathematics*.

timizationof [49] Rockafellar, R. T. and S. Uryasev (2000). Optimization of conditional value-at-risk. *Journal of Risk 2*, 21–41.

kafellar2006 [50] Rockafellar, R. T., S. Uryasev, and M. Zabarankin (2006). Generalized deviations in risk analysis. *Finance and Stochastics 10*(1), 51–74.

ilitypremium [51] Rolski, T., H. Schmidli, V. Schmidt, and J. Teugels (2008). *Stochastic Processes for Insurance & Finance*, Chapter Premiums and Ordering of Risks, pp. 79–98. John Wiley and Sons, Inc.

optimization [52] Ruszczynski, A. and A. Shapiro (2006). Optimization of convex risk functions. *Mathematics of Operations Research 31*(3), 433–452.

Shapiro1991 [53] Shapiro, A. (1991). Asymptotic analysis of stochastic programs. *Annals of Operations Research 30*(1), 169–186.

fanovits2010 [54] Stefanovits, D. (2010). Equal contributions to risk and portfolio construction. Master's thesis, ETH, Zurich.

Tutuncu2004 [55] Tütüncü, R. and M. Koenig (2004). Robust asset allocation. *Annals of Operations Research 132*(1), 157–187.

Yitzhaki1982 [56] Yitzhaki, S. (1982). Stochastic dominance, mean variance, and Gini's mean difference. *American Economic Review 72*(1), 178–85.

Zhu09 [57] Zhu, S. and M. Fukushima (2009). Worst-case conditional value-at-risk with application to robust portfolio management. *Operations Research 57*(5), 1155–1168.

Ziegel2013 [58] Ziegel, J. F. (2016). Coherence and elicitability. *Mathematical Finance 26*(4), 901–918.