University of Pisa

Ph.D. Program in Mathematics for Economic Decisions

Leonardo Fibonacci School

Ph.D. Thesis

Robust asset allocation problems: a new class of risk measure based models

SSD: MAT/09

Raffaella Recchia

Ph.D. Supervisor:

Prof. Maria Grazia Scutellà

"Nelle Tue mani è la mia vita."
(Sal.15,5)

Acknowledgments

This dissertation would not have been possible without the help and encouragement of several people. First of all, I thank my supervisor Maria Grazia Scutellà. She has been a source of support, ideas and encouragement. I am very grateful to her because she has been an example for me.

I am greatly indebted to Professor Riccardo Cambini, his availability has been providential every times I needed.

I wish to thank the referees Professors Reha Tütüncü and David B. Brown for their helpful comments and suggestions; as well as I thank Professor Tütüncü and Professor Byrne who kindly provided me with used data for the computational analysis.

I heartily thank my colleague and friend Francesca Salvi; we spent a lot of time together, overlapping troubles and sharing joys. She has often been my source of inspiration and my model. Affectionately, I want to thank Fausto and Lorenzo for their advice, their help and their liking, as well as Roberta, Rossana and Mauro; finally all my colleagues of the Department of Computer Sciences where I worked during these years and those one of the Department of Statistics and Applied Mathematics.

A sincere thanks to my family: they have never once failed to support me when I needed. I am grateful to my mother, her love has been my greatest strength and it has been of use to go on in the difficult situations; I thank my father for believing in me all times and my brother for his ever present good humor and for the technical (and not only) support. I am grateful to Aldo who supported me unconditionally throughout these years with patience and affect; and I thank Anna, she has always cheered up me even if she was far.

Finally, I thank all friends and people that have touched my life in special ways.

Ringraziamenti

Questo lavoro non sarebbe stato possibile senza l'aiuto e il sostegno di diverse persone. Prima di tutto, vorrei ringraziare il mio supervisore, la Professoressa Maria Grazia Scutellà per il supporto continuo, l'attenzione, le idee e l'incoraggiamento che non hai mai smesso di darmi.

Un ringraziamento particolare al Professor Riccardo Cambini per essersi messo a mia disposizione in ogni occasione a volte quasi in maniera provvidenziale.

Sono grata ai Professori Tütüncü e Brown per aver accettato di revisionare la mia tesi e per avermi offerto utili commenti e suggerimenti; il Professor Tütüncü insieme al Professor Byrne mi hanno poi gentilmente fornito i dati attraverso i quali è stato possibile portare avanti l'analisi computazionale e per questo li ringrazio.

Di cuore, vorrei poi ringraziare la mia collega e amica Francesca Salvi; insieme abbiamo vissuto tanti momenti di difficoltà e condiviso tante gioie. La ringrazio per essere stata spesso la mia fonte di ispirazione e il mio modello e soprattutto per aver sempre trovato le parole giuste per risollevarmi. Con grande affetto vorrei ringraziare Fausto e Lorenzo per i loro consigli, il loro aiuto e la loro simpatia, così come Roberta, Rossana e Mauro e tutti i colleghi del Dipartimento di Informatica dove ho lavorato tutti questi anni e quelli del Dipartimento di Statistica e Matematica Applicata.

Un grazie particolare alla mia famiglia che non ha mai ostacolato le mie scelte, mi ha sempre appoggiato ed è stata continuamente presente. A mia madre, il suo amore è stato la mia più grande forza e mi ha fatto andare avanti anche nei momenti più critici; a mio padre che ha sempre creduto in me e mi ha incoraggiato e a mio fratello per il suo humor e il suo supporto tecnico (e non solo).

Un ringraziamento speciale ad Aldo, per avermi sostenuto incondizionatamente in tutti questi anni con pazienza e affetto, per aver dovuto subire tutti i miei sbalzi di umore e per non aver mai perso fiducia in me.

Con tanta stima e affetto vorrei poi ringraziare Anna, guida costante e discreta che ha continuato a regalarmi serenità, sostegno e attenzione anche quando lontana.

Infine, un grazie a tutte quelle persone che, in questi anni, hanno toccato la mia vita in maniera unica e speciale.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Optimization under uncertainty

Optimization problems very often involve input data and parameters that are uncertain due to measurement or modeling errors, or simply because they correspond to quantities that will only be realized in the future, or cannot be known exactly at the time the problem must be formulated and solved. Over the years, the interest about this kind of problems has increased, and a certain number of methodologies have been developed for their solution.

The oldest such method, the "sensitivity analysis" deals with uncertainty after an optimal solution is obtained. The aim is to determine the parameter ranges over which the current solution remains optimal, assuming that one or more parameters can deviate from their nominal value.

Other methods incorporate the uncertainty directly into the computation of the optimal solution.

The *Stochastic Programming* approach for example utilizes an underlying probability model to handle the uncertainty. This approach has been successfully applied in a different number of areas, but it remains challenging to implement it because probability distributions are often unknown in practice and also because of the intractability of the models when the problem size grows.

Robust Optimization is a relatively recent framework in which the uncertainty is treated as *deterministic*. Unlike the traditional approach, robust optimization incorporates the notion that inputs have been estimated with errors. In this case, the inputs are not the traditional forecasts, but rather uncertainty sets including these point estimates. The optimal solution to the optimization problem is required to remain feasible for any values of the parameters that fall within these specified uncertainty sets.

In this work we will focus on the Robust Optimization approach in the context of asset allocation problems, considering however an extension of the standard notion of robustness that involves also some probability aspects.

1.2 Motivation and overview

Many problems, particularly in financial field, involve necessarily uncertain parameters, like future values of security prices, interest rates and exchange rates. This is true, for example, in asset allocation problems. Our focus is to handle the uncertainty in this kind of problems using a flexible robust approach [5], in which robustness incorporates some probabilistic aspects.

The remainder is organized as follows. Chapter 2 provides an overview of the theory and the applications of robust optimization in the field of robust asset allocation problems. Starting from the classical *mean-variance* optimization problems introduced by Markowitz, we review several robust models proposed in literature, from the most traditional one to the most recent developments. We also give an overview of some of the algorithmic approaches and relative computational results.

By emphasizing the role played by convex risk measures in this environment, we then describe an innovative approach to robustness which relaxes the traditional notion of robustness and that specifies not only the values of the uncertainty parameters, but also their degree of feasibility.

Based on the relaxed robustness related to convex risk measures, Chapter 3 proposes a new family of models that we call *norm-portfolio* models, that include as special cases linear programming (LP) and second order cone programming (SOCP) problems, i.e., computational tractable models. To define these models we focus on the notion of the penalty function appearing in the characterization of convex risk measures, and we propose models in which this function is defined in terms of the general norms in order to obtain models which are computationally tractable.

Then we study a variant of the proposed family, i.e., the case in which the used risk measure is a coherent one (a sub-case of the convex one) and we conclude the chapter with considerations about some used parameters, that describe an interesting link between this coherent variant of the *norm-portfolio* models and one of the most known coherent risk measures studied in literature, i.e., the Conditional Value at Risk (CVaR).

Chapter 4 provides the implementation of some described models, with real market data. The aim of the computational analysis is to observe how different risk measures utilize the scenario information based on past history in producing a successful portfolio. We work on three different data sets and for each of them we lead a twofold analysis: an *in-sample* analysis through that we determine suitable values of the parameters in the models, and an *out-of-sample* analysis in which we present in a certain sense the "actual" performance of the models.

Finally, in order to conduct a preliminary comparison among several robustness types, we compare some *norm-portfolio* models and their *coherent* variant with the following models described in literature: the classical CVaR, some standard robust models and a soft robust model. Our aim is to understand how the greater flexibility influences the optimal portfolio values. To compare the models we use the following statistics: the out-of-sample mean, variance, Sharpe Ratio of realized return, portfolio turnover and the computational cost (i.e., the time needed to solve each model).

In general, from the computational analysis the following key observations are drawn: in the first experiment between a traditional robustness and a more relaxed robust approach, the second type always outperforms the first one. In other words, by relaxing the robustness constraints in a flexible way, one can potentially gain out-of-sample performance for not too high of a price. Among the flexible type robustness, the soft one (incorporated by a so-called *entropic* model) is without doubts the best approach in terms of almost all criteria adopted in order to compare the models, even if it results the most expensive in terms of the transaction costs, i.e., by evaluating the variability of assets into portfolio and above all in terms of the computational cost.

In the second computational test, the standard robustness shows a better performance than relaxed one in terms of variability, risk-return relation and portfolio turnover. The standard approach results more robust than relaxed one, although soft robustness incorporated by the *entropic* model produces by far the best performance.

In the last computational test the related robustness incorporated by the *norm-portfolio* approach and the soft one incorporated by the *entropic* model provide a better performance with respect to the traditional approach. In particular, traditional robustness produce the worst performance at the highest cost in terms of portfolio turnover.

Finally, Chapter 5 presents some final comments and directions for future research.

Chapter 2

Robust asset allocation problems: a literature review

2.1 Asset Allocation Problems

Portfolio selection problems were formulated for the first time by Markowitz in 1952. They consist in allocating capital over a number of available assets in order to maximize the "return" on the investment while minimizing the "risk" using mathematical techniques. In the proposed models, the return is measured by the expected value of the random portfolio return, while the risk is quantified by the variance of the portfolio (*mean-variance models*).

Despite the strong theoretical support proved by mean-variance models, their elegance and the availability of efficient computer codes for solving them, these models present various practical pitfalls: optimal portfolios are not well diversified, in fact they tend to concentrate on a small subset of the available securities and, above all, they are often very sensitive to changes in input parameters.

2.1.1 Formulations of asset allocation problems

Let n be the number of available assets and X be the set of feasible portfolios defined as

$$X = \left\{ x \in \mathbb{R}^n | \sum_{j=1}^{n} x_i = 1,\ x_j \geq 0, j = 1, \cdots, n \right\} \tag{2.1}$$

i.e., a non-empty and bounded set where no short sales are allowed.

The set X in (2.1) is only an example of feasible portfolio set, additional constraints in fact, could be added to describe it.

Then, the classical Markowitz' *mean-variance optimization (MVO) models* can be formulated as follows:

1) maximize the expected return subject to an upper limit on the variance:

$$\begin{aligned} \max \quad & \mu^T x \\ s.t. \quad & x^T Q x \leq \sigma \\ & x \in X; \end{aligned} \tag{2.2}$$

2) minimize the variance subject to a lower limit on the expected return:

$$\begin{aligned} \min \quad & x^T Q x \\ s.t. \quad & \mu^T x \geq R \\ & x \in X; \end{aligned} \tag{2.3}$$

3) maximize the *risk-adjusted expected return*:

$$\begin{aligned} \max \quad & \mu^T x - \lambda x^T Q x \\ s.t. \quad & x \in X \end{aligned} \tag{2.4}$$

where μ and Q denote the estimated expected return vector and the covariance matrix of the given assets respectively and $\lambda \in \mathbb{R}$ indicates a risk-aversion parameter.

These three models are parametrized by the variance limit, the expected return limit and the risk-aversion parameter, respectively. Let us observe that while the first formulation can not be classified as convex QP problem because of nonlinear variance constraint, the latter two are convex QP problems [53, 79].

Mean-variance portfolios generated using the sample expected return and covariance matrix of the asset returns perform poorly out of sample due to estimation errors [17, 20, 25, 56].

A study of Black and Litterman [17] for example demonstrated that small changes in the expected returns, in particular, had a substantial impact in the portfolio composition. It is indeed commonly accepted that estimation errors in the sample expected return are much larger than in sample covariance matrix [24, 45]. For this reason, researchers have recently focused on the so-called *minimum-variance optimization models*:

$$\begin{aligned} \min \quad & x^T Q x \\ s.t. \quad & x \in X, \end{aligned}$$

which rely solely on estimates of the covariance matrix.

However, also these portfolios are quite sensitive to estimation errors and have unstable weigths that fluctuate substantially over time. The main motivation is that the sample covariance matrices, on which minimum-variance portfolios are based on, are the maximum likelihood estimators (MLE) for normally distributed returns, and the efficiency of these estimators is highly sensitive to deviations of the asset-return distribution from the implicitly assumed normal distribution. This is particular relevant for portfolio asset allocation, where extensive evidence shows that the empirical distribution of the asset returns usually deviates from the normal distribution, [28].

For this, incorporating the uncertainty about the accuracy of the estimates in the portfolio optimization process becomes crucial for practical applications.

A way to handle the uncertainty in minimum-variance models is to use suitable robust estimators of the portfolio return characteristics, that allow one to generate portfolios with better stability properties. A *robust estimator* is one which should have good properties not only for the assumed (normal) distribution, but also for distributions in a neighborhood of the assumed one (see [44]). The class of models based on robust estimators will be not reviewed in this dissertation. The interested reader is referred to [22, 62, 55] for two-steps approach and [28, 49] for innovative one-step approach.

2.2 Robust Optimization

Robust Optimization refers to modeling of optimization problems with uncertain data to obtain a solution that is guaranteed to be "good" for all or most possible realizations of the uncertain parameters. Uncertainty in the parameters is described through uncertainty sets that contain many possible values that may be realized for the uncertain parameters. The size of the uncertainty set is determined by the level of desired robustness.

2.2.1 Absolute robustness

Let us consider a general formulation of an optimization problem:

$$\min \{c(x) \mid x \in X\} \tag{2.5}$$

where $c(x)$ is a generic objective function, $x \in \mathbb{R}^n$ is a decision vector and X represents the set of feasible solutions.

In general, uncertainty could affect either the constraints (constraint robustness) or the objective function (objective robustness); in the first case we seek solutions that are feasible for all possible values of the uncertain inputs, in the second one we seek solutions that minimize the maximum value of the objective function value considering all possible realizations of the uncertain parameters inside the uncertainty set.

Let us consider an optimization problem in which the uncertainty appears in constraints, i.e.:

$$\begin{aligned} &\min_x \; c(x) \\ &s.t. \; G(x,p) \in K \end{aligned} \tag{2.6}$$

where like before x and $c()$ are the decision vector and the objective function, respectively; G and K are the structural elements of the constraints that are assumed to be certain. Let us consider an uncertainty set $\mathcal{U}$ that contains all possible values of the uncertain parameters p, then a constraint-robust optimal solution can be found by solving the following problem:

$$\begin{aligned} &\min_x \; c(x) \\ &s.t. \; G(x,p) \in K, \quad \forall p \in \mathcal{U}. \end{aligned} \tag{2.7}$$

The vector x^* that solves (2.7) is called a *robust solution* of the uncertain problem. A *robust feasible* solution to the robust counterpart (2.7) should, by definition, satisfy all realizations of the constraints from the uncertainty set $\mathcal{U}$, and a *robust optimal* solution to (2.7) is a robust feasible solution with the best possible value of the objective function.

Now, let us observe that an optimization problem in which uncertain parameters appear in both the objective function and constraints can be easily reformulated like (2.7) simply introducing an auxiliary variable. Indeed, let us consider the following problem

$$\begin{aligned} &\min_x \; c(x,p) \\ &s.t. \; G(x,p) \in K \end{aligned} \tag{2.8}$$

in which the uncertain parameters p appear both in the objective function and in the constraints. Problem (2.8) is equivalent to the following one:

$$\begin{aligned} &\min_{t,x} \; t \\ &s.t. \; t - c(x,p) \geq 0 \\ &\qquad G(x,p) \in K \end{aligned} \tag{2.9}$$

i.e., a problem in which all uncertainties are in the constraints.

There is an alternative procedure to handle uncertainty in the case in which it affects the objective function; this procedure consists in finding solutions whose worst-case behaviour is optimized. The worst-case behaviour of a solution corresponds to the value of the objective function for the worst possible realization of the uncertain data for that particular solution. Let us consider the following optimization problem:

$$\begin{aligned} &\min_x \; c(x,p) \\ &s.t. \; x \in X. \end{aligned} \tag{2.10}$$

As before, let us denote by $\mathcal{U}$ the uncertainty set that contains all possible values of the uncertain parameter p. Then an objective robust solution can be obtained by solving:

$$\min_{x \in X} \max_{p \in \mathcal{U}} c(x,p). \tag{2.11}$$

2.2.2 Relative robustness

Measuring the worst case in absolute way as seen in the previous subsection is much conservative and it is not consistent with the risk tolerances of many decision-makers. This is the starting observation of the relative robustness approach. The idea of relative robustness is to measure the worst case in a relative manner, i.e., relative to the best possible solution under each scenario.

In other words, the relative robustness criteria, also called *minimax regret*, consists in finding a solution $x^* \in X$ such that the maximum deviation between the objective function $c(x,p)$ and the optimal value function $z^*(p)$ is minimized under each scenario, i.e., a *relative-robust* solution is a vector x that minimizes the maximum regret:

$$\min_{x \in X} \max_{p \in \mathcal{U}} c(x,p) - z^*(p) \tag{2.12}$$

where

$$z^*(p) = \min_{x \in X} c(x,p) \; .$$

While it is intuitively attractive, relative robust formulations can also be more difficult than the standard absolute-robust formulations. Indeed, since $z^*(p)$ is the optimal-value function and involves an optimization problem itself, problem (2.12) is a three-level optimization problem as opposed to the two-level problems in absolute-robust formulations. Furthermore, the optimal value function $z^*(p)$ is rarely available in analytic form and it is hard to analyze.

2.2.3 Adjustable robust optimization

Robust optimization formulation we described above assumes that uncertain parameters will not be observed until all decision variables are determined and therefore do not allow for recourse actions that may be based on realized values of some of these parameters. Indeed, for example, multi-period decision models involve uncertain parameters, some of which are revealed during the decision process. *Adjustable robust optimization* (ARO) formulations model these decision environment and allow recourse action. These models are closely related to stochastic programming formulations with recourse. Let us consider a two-stage linear optimization problem whose first-stage decision variables x_1 need to be determined now, while the second-stage decision variables x_2 can be chosen after the uncertain parameters of the problem, A_1, A_2 and b, are realized:

$$\min_{x_1,x_2} \left\{ c^T x_1 : A_1 x_1 + A_2 x_2 \leq b \right\}. \tag{2.13}$$

Let $\mathcal{U}$ denote the uncertainty set for parameters A_1, A_2 and b.

Instead of the standard constraint-robust optimization formulation in which both sets of variables must be chosen before the uncertain parameters can be observed and therefore cannot depend on these parameters, the adjustable robust optimization formulation allows to choice the second-period variables x_2 on the basis of the realized values of the uncertain parameters. As a result, the adjustable robust counterpart problem is given as follows:

$$\min_{x_1} \left\{ c^T x_1 : \forall \left(A_1, A_2, b\right) \in \mathcal{U}, \exists x_2 \equiv x_2(A_1, A_2, b) : A_1 x_1 + A_2 x_2 \leq b \right\}.$$

Soyster [75] was one of the first researchers to investigate explicit approaches to robust optimization.

Then, Ben-Tal and Nemirovski in [7, 8] as well as separately El Ghaoui, Oustry and Lebret in [32, 33] continued to investigate this innovative approach focusing also on computational issues.

Robust optimization techniques have been applied in several fields: telecommunications, control theory, network flow problems, engineer design, supply chain problems and finance.

In the whole work we will present, we will focus on models characterized by a robustness of absolute type applied in the financial optimization problems field in which future values of security prices, interest rates and exchange rates are not known in advance, but can only be forecast or estimated.

2.2.4 Robust mean-variance models

Recalling the portfolio asset allocation problems reviewed in section 2.1.1, let us assume now that the uncertain expected return vector μ and the uncertain covariance matrix Q of the asset returns belong to uncertainty sets having the following form of intervals:

$$\mathcal{U}_\mu = \left\{\mu : \mu^L \leq \mu \leq \mu^U\right\} \text{ and } \mathcal{U}_Q = \left\{Q : Q \succeq 0, Q^L \leq Q \leq Q^U\right\}, \quad (2.14)$$

where the relation $\leq$ is intended to hold true componentwise (both considering vectors and matrices), while the restriction $Q \succeq 0$ is a necessary condition so that Q is a covariance matrix, i.e., we assume that Q is a positive semidefinite matrix.

Based on the above introduced uncertainty sets, in [79] Tütüncü and Koenig have formulated some robust counterparts of problems (2.4) and (2.3) by exploiting formulations previously introduced in [41, 43]. In order to generate the extreme values of the intervals (2.14), the authors use percentiles techniques, but several other methods could be used to do this, both using the available time series and by generating estimates of parameters needed.

The first robust model looks for a feasible portfolio x such that its minimum risk-adjusted expected return, when both parameters vary in the given uncertainty sets, is the maximum one among the feasible portfolios. On the other hand, the latter robust model looks for a feasible portfolio which guarantees the lower limit R on the expected return also in the worst case, i.e., for the worst realization of parameter μ in $\mathcal{U}_\mu$, and which minimizes the variance in the worst realization of parameter Q in the uncertainty set $\mathcal{U}_Q$. The resulting robust counterparts are therefore:

$$\max_{x \in X} \left\{ \min_{\mu \in \mathcal{U}_\mu, Q \in \mathcal{U}_Q} \mu^T x - \lambda x^T Q x \right\} \quad (2.15)$$

and

$$\begin{aligned} \min \max_{Q \in \mathcal{U}_Q} \quad & x^T Q x \\ s.t. \min_{\mu \in \mathcal{U}_\mu} \quad & \mu^T x \geq R, \\ & x \in X \end{aligned} \quad (2.16)$$

Under certain simplifying assumptions, that is when Q^U is a positive semidefinite matrix, these robust problems can be reduced to pure MVO problems. In such a special case, the best asset allocation can in fact be determined by first fixing the worst-case input data in the considered uncertainty sets, that is μ^L for the uncertain mean return vector μ and Q^U for the uncertain covariance matrix Q, and then solving the resulting QP problems [79]. Without these assumptions, it is not possible to solve the robust asset allocation problems as standard QPs. In the general case, the robust counterparts (2.15) and (2.16) can be solved using a nonlinear saddle-point formulation that involves semidefinite constraints, [43].

Let $\Psi_\lambda(x, \mu, Q) = \mu^T x - \lambda x^T Q x$ be the objective function of the problem (2.15), with $x \in X$ and $(\mu, Q) \in \mathcal{U} = (\mathcal{U}_\mu \times \mathcal{U}_Q)$. For fixed $(\mu, Q) \in \mathcal{U}$ and given $\lambda \geq 0$, Ψ_λ is a concave quadratic function of x. On the other hand, fixed x and λ, the function Ψ_λ is a linear function of μ and Q. Moreover, X and $\mathcal{U}$ are non-empty and bounded sets. Then, from [43] (see Lemma 2.3), we have that the optimal values of the following pair of primal and dual problems

$$\max_{x \in X} \left\{ \min_{(\mu,Q)} \Psi_\lambda\left(x, \mu, Q\right) \right\} \quad \min_{(\mu,Q)\in\mathcal{U}} \left\{ \max_{x \in X} \Psi_\lambda(x, \mu, Q) \right\} \tag{2.17}$$

are equal, and they are obtained at a saddle-point of the function $\Psi_\lambda(x, \mu, Q)$. That is, there exists a vector $\bar{x} \in X$ and a vector-matrix pair $(\bar{\mu}, \bar{Q}) \in \mathcal{U}$ such that:

$$\Psi_\lambda(x, \bar{\mu}, \bar{Q}) \leq \Psi_\lambda(\bar{x}, \bar{\mu}, \bar{Q}) \leq \Psi_\lambda(\bar{x}, \mu, Q), \quad \forall x \in X,\ (\mu, Q) \in \mathcal{U}. \tag{2.18}$$

Moreover, $\bar{x} \in X$ and $(\bar{\mu}, \bar{Q}) \in \mathcal{U}$ solve both problems in (2.17). Therefore, the robust counterpart (2.15) can be solved using the literature on saddle-point problems and related algorithmic approaches, such as the one in [43]. In a similar way it is possible to solve the robust counterpart (2.16).

From a computational point of view, the authors describe an algorithm to generate the so-called robust efficient frontier. Given μ, Q and X, the efficient frontier is the collection of portfolios that are Pareto-optimal solutions to problem (2.16) when R varies (or to problem (2.15) when λ varies), [53]. Since the methods proposed in the literature to generate the efficient frontier, such as the method of critical lines in [53], are either not available or not implemented for the robust case, the authors propose and implement the following approach, that generates a discrete approximation to the robust efficient frontier. Firstly they determine the robust efficient portfolios with the lowest and highest expected returns, discretize the range between these two extremes to obtain a finite number of set levels of the expected return, and then solve problem (2.16) for each level of the expected return. To obtain the robust efficient portfolios with the lowest and highest expected returns, as well as to solve problem (2.16) for each intermediate value, they use the saddle-point algorithm developed by Halldórsson and Tütüncü in [43] which is an interior-point path-following method with computationally attractive polynomial-time convergence guarantees. The authors apply the robust asset allocation methods to two market data sets, and compare the robust efficient frontier to the classical efficient frontier, generated solving model (2.3). The first data set is composed of 5 asset classes, spanning the period January 1979 to July 2002, for a total of 283 months. The second data set addresses 8 asset classes over the period July 1983 to July 2002, for a total of 229 months. The computational analysis demonstrates that the portfolios generated using the proposed robust techniques have a significantly better worst-case behaviour than the classical MVO portfolios, i.e., the robust approaches guarantee a risk reduction for worst-case scenarios. For example, in one considered data set, the robust efficient portfolio achieving a 7.5% worst-case annualized expected return has an 8% standard deviation, while the classical efficient portfolio with a 7.5% worst-case annualized expected return has approximately 12% standard deviation, indicating that it is significantly riskier. Moreover, the generated robust portfolios show more stability over time, that is, they remain relatively unchanged over long periods of time. The overall conclusion of the authors is that, "by directly addressing some of the weaknesses of classical MVO models, the proposed robust models provide a valuable asset allocation vehicle to conservative investors ".

An alternative method for modeling the uncertainty was proposed in [41] using the factor model.

Let us consider a standard factor model for representing the return vector $r \in \mathbb{R}^n$, that is

$$r = \mu + V^T f + \varepsilon$$

where, according to the previous notation, μ is the expected return vector. In the formula, $f \in \mathbb{R}^m$ is the vector of random returns of the m $(< n)$ factors that drive the market, $V \in \mathbb{R}^{m \times n}$ denotes the matrix of factor loading and ε is the vector of residual returns. The parameters (μ, V) are estimated by linear regression. Given market data consisting of samples of asset return vectors and the corresponding factor returns, the linear regression procedure computes the least squares estimates (μ_0, V_0) of (μ, V).

The covariance matrix of the returns, Q, can be expressed as

$$Q = V^T F V + D,$$

where F is the covariance matrix of the factor returns and D is the diagonal matrix of the error term variances.

The individual elements d_i of the diagonal matrix D are assumed to belong to intervals $[\underline{d}_i, \bar{d}_i]$, i.e., the uncertainty set S_d for the error matrix D is given by:

$$S_d = \left\{ D : D = diag(d), d_i \in [\underline{d}_i, \bar{d}_i], i = 1, \dots, n \right\}.$$

On the other hand, the uncertainty set for the matrix of factor loading V is the following ellipsoidal set:

$$S_v = \{V : V = V_0 + W, \|W_i\|_G \leq \rho_i, i = 1, \dots, n\}$$

where $V_0 = [\bar{V}_1 \cdots \bar{V}_n]$ is the least square estimate of V, ρ_i are given bounds and W_i denotes the i-th column of W, for $i = 1, \cdots, n$, while $\|w\|_G = \sqrt{w^T G w}$ denotes the Euclidean (elliptic) norm of w with respect to a symmetric positive definite matrix G [1]. Finally, the expected return vector μ is assumed to lie in the uncertainty set S_m given by:

$$S_m = \{\mu : \mu = \mu_0 + \xi, |\xi_i| \leq \gamma_i, i = 1, \dots, n\}$$

where γ_i are given bounds (i.e., each component of μ is assumed to lie within a certain interval).

According to the factor model, the return of a portfolio x is:

$$r_x = r^T x = \mu^T x + f^T V x + \varepsilon^T x$$

where μ and V are the uncertain parameters. Similarly, its variance is:

$$x^T Q x = x^T V^T F V x + x^T D x,$$

where V and D are assumed to be uncertain.

By considering the uncertainty sets S_d, S_v and S_m above defined, Goldfarb and Iyengar derive the following robust analog of the Markowitz's mean-variance optimization problem (2.3):

$$\begin{aligned} \min \quad & \max_{\{V \in S_v, D \in S_d\}} x^T Q x \\ s.t. \quad & \min_{\{\mu \in S_m\}} \mathbb{E}(r_x) \geq R, \\ & x \in X \end{aligned} \tag{2.19}$$

where $\mathbb{E}(r_x)$ represents the expected value of the portfolio return.

At the same way, the authors derive the following robust counterpart of problem (2.2):

$$\max \min_{\{\mu \in S_m\}} \mathbb{E}(r_x)$$

[1] "*A way to define G is related to probabilistic guarantees on the likelihood that the actual realization of the uncertain coefficients will lie in the ellipsoidal uncertainty set S_v. Specifically, the definition of matrix G can be based on the data used to produce the estimates of the regression coefficients of the factor model*" [35].

$$\max_{\{V \in S_v, D \in S_d\}} x^T Q x \leq \sigma \tag{2.20}$$
$$x \in X$$

Under the hypothesis of normality of r_x (it is in fact assumed that r_x follows a normal distribution $\mathcal{N}(\mu^T x, x^T(V^T F V + D)x))$, and considering the uncertainty sets above defined, the authors prove that the robust optimization problem (2.19) can be reduced to a second order cone programming problem (SOCP) that can be solved very efficiently using interior point algorithms,[1, 51, 60, 76].

In fact, both the worst case and the practical computational effort required to solve a second order cone programming problem is comparable to that for solving a convex quadratic program of similar size and structure; in practice, the computational effort required to solve these robust portfolio selection problems is comparable to that required to solve the classical Markowitz mean-variance portfolio selection problems. The reduction to a second order cone programming problem is subsequently shown for the robust counterpart (2.20).

In [41], the authors further improve the factor model by addressing uncertainty also in the factor covariance matrix, and show that, for natural classes of uncertainty sets, all robust counterparts continue to be second order cone programming problems. They also show that these natural classes of uncertainty sets correspond to the confidence regions associated with the maximum likelihood estimation of the covariance matrix. Finally, computational experiments are performed using the Sharpe ratio variant of the proposed robust framework. This variant and the related computational results will be reviewed in Section 2.2.5.

2.2.5 The Sharpe ratio problem and its robust counterparts

A way to incorporate in asset allocation models also assets considered essentially riskless is to study the following optimization problem, called the *Sharpe Ratio Problem*:

$$\max \frac{\mu^T x - r_f}{\sqrt{x^T Q x}}$$
$$x \in X,$$

where r_f represents the known return on a riskless asset. The Sharpe ratio, i.e., $h(x) = \frac{\mu^T x - r_f}{\sqrt{x^T Q x}}$, is a measure that evaluates the excess return per unit of risk.

Since this maximization problem has a nonlinear and nonconcave objective function, and therefore it may be difficult to solve it directly in [41] the authors proposed an elegant argument to formulate the problem in terms of a convex minimization problem. All is based on the observation that $e^T x = 1$ whenever $x \in X$ (e represents an n-dimensional vector of 1's) since proportions in all securities must sum 1. Therefore the Sharpe ratio $h(x)$ can be rewritten as a homogeneous function of x, say $g(x)$, as follows:

$$h(x) = \frac{\mu^T x - r_f}{\sqrt{x^T Q x}} = \frac{(\mu - r_f e)^T x}{\sqrt{x^T Q x}} =: g(x) = g\left(\frac{x}{k}\right) \qquad k > 0. \tag{2.21}$$

The term $\mu - r_f e$ is the excess return of the risk-free rate. In [41], the authors proved that when X has the form (2.1), the optimal solution is not influenced if the normalization constraint $e^T x = 1$ is replaced with constraint $(\mu - r_f e)^T x = 1$.

In [79] the authors proved that a similar reduction can be achieved even when X is not in the form in (2.1), as long as $x \in X$ implies $e^T x = 1$. Under this assumption, and using the observation that $h(x) = g(x)$ and $g(x)$ is homogeneous, a portfolio x^* with the maximum Sharpe ratio can be found by solving the following problem:

$$\begin{aligned} &\min \ x^T Q x \\ &s.t \ (\mu - r_f e)^T x = 1 \\ &(x, k) \in X^+ \end{aligned} \tag{2.22}$$

where X^+ is a cone that lives in a one higher-dimensional space than X, and which is defined as follows:

$$X^+ = \left\{ x \in \mathbb{R}^n, k \in \mathbb{R} \mid k > 0, \frac{x}{k} \in X \right\} \cup \{(0,0)\}. \tag{2.23}$$

Moreover, the normalizing constraint can be relaxed to $(\mu - r_f e)^T x \geq 1$ by recognizing that this constraint will always be tight at an optimal solution.

Based on this observation, in [79] is proposed the following robust counterpart of the relaxed maximum Sharpe ratio problem, where the sets $\mathcal{U}_\mu = \left\{\mu : \mu^L \leq \mu \leq \mu^U\right\}$ and $\mathcal{U}_Q = \left\{Q : Q \succeq 0, Q^L \leq Q \leq Q^U\right\}$ are previously defined:

$$\begin{aligned} &\min \left\{ \max_{Q \in \mathcal{U}_Q} x^T Q x \right\} \\ &(x, k) \in X^+ \\ &\min_{\mu \in \mathcal{U}_\mu} (\mu - r_f e)^T x \geq 1. \end{aligned} \tag{2.24}$$

This robust counterpart, too, can be solved by reducing it to a second order cone programming problem [41], [79].

Let us consider the computational results about the robust counterpart (2.22) as reported in [41]. Like in the computational experiments performed in [79] for problems (2.15) and (2.16), the aim is to contrast the performance of the classical portfolio selection strategies with that of the robust portfolio selection strategies. Two types of computational tests are conducted, on simulated data and on real market data, by selecting the classical and robust portfolios via the corresponding maximum Sharpe ratio problems.

Consider the simulated data. By varying a confidence threshold ω from 0.01 to 0.95, where ω is used to generate the uncertainty sets the authors perform some independent runs; in particular, if ω is chosen very high, the uncertainty sets will be very large, leading to very conservative portfolios; on the other hand, if ω is chosen too low, the portfolio choice will not be robust enough. Then, they compare the mean Sharpe ratio of the robust portfolios to that of the classical portfolios, as well as the worst-case Sharpe ratio of the robust portfolios to that of the classical portfolios, where the worst-case Sharpe ratio of a portfolio is the minimum Sharpe ratio when the parameters vary in the uncertainty sets S_d, S_v and S_m introduced in Section 2.2.4.

The main result is that the mean performance of the robust portfolios does not significantly degrade as ω increases. On the other hand, the worst-case performance of the robust portfolios is about 200% better than the one of the classical portfolios. Moreover, robust portfolios are able to withstand noisy data considerably better than classical portfolios.

Concerning the real data, the authors perform an out-of-sample analysis in order to compare the classical and the robust strategies on a universe of 43 assets and 5 factors, spanning the period January 1997 to December 2000. Specifically, they divide the whole time series into investment periods of length 90 days; for each investment period they estimate the covariance matrix and the factor matrix, then setting all other parameters required in their model, they compute the robust and the classical portfolios solving the robust and the classical maximum Sharpe ratio problems, respectively. Finally, they compute the "actual" return of the generated portfolios by evaluating the return of the portfolios generated at period t with the return data available at $t+1$, and plot the relative performance of the robust strategy with respect to the classical strategy. For the particular data sequences used, the robust strategy appears to be clearly superior; i.e., it generates a larger return at a smaller computational cost when ω is sufficiently large. On the other hand, for small values of ω, i.e., when the portfolios generated through the robust strategy are not robust enough, no discernible trend is observed by the authors.

2.3 Robustness and risk measures

2.3.1 Risk measures

Recently, there has been an increasing interest in defining quantitative methods for assessing the risk of financial positions. Simplistically speaking, it is possible to distinguish between two types of risk measures: *dispersion* and *downside* measures [35]. Dispersion measures consider both positive and negative deviations from the expected return, and treat these deviations as equally risky. A well-known dispersion measure is portfolio standard deviation (and portfolio variance), used in the previously reviewed minimum-variance and mean-variance optimization models. On the other hand, downside risk measures traditionally address the probability that the portfolio return is above a minimal acceptable level.

The theory of risk measures has made significant progress in recent years. Here we introduce some advanced concepts of risk measures, together with related mean-risk optimization models.

Let Y be a real-valued function on a set Ω of possible scenarios that represents the return from an investment portfolio over a fixed period of time. A negative value for Y indicates loss. Then, a quantitative *measure of risk* can be modelled as a mapping ρ from the space of these return functions into the real line. This is the classical definition of measure of risk, as provided in [2] in their seminal contribution. In that paper, the authors have initiated a systematic analysis of the concept of risk measure, by formulating certain axioms which should be satisfied by any reasonable measure of risk. See also [30] for a review of risk measures in the framework of setting solvency capital requirements for a risky business.

In this paper, we introduce the definition of "monetary measure of risk" as in [38], as well the corresponding definitions of convex and coherent risk measures. From these definitions, Föllmer and Schied in [38] provided some risk measure characterizations that have been used, quite recently, to model interesting robust optimization models for asset allocation problems.

Let us denote by Φ the set of all bounded measurable functions on a set of scenarios Ω, and $\mathcal{P}$ be the set of all probability measures on Ω. For any $Y, Z \in \Phi$, the shortand notation $Y \leq Z$ denotes $Y(\omega) \leq Z(\omega)\ \forall \omega \in \Omega$. We define the following, [38]:

Definition 2.1. A mapping $\rho : \Phi \rightarrow \mathbb{R}$ is called a *monetary measure of risk* if it satisfies the following conditions for all $Y, Z \in \Phi$:

Monotonicity : if $Y \leq Z$, then $\rho(Y) \geq \rho(Z)$.

Cash Invariance: if $m \in R$, then $\rho(Y + m) = \rho(Y) - m$.

The financial meaning of monotonicity is that the downside risk of a position is reduced if the payoff profile is increased. Cash invariance, also called translation invariance, is motivated by the interpretation of $\rho(Y)$ as a capital requirement. Thus, if the amount m is added to the position and invested in a risk-free manner, the capital requirement is reduced by the same amount.

Definition 2.2. A monetary measure of risk $\rho : \Phi \to \mathbb{R}$ is called a *convex measure of risk* if it satisfies

Convexity: $\rho(\lambda Y + (1-\lambda) Z) \leq \lambda\rho(Y) + (1-\lambda)\rho(Z)\, \forall 0 \leq \lambda \leq 1$, $\forall Y, Z \in \Phi$.

The axiom of convexity states that diversification should not increase the risk.

Definition 2.3. A convex measure of risk ρ is called a *coherent risk measure* if it satisfies [2]

Positive homogeneity: if $\lambda \geq 0$, then $\rho(\lambda Y) = \lambda\rho(Y)$.

If a measure of risk ρ is positively homogeneous, then it is normalized, i.e., $\rho(0) = 0$.

In [2], Artzner et al. provided the following characterization of coherent risk measures (also recalled in [38]):

Theorem 2.3.1. *A functional $\rho : \Phi \to R$ is a coherent measure of risk if and only if there exists a subset $\mathcal{Q} \subseteq \mathcal{P}$ such that*

$$\rho(Y) = \sup_{q \in \mathcal{Q}} E_q[-Y], \quad Y \in \Phi, \tag{2.25}$$

where $E_q[-Y]$ denotes the mean value of $-Y$ (i.e., the expected loss) with respect to the probability q.

Such a characterization can be generalized to convex measures of risk in the following way, [36]:

[2] In [2], a mapping $\rho : \Phi \to \mathbb{R}$ is called a coherent measure of risk it it satisfies the four axioms of translation invariance, positive homogeneity, monotonicity and subadditivity, where subadditivity states that, $\forall Y, Z \in \Phi$, $\rho(Y+Z) \leq \rho(Y) + \rho(Z)$; convexity is a consequence of these axioms.

Theorem 2.3.2. *Suppose that* Ω *is a finite set* [3]*. Then* $\rho : \Phi \rightarrow R$ *is a convex measure of risk if and only if there exists a penalty function* $\alpha : \mathcal{P} \rightarrow R \cup (-\infty, +\infty]$ *such that*

$$\rho(Y) = \sup_{q \in \mathcal{P}} \left(E_q \left[-Y \right] - \alpha \left(q \right) \right). \tag{2.26}$$

The function α *satisfies* $\alpha(q) \geq -\rho(0)$ *for any* $q \in \mathcal{P}$*, and it can be taken to be convex and lower semicontinuous on* $\mathcal{P}$*.*

In the above characterization, function α has the role to assign a possibly different weight to the probabilities in $\mathcal{P}$, by suitably penalizing some of them. In particular, by choosing $\alpha(q) = 0$ for all $q \in \mathcal{Q}$, and $+\infty$ otherwise, the characterization of coherent measures of risk stated by Theorem 2.3.1 can be derived as special case. The characterization expressed by Theorem 2.3.2 has been recently exploited in order to propose more flexible robust models for asset allocation problems. This will be the subject of Section 2.3. Here let us review some well-known measures of risk together with related mean-risk optimization models.

In the literature, a well-known measure of risk is *Value at Risk* (VaR), developed by engineers at J.P. Morgan. VaR represents the predicted maximum loss with a specified probability level over a certain period of time. Let $f_x(\omega)$ denote the loss function of a portfolio $x \in X$ when ω is the realization of some random events (so, $f_x(\omega) = -Y$ according to our previous notation). Then, the α-VaR risk measure of x is defined as follows:

$$VaR_\alpha(x) = \min \left\{ \gamma : Prob \left(f_x(\omega) \geq \gamma \right) \leq 1 - \alpha \right\},$$

where α is a given probability level, and *Prob* denotes the probability with respect to a given reference probability on the set of scenarios Ω, say $p \in P$. In other words, VaR is defined as the minimum level γ such that the probability that the portfolio loss exceeds γ is less than or equal to $1 - \alpha$.

Some practical and computational issues related to VaR are discussed in [39], where the authors describe a method of calculating the portfolio giving the smallest VaR among those which yield at least a specified expected return. The method consists in approximating the historic VaR by a "smoothed" VaR which filters out local irregularities. Therefore, it is based on historic simulation. Several other approaches to VaR optimization are used in practice; in some contexts it is proved that VaR can be a suitable measure of risk, as in [59].

[3]There exists also a generalization of this theorem to the case where Ω is an infinite set [36].

However, VaR has also some undesirable properties as a risk measure. First, if studied in the framework of coherent risk measures [2], it lacks subadditivity (and therefore convexity). An additional difficulty with VaR may be in its computation and optimization. In fact, when VaR is calculated from generated scenarios, it turns out to be a nonsmooth and nonconvex function of the positions in the investment portfolio.

Another criticism of VaR is that it pays no attention to the magnitude of losses beyond the VaR value. This and other undesirable features of VaR led to the development of alternative risk measures. One well-known modification of VaR is the *Conditional Value at Risk* (CVaR), which measures the expected loss exceeding VaR. With respect to the classification of risk measures previously reviewed, CVaR classifies as a coherent risk measure (it is also a concave distortion risk measure), whereas VaR is a distortion risk meausure but it is not coherent (and therefore it is not a concave distortion risk measure).

Specifically, given a probability level α, the α-CVaR associated with a portfolio x is defined as follows:

$$CVaR_{\alpha}\left(x\right) = \frac{1}{1-\alpha}\int_{f_x(\omega)\geq VaR_{\alpha}(x)} f_x(\omega)p\left(\omega\right)d\omega$$

where, as before, $f_x(\omega)$ denotes the loss function when the portfolio x is choosen from the set X of feasible portfolios and ω is the realization of the random events, while $p\left(\omega\right)$ denotes the reference probability of ω.

In [70], the authors showed that minimizing CVaR can be achieved by minimizing a more tractable auxiliary function without predetermining the corresponding VaR first. They introduced the following simpler auxiliary function

$$F_{\alpha}(x,\gamma) = \gamma + \frac{1}{1-\alpha}\int_{f_x(\omega)\geq\gamma} f_x\left(\omega\right)p\left(\omega\right)d\omega. \tag{2.27}$$

This formulation can be written in the following equivalent way:

$$F_{\alpha}(x,\gamma) = \gamma + \frac{1}{1-\alpha}\int \left(f_x(\omega)-\gamma\right)^{+}p\left(\omega\right)d\omega \tag{2.28}$$

where $a^{+} = \max\{a,0\}$. The authors showed that $F_{\alpha}(x,\gamma)$ verifies some interesting properties such that minimizing CVaR is equivalent to minimize the auxiliary function $F_{\alpha}(x,\gamma)$, i.e., :

$$\min_{x\in X} CVaR_{\alpha}(x) = \min_{x\in X,\gamma} F_{\alpha}(x,\gamma).$$

Moreover, if $f_x\left(\omega\right)$ is a convex (linear) function of the portfolio variables x, then $F_{\alpha}(x,\gamma)$ is also a convex (linear) function of x. In this case, considering that the feasible portfolio set X is also convex, the above optimization problem is a smooth convex optimization problem that can be solved using well-known optimization techniques. In particular, the authors formulated the problem in the discrete case, obtaining a tractable formulation. Assume that the set of scenarios Ω comprises N scenarios $\omega_1, \ldots, \omega_N$, and that all scenarios have the same probability (so, $p(\omega) = \frac{1}{N}\ \forall\omega$). In this case the auxiliary function $F_\alpha(x,\gamma)$ can be approximated by the following function:

$$\tilde{F}_\alpha(x,\gamma) = \gamma + \frac{1}{(1-\alpha)N}\sum_{k=1}^{N}\left(f_x(\omega_k) - \gamma\right)^+ . \tag{2.29}$$

Hence, the problem $\min_{x\in X} CVaR_\alpha(x)$ can be approximated by replacing $F_\alpha(x,\gamma)$ with $\tilde{F}_\alpha(x,\gamma)$, obtaining the following formulation:

$$\begin{aligned}
&\min_{x,z,\gamma} \gamma + \frac{1}{(1-\alpha)N}\sum_{k=1}^{N} z_k \\
&s.t.\ z_k \geq 0, \quad k = 1,\ldots,N \\
&\qquad z_k \geq f_x(\omega_k) - \gamma, \quad k = 1,\ldots,N \\
&\qquad x \in X,
\end{aligned}$$

where z_k are artificial variables used to model $(f_x(\omega_k) - \gamma)^+$, $k = 1,\ldots,N$.

This formulation of CVaR usually results in convex programs and even linear programs (when $f_x(\omega)$ is a linear loss function). Thus, Rockafellar and Uryasev's work opened the door to the application of CVaR to financial optimization and risk management in practice.

In the following section we review some robust models that are based on VaR and CVaR.

2.3.2 Robust VaR and CVaR models

According to the previous definition, the portoflio α-VaR optimization problem consists in determining a portfolio x which minimizes the VaR risk measure with respect to a given confidence level α, i.e. :

$$\begin{aligned}
&\min \gamma \\
&s.t.\ Prob(f_x(\omega) \geq \gamma) \leq 1-\alpha \\
&x \in X.
\end{aligned} \tag{2.30}$$

Assuming that the asset returns are multivariate normally distributed (Normal VaR), in [41] the authors derived a robust counterpart of this formulation using the factor model to describe the future returns, and assuming that there is some errors in the estimation of the expected return vector and of the covariance matrix. No computational results have been provided by the authors concerning this robust counterpart.

However, in some contexts it may be not realistic to assume that the asset returns follow a multivariate normal distribution, as observed before. For this reason, El Ghaoui et al. in [34] proposed an alternative formulation. The authors assume that the distribution of the returns is partially known, in the sense that only bounds on the first two moments (mean and covariance matrix) are available. In particular, let $\mathcal{Q}$ denote the set of allowable distributions. For example, $\mathcal{Q}$ could represent the set of Gaussian distributions with mean $\hat{x}$ and covariance matrix Γ, where $\hat{x}$ and Γ are only known up to a given componentwise bound.

Then, given a probability level $\alpha \in (0, 1)$ and a portfolio $x \in X$, the authors introduce the notion of *Worst-case VaR* with respect to $\mathcal{Q}$ as follows:

$$\begin{aligned} &\min \gamma \\ &s.t. \sup_{q \in \mathcal{Q}} \; Prob_q(f_x(\omega) \geq \gamma) \leq 1 - \alpha, \end{aligned} \tag{2.31}$$

where $Prob_q$ denotes here the probability with respect to q. Calling $V_{\mathcal{Q}}(x)$ the Worst-case VaR, i.e., the optimum value of problem (2.31), the corresponding robust portfolio optimization problem is defined as follows:

$$\min V_{\mathcal{Q}}(x) \quad s.t. \quad x \in X. \tag{2.32}$$

In [34], El Ghaoui et al. prove that, for a large class of allowable probability distribution sets $\mathcal{Q}$, problem (2.32) can be solved via a semidefinite programming reformulation (SDP). Specifically, the authors examine two kinds of bounds, i.e., polytopic and componentwise, as well as uncertainty structures arising from factor models. A numerical example is then provided which compares the worst-case VaR of a nominal portfolio with that of its robust counterpart, computed via SDP. The results show that, if one chooses the nominal portfolio, data errors can have a dramatic impact on the VaR. On the other hand, taking into account the uncertainty via problem (2.32) dampens such a negative effect. For a review of SDP the interest reader is referred to [61, 71, 78, 80].

An attempt of robust optimization of the CVaR was proposed in [65], where the authors implemented in a robust way the bicriteria model [4] proposed by Rockafellar and Uryasev in [70], obtaining a robust linear reformulation of the problem. Three different versions are implemented and tested

[4]The goal is to form a portfolio in which the expected return is maximized, while some index of risk is minimized.

using market data. Also in this case, the experiments show that the robust counterpart of each model is able to select qualitatively good portfolios in terms of realized return.

Another variant of the CVaR optimization problem was formulated in [85], where the concept of Worst-case CVaR (WCVaR) is introduced. Given a probability threshold $\alpha > 0$, the *Worst-case CVaR* of a portfolio $x \in X$, with respect to a certain set $\mathcal{Q}$ of probability distributions, is defined as:

$$WCVaR_\alpha(x) = \sup_{p \in \mathcal{Q}} CVaR_\alpha(x).$$

In other words, as for the Worst-case VaR definition, the α-CVaR is computed for each probability $p \in \mathcal{Q}$, and the supremum value is then returned. Like CVaR, Worst-case CVaR remains a coherent risk measure.

In [85] the authors investigated the problem of minimizing WCVaR for several structures of $\mathcal{Q}$, reformulating the problem in a tractable form that can be efficiently solved. First, the authors consider the Worst-case CVaR in the case in which only partial information on the underlying probability distribution is given, i.e., the distribution p is known to belong to a set of distributions which consists of all convex combinations of some predetermined likelihood distributions p^h, $h = 1, ..., H$:

$$p \in \mathcal{Q}_M = \left\{ \sum_{h=1}^{H} \lambda_h p^h : \sum_{h=1}^{H} \lambda_h = 1, \lambda_h \geq 0, h = 1, \dots H \right\}. \tag{2.33}$$

Following Rockafellar and Uryasev's approach in [69], Zhu and Fukushima in [85] introduce the auxiliary functions

$$F_\alpha^h(x, \gamma) = \gamma + \frac{1}{1-\alpha} \int_\omega [f_x(\omega) - \gamma]^+ p^h(\omega) d\omega \tag{2.34}$$

where, as before, ω denotes the random events. Then, for each portfolio x, they prove that

$$WCVaR_\alpha(x) = \min_\gamma \max_{h \in \mathcal{H}} F_\alpha^h(x, \gamma)$$

where $\mathcal{H} = \{1, 2, \cdots, H\}$. Denoting $F_\alpha^{\mathcal{H}}(x, \gamma) = \max_{h \in \mathcal{H}} F_\alpha^h(x, \gamma)$, minimizing the Worst-case CVaR over X can then be achieved by minimizing $F_\alpha^{\mathcal{H}}(x, \gamma)$ over $X \times \mathbb{R}$, i.e., :

$$\min_{x \in X} WCVaR_\alpha(x) = \min_{(x,\gamma) \in X \times \mathbb{R}} F_\alpha^{\mathcal{H}}(x, \gamma). \tag{2.35}$$

Under the hypothesis that each likelihood distribution is characterized by a finite set of possible scenarios, and denoting by N_h the number of scenarios related to p^h, through a discretization procedure similar to that in [69] the authors obtain the following formulation, which is linear when the loss function $f_x(\omega)$ is linear, too, and the set X is a polyhedron:

$$\begin{aligned} & \min \theta \\ s.t. \quad & x \in X \\ & \gamma + \frac{1}{1-\alpha}(p^h)^T u^h \leq \theta \quad h = 1, \dots H \\ & u_k^h \geq f_x(\omega_{[k]}^h) - \gamma \quad k = 1, \cdots N_h \ h = 1, \dots H \\ & u_k^h \geq 0, \quad k = 1, \cdots N_h, \ h = 1, \dots H, \end{aligned} \tag{2.36}$$

where u_k^h is the k-th component of u^h and $\omega_{[k]}^h$ denotes the k-th sample of the likelihood distribution p^h.

In the second part of their work, the authors generalize the above model and, among various uncertainty structures, they focus on the minimization of WCVaR under box and ellipsoidal uncertainty sets associated with the distribution p. The considered box uncertainty set is the following:

$$p \in \mathcal{Q}_B = \left\{ p : p = p^0 + \eta, e^T \eta = 0, \underline{\eta} \leq \eta \leq \overline{\eta} \right\} \tag{2.37}$$

in which p^0 is a nominal distribution that represents the most likely distribution of the random component ω, e denotes the vector of ones, and $\underline{\eta}$ and $\overline{\eta}$ are given lower and upper bound vectors; in this case the problem is linear when $f_x(\omega)$ is a linear function and it is convex when the function is convex.

The ellipsoidal set is defined as follows:

$$p \in \mathcal{Q}_E = \left\{ p = p^0 + A\eta, e^T A\eta = 0, p^0 + A\eta \geq 0, \|\eta\| \leq 1 \right\} \tag{2.38}$$

in which p^0 is a nominal distribution that is the center of the ellipsoid, A is the scaling matrix of the ellipsoid and $\|\eta\| = \sqrt{\eta^T \eta}$. In this case the problem is convex when the loss function $f_x(\omega)$ is convex and it is a second order cone program when the function is linear.

The authors discuss robust portfolio selection problems corresponding to the types of uncertainties just described through two numerical examples. Market data simulation analysis and Monte Carlo simulation analysis are presented. Concerning the market data simulation, four sectoral sub-indices of Hang Seng Index of Hong Kong Stock Exchange (SEHK) are chosen, for which it is reasonable to assume a mixture distribution of the random returns, and therefore it makes sense to perform a Worst-case CVaR minimization. Numerical experiments for the nominal and the robust portfolio optimization problems are performed via the linear programming model (2.36), where it is assumed that the investor has an initial wealth, bound constraints are imposed on the portfolio, and a minimum expected return R is required. The former model employs the original CVaR as the risk measure to minimize, while the latter minimizes the Worst-case CVaR. Various nominal portfolio strategies and robust portfolio strategies are computed by setting different values of R. The authors show that the robust optimal portfolio almost always outperforms the nominal optimal portfolio in terms of portfolio value.

Concerning the Monte Carlo simulation analysis, the authors investigate the robust portfolio optimization model under the ellipsoidal uncertainty set (2.38). Considering the example given in [70], where the portfolio is to be constructed by three assets, the samples are generated via the Monte Carlo simulation approach by assuming a joint normal distribution. The scaling matrix of the ellipsoid, i.e., A, is assumed to be a diagonal matrix ρI, where ρ is a non-negative parameter which models uncertainty aspects. In particular, the nominal optimal portfolio is obtained by setting $A = 0$, i.e., $\rho = 0$. The main result is that the gap between the two curves (nominal versus robust portfolio value) becomes larger as ρ increases, which demonstrates the advantage of the robust optimization formulation in the situation where the uncertainty grows.

The numerical experiments thus imply that the portfolio selection models using the Worst-case CVaR as the risk measure in minimizing perform robustly in practice and provide flexibility in portfolio decision analysis; moreover, the experiments confirm that the specification of the uncertainty set is the key issue for successful practical applications.

2.3.3 Alternative robust models

Some critical aspects of traditional robust optimization took some authors to propose alternative robust models. Traditional robust approach is sometimes criticized for being overly conservative; moreover, this approach only guards against data realizations that are allowed by the given uncertainty set, while potentially becoming very vulnerable to realizations outside of the uncertainty sets considered. Finally, it tends to give the same weight to all possible data realizations (within the considered uncertainty sets), which may be unrealistic in practice.

A major critique that Bienstock moves in [16], and that he wants to overcome with his work, is tied to the notion of tractability of the models. The notion of tractability is without doubts a positive attribute, and certainly when a model is relevant, then ease of solution should be a worthwhile goal. However (and this is the critique moved by Bienstock) the solution methodology is often chosen at the expense of the richness and flexibility of the uncertainty model. Further, the underlying assumptions that often are used to justify convex uncertainty sets (such as normally distributed asset returns) potentially expose the user to a structural lack of robustness. Bienstock's view is that it is preferable to rank the risk modeling flexibility higher than the theoretical algorithmic performance.

Based on these observations, the author presents two non standard robust models for shortfalls in asset returns, which achieve an enhanced risk modeling flexibility by incorporating probability distribution and risk measure elements into the models. In constructing the models, the author assumes that a time series is available from which expected returns (and variances) are computed. This time series is used to construct rough distributions of the return shortfalls. In particular, Bienstock does not assume that the returns are normally distributed.

In the first of these models, called the *histogram* model, returns are segmented into a fixed number of categories, or bands, according to the magnitude of each shortfall; the distribution is obtained by employing an approximate count of the number of assets falling within each band. In the second model, the *ambiguous chance constrained* model, suitable probability distributions for the shortfalls are generated. Both models are solved using a cutting plane algorithm that runs a sequence of two step iterations: the first one solves an *implementor problem* which picks values for the decision variables of the model, the second one solves an *adversarial problem* which finds the worst-case data corresponding to the decision variables just selected by the implementor problem. The adversarial problem, in both cases, is a mixed-integer linear program. On the contrary, the implementor problem is, in the first case, a quadratic convex problem, while in the second case it is a quadratically constrained program solvable using SOCP techniques.

Starting from the risk-adjusted expected return model of Markowitz, Bienstock proposes the following formulation of the histogram problem:

$$\min_x \max_{\mu \in \mathcal{U}} \left\{ \lambda x^T Q x - \mu^T x \right\}$$

where $\mathcal{U}$ is the uncertain set of allowable return vectors. This problem is equivalent to

$$\min_x \left\{ \lambda x^T Q x - \min_{\mu \in \mathcal{U}} \mu^T x \right\} = \min_x \left\{ \lambda x^T Q x - \mathcal{A}(x) \right\}$$

where $\mathcal{A}(x)$ denotes the worst case return achieved by the asset vector x in the uncertainty set $\mathcal{U}$. The implementor problem consists on approximatively solving the problem

$$\min_x \left\{ \lambda x^T Q x - \min_{\mu \in \mathcal{U}} \mu^T x \right\}, \tag{2.39}$$

generating a portfolio x^*. Then, the corresponding adversarial problem yields a vector of asset returns corresponding to $\mathcal{A}(x^*)$, which is incorporated into the implementor problem. Specifically, let us consider the iteration h of the basic cutting plane algorithm, and let $\mu_{(j)}$, $1 \leq j \leq h-1$ be the return vectors produced by the priors runs of the adversarial problem. Then the implementor model at the iteration h is the following:

$$\begin{aligned} \min_x \quad & \lambda x^T Q x - r \\ s.t. \quad & Ax \geq b \\ & r - \mu_{(j)} x \leq 0 \quad 1 \leq j \leq h-1 \end{aligned}$$

where the constraints $Ax \geq b$ reflects a variety of restrictions on decisions. As we have said before, this is a convex quadratic program. After the solution of this model, the adversarial problem generates a new return vector, say $\mu_{(h+1)}$, which is the worst case return achieved by the current optimum portfolio in the considered uncertainty set $\mathcal{U}$. This return vector determines an additional constraint, $r - \mu_{(h+1)} x \leq 0$, which is incorporated into the implementor model at the next iteration.

The nature of the uncertainty set in the histogram model is tied to the concept of discretization of the risk, which is obtained by constructing suitable bands around an estimate of the expected value of the return shortfalls. Specifically, the return shortfalls are classified into a certain number of bands, whose width depends on the target of risk protection of the modeler. Then, given the specification of such an uncertainty set $\mathcal{U}$, and given the current optimum portfolio x^*, the adversary behavior is modeled via a mixed-integer linear programming formulation, which generates the worst case return achieved by x^* in $\mathcal{U}$, i.e., $\mathcal{A}(x^*)$. For more details and formal proofs, the interested reader is referred to [16].

In the *ambiguous chance constrained* model, risk measures such as VaR and CVaR are used to modeling the risk. In this model, at each iteration the adversary produces a probability distribution according to the so-called random model and, on the basis on such a distribution, he generates an adversarial return vector. This is obtained via a mixed-integer linear programming formulation, [16]. Then, the implementor chooses a portfolio x^* that minimizes the Worst-case VaR, say $VaR^{max}(x)$ (or the Worst-case CVaR, say $CVaR^{max}(x)$), which is the maximum VaR (CVaR) that can be incurred by the portfolio under the random model, that is under the choices made by the adversary. The implementor problem in the VaR (or CVaR) case has the following form:

$$\min V$$

$$
\begin{aligned}
s.t. \quad & \lambda x^T Q x - \mu^T x \le v^* + \varepsilon && (2.40)\\
& Ax \ge b && (2.41)\\
& V \ge VaR^{max}(x) && (2.42)\\
& (V \ge CVaR^{max}(x)) && (2.43)
\end{aligned}
$$

where $v^* = \min \lambda x^T Q x - \mu^T x$ subject to $Ax \ge b$, and $\varepsilon > 0$ is a positive tolerance parameter chosen by the modeler.

Constraint (2.40) formulates a concept of near-optimality, i.e., rather than focusing only on the minimization of $VaR^{max}(x)$ (or $CVaR^{max}(x)$), producing a possible conservative solution, the model finds solutions that are simultaneously near-optimal (with respect to the risk-adjusted expected return measure) and robust. Note that constraint (2.40) is a convex quadratic constraint, whereas constraint (2.42) (and its analogous (2.43)) is non-convex. To handle this kind of constraints, the author suitably approximates constraint (2.42) (and with analogous procedure constraint (2.43)), leading to a convex, quadratically constrained program, which depends on the adversarial returns which have been generated until the current iteration.

As far as the histogram model is concerned, using real-life data the author compares the robust portfolios generated via the histogram model to the classical optimal mean-variance portfolios, by comparing their composition and their nominal and worst-case returns. The performed experimentation shows that the robust portfolios take more positions than the mean-variance portfolios, and generally provide a higher worst-case return. On the other hand, considering the ambiguous chance constrained model, the author evaluates the performance of the model when some parameters change and for some data perturbation, by comparing the $VaR^{max}(x)$ and the $CVaR^{max}(x)$ formulations. The conclusion is that, in all cases, the proposed cutting plane approach is very fast and accurate, often requiring few seconds of computation, and it is capable of handling realistic uncertainty models with explicit non-convexities.

2.3.4 Risk measures theory and robust optimization

In a recent work [5], Ben-Tal et al. proposed a framework for robust optimization which relaxes the standard notion of robustness. Specifically, the authors focus on a relaxed approach in which not only the values of the uncertain parameters, but also their degree of feasibility are specified. The authors take the soft robust approach as starting point, link it to the theory of convex risk measures and then focus on tractability, structural properties, conservatism results and probabilistic implications in the framework of optimization under ambiguity.

Let us introduce such an idea of relaxed robustness in the context of loss functions. Given n assets, let $\tilde{r}$ denote the corresponding random return vector. Then, given a feasible portfolio $x \in \mathbb{R}^n$, define its associated loss function as $f_x(\omega) = -\tilde{r}^T x$ (so, the random events ω in the general expression $f_x(\omega)$ are modeled here in terms of random returns). Consider the following probabilistic constraint related to the loss of x:

$$-\tilde{r}^T x \leq b, \tag{2.44}$$

stating that the random loss of portfolio x must not exceed a threshold b, where b is a generic linear expression. According to the classical notion of robustness, the standard robust counterpart of constraint (2.44) should have the form:

$$-r^T x \leq b \quad \forall r \in \mathcal{U} \tag{2.45}$$

where $\mathcal{U}$ denotes a given uncertainty set for the random return vector.

The robust counterpart (2.45) can be equivalently written as:

$$-r^T x \leq b + \beta(r) \quad \forall r \in \mathbb{R}^n \tag{2.46}$$

where $\beta(r) = 0$ if $r \in \mathcal{U}$ and $+\infty$ otherwise. In other words, we can say that the robust counterpart's feasibility is ensured by employing a particular "extreme" penalty function, the indicator of $\mathcal{U}$.

The main observation in [5] is that milder penalty functions could be used in constraints like (2.46). These penalty functions arise by considering the risk-aversion of the investor. Specifically, the choice of alternative penalty function leads to penalize certain solutions rather than others. In presenting the soft robustness approach, we refer here to the Ph.D. Dissertation of D. Brown [21], where emphasis is just put on the concept of penalty functions.

The use of different "milder" penalty functions is related to the theory of convex risk measures developed by Föllmer and Schied in [38], as shown below. Let us assume that we have a number N of scenarios, say $\{r_1, \ldots, r_N\}$, that represent the possible realizations of the uncertain parameter $\tilde{r}$; let us denote by $\mathcal{P}$ the set of all probability measures on this discrete set of scenarios and let $\mathcal{R}$ be the convex hull of $\{r_1, \ldots, r_N\}$, so that any return vector $r \in \mathcal{R}$ can be expressed as $r = \sum_{r=1}^{N} r_i q_i$, $\sum_{r=1}^{N} q_i = 1$, $q_i \geq 0, i = 1, \ldots, N$. Then:

Theorem 2.3.3. *Let ρ be a convex risk measure, $\alpha(q)$ be the penalty function associated with ρ according to Theorem 2.3.2 and $\tilde{r}$ be the uncertain return vector. Then the following relations are equivalent:*

(A) $\rho(\tilde{r}^T x) \leq b$

(B) $-r^T x \leq b + \beta(r) \quad \forall r \in \mathcal{R}$

where

$$\beta(r) = \inf \left\{ \alpha(q) | q \in \mathcal{P}, \; r = \sum_{r=1}^{N} r_i q_i \right\}$$

Proof. Let $\mathcal{P}(r) = \left\{ q \in \mathcal{P} | r = \sum_{r=1}^{N} r_i q_i \right\}$.

$$\begin{aligned}
\rho(\tilde{r}^T x) \leq b &\Leftrightarrow \sup_{q \in \mathcal{P}} \left\{ E_q \left(-\tilde{r}^T x \right) - \alpha\left(q\right) \right\} \leq b \quad \textit{(from Theorem 2.3.2)} \\
&\Leftrightarrow E_q \left(-\tilde{r}^T x \right) - \alpha\left(q\right) \leq b, \quad \forall q \in \mathcal{P} \\
&\Leftrightarrow -\sum_{i=1}^{N} r_i^T x q_i - \alpha\left(q\right) \leq b, \quad \forall q \in \mathcal{P} \\
&\Leftrightarrow -r^T x - \alpha(q) \leq b, \quad \forall q \in \mathcal{P}(r), \; \forall r \in \mathcal{R} \\
&\Leftrightarrow -r^T x \leq b + \alpha(q), \quad \forall q \in \mathcal{P}(r), \; \forall r \in \mathcal{R} \\
&\Leftrightarrow -r^T x \leq b + \inf_{q \in \mathcal{P}(r)} \alpha\left(q\right), \quad \forall r \in \mathcal{R} \\
&\Leftrightarrow -r^T x \leq b + \beta(r), \quad \forall r \in \mathcal{R}.
\end{aligned}$$

□

Theorem 2.3.3 thus states that the relaxed notion of robustness above introduced corresponds to defining probabilistic type constraints, based on convex risk measures. The authors analyze different penalty functions in such convex risk measure constraints, so deriving different types of relaxed robustness. In particular, given a non-negative value δ, they consider particular penalty functions defined only for values less than or equal to δ, i.e., they focus on the following kind of constraints:

$$\sup_{\{q:\alpha(q)\leq\delta\}} \left(E_q(-\tilde{r}^T x) - \alpha(q) \right) \leq b. \tag{2.47}$$

In [5] they prove that the left hand side of (2.47) is equivalent to:

$$\begin{aligned}
&\min_{c\geq 0} \left\{ c\delta + \sup_{q\in\mathcal{P}} \left(E_q(-\tilde{r}^T x) - (c+1)\alpha(q) \right) \right\} = \\
&= \min_{c\geq 0} \left\{ c\delta + (c+1)\;\rho\left(\frac{\tilde{r}^T x}{c+1} \right) \right\}
\end{aligned} \tag{2.48}$$

where ρ is the normalized convex risk measure induced by the penalty function α.

As proved by the authors, the function in (2.48) is jointly convex in (c, x).

In the case in which we pose $b = 0$ in constraint (2.47), for each $x \in X$ and $\delta > 0$, the expression

$$\min_{c \geq 0} \left\{ c\delta + (c+1)\ \rho \left(\frac{\tilde{r}^T x}{c+1} \right) \right\} \leq 0 \tag{2.49}$$

is equivalent to choose a portfolio $x \in X^S_\delta$, where X^S_δ, called the *soft robust set*, is defined as follows:

$$X^S_\delta = \left\{ x \in X : \inf_{q \in \mathcal{P}(\varepsilon)} \mathbb{E}_q \left[\tilde{r}^T x \right] \geq -\varepsilon \quad \forall \varepsilon \in [0, \delta] \right\}. \tag{2.50}$$

The sequence of families $\mathcal{P}(\varepsilon) \subseteq \mathcal{P}$ represents sets of probability measures non decreasing on $\varepsilon \geq 0$.

Working on the particular penalty functions described above (i.e., those ones defined for values less than or equal to the threshold δ), the authors suggest several forms of soft robustness. Several chosen penalty functions are described in terms of the so-called ϕ-divergence. One of them is the following one:

$$\alpha(q) = \begin{cases} \sum_{i=1}^{N} p_i \phi\left(\frac{q_i}{p_i}\right) & \text{if } \sum_{i=1}^{N} p_i \phi\left(\frac{q_i}{p_i}\right) \leq \delta \\ +\infty & \text{otherwise} \end{cases} \tag{2.51}$$

where $\delta \geq 0$, while $\phi : \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ is a closed and convex function such that $\phi(0) = 1$ and *dom* $\phi \subseteq \mathbb{R}^+$. The term $\sum_{i=1}^{N} p_i \phi\left(\frac{q_i}{p_i}\right)$ is called the ϕ-divergence from q to the reference probability p. It represents a distance-like measure from q to p.

In their work, Ben Tal et al. study the following ϕ-divergence function $\phi(t) = \gamma^{-1} \left(t\ log(t) - t + 1 \right)$ at level γ. Scaling the ϕ-divergence penalty function by a positive factor $\frac{1}{\gamma}$, they study the relative entropy from q to p, i.e., $\alpha(q) = \frac{1}{\gamma} \sum_{i=1}^{N} log \frac{q_i}{p_i}$, and they show that, in such a case, the corresponding convex risk measure under consideration is $\rho_\gamma(\tilde{r}^T x) = \frac{1}{\gamma} log(\sum_{i=1}^{N} p_i e^{-\gamma r_i^T x})$, known as the *entropic risk measure at level* γ.

Now, suppose we wish to protect against all distributions q contained within δ-relative entropy of p in a soft way. Then the equivalence

$$x \in X^S_\delta \Leftrightarrow \min_{c \geq 0} \left\{ c\delta + (c+1)\rho \left(\frac{\tilde{r}^T x}{c+1} \right) \right\} \leq 0 \tag{2.52}$$

(recalled before), becomes the following one:

$$x \in X_{\delta}^{S} \Leftrightarrow \min_{c \geq 0} \left\{ c\delta + (c+1)\ log\ \mathbb{E}_p \left(e^{\frac{-\tilde{r}^T x}{c+1}} \right) \right\} \leq 0 \tag{2.53}$$

where like before the function $c\delta + (c+1)\ log\ \mathbb{E}_p \left(e^{\frac{-\tilde{r}^T x}{c+1}} \right)$ is jointly convex in (c, x).

In this way, Ben-Tal et al. use the relative entropy as the basis for deriving a specific soft robust approach.

In [5] the authors show that the complexity of this approach is equivalent to that of solving a small number of standard robust approaches and they illustrate the methodology on asset allocation example consisting of historical market data. Their focus is to investigate the performance of using a soft robust approach as opposed to a standard robust one.

Let us consider an investor who wishes to allocate wealth among n assets; the decision vector $x \in \mathbb{R}^n$ denotes the vector of weights the investor allocates to each asset, and X is the set of the feasible decision vectors x. Let us denote by $\tilde{r}$ the random return vector, as before. Then the investor solves the following problem:

$$\begin{aligned} &\max\ \mathbb{E}(\tilde{r}^T x) \\ &s.t.\ x \in X_{\delta}^{S} \end{aligned} \tag{2.54}$$

for some value δ chosen on the basis of the α value (the confidence level characterizing the CVaR model).

The above problem is then solved using monthly historical data related to 11 publicly traded asset classes spanning the period April 1981 to February 2006. The obtained portfolios are compared, via an out-of-sample analysis (such as the one described in Section 2.2.5) to the portfolios obtained by a variant of the soft approach called the "comprehensive soft robust approach" and to the portfolios generated imposing that CVaR is less than or equal to zero. The main result is that, on the considered data set, relaxing the standard robustness constraints to soft robustness constraints, such as the one expressed by the entropic risk measure constraint, guarantees a higher out-of-sample performance, expressed in terms of realized expected return, for not too high of a price in increased downside risk, expressed in terms of realized CVaR.

Chapter 3

A new class of robust asset allocation problems

3.1 Risk measure based models

A new family of robust asset allocation problems is here proposed. Starting from the relation between a relaxed type robustness and the convex risk measures recalled in previous chapter, we formulate risk measure based models in which alternative penalty functions $\alpha(q)$ are chosen.

The idea is to define penalty functions based on distance-like measures such as general norms, in order to propose portfolio optimization models that are computationally tractable.

The chapter begins with a more general formulation of Markowitz's problems, by exploiting then the relation between a relaxed robustness and convex risk measures, the so-called *norm-portfolio* family is introduced. Working on the penalty functions, we prove that this family includes, as special cases, linear programming (LP) and second order cone programming (SOCP) problems, i.e., tractable computational models.

Finally, we study a *coherent* variant of norm-portfolio family, i.e., we focus on the case in which the considered risk measures are also coherent. Some of the models proposed in this chapter will be tested with real market data in Chapter 4 in which also a comparison with different robust approaches described in literature is conducted.

3.2 The norm-portfolio models

Let us consider an environment with a set of n risky financial securities. Let $x \in \mathbb{R}^n$ represent a portfolio of n securities where x_j corresponds to the amount invested in security j. As in Chapter 2, let us denote by μ the mean return vector associated with the given positions and let $\tilde{r}$ be the vector of the random returns of the n positions.

Let us consider the following generalization of Markowitz's formulations, i.e., :

$$(P1)\quad \begin{array}{c} \sup\ \mu^T x \\ s.t.\ g(\tilde{r}^T x) \leq b \\ x \in X \end{array} \qquad (P2)\quad \begin{array}{c} \inf\ g(\tilde{r}^T x) \\ s.t.\ \mu^T x \geq R \\ x \in X \end{array} \qquad (P3)\quad \begin{array}{c} \sup\ \mu^T x - \lambda g(\tilde{r}^T x) \\ s.t.\ x \in X \end{array}$$

where $g(\tilde{r}^T x)$ is a generic convex risk function of portfolio's return, b is a generic linear expression which upper bounds the risk associated with the feasible portfolios, R is a threshold value on the expected return and

$$X = \left\{ x \in \mathbb{R}^n | \sum_{j=1}^{n} x_j = 1, x_j \geq 0, j = 1, \ldots, n \right\} \tag{3.1}$$

is the feasible portfolio set.

Let us observe that by setting $g(\tilde{r}^T x) = x^T Q x$ with Q (symmetric) positive semidefinite covariance matrix of $\tilde{r}$, then (P1), (P2) and (P3) collapse to the classical Markowitz's models for portfolio optimization.

In this work we investigate the problem (P2) when $g(\tilde{r}^T x) = \rho(\tilde{r}^T x)$ is a given convex risk measure, i.e.,:

$$\begin{array}{ll} \inf & \rho\left(\tilde{r}^T x\right) \\ s.t. & \mu^T x \geq R \\ & x \in X. \end{array} \tag{3.2}$$

Let us reformulate the problem (3.2) by introducing an auxiliary variable γ as follows:

$$\begin{array}{l} \inf\ \gamma \\ s.t.\ \rho\left(\tilde{r}^T x\right) \leq \gamma \\ \mu^T x \geq R \\ x \in X. \end{array} \tag{3.3}$$

Let us now exploit the relation between convex risk measures and the type of probabilistic constraints, previously reviewed.

Let us assume to know a set of return vectors $r_1, \ldots, r_N$, and let us denote by $\mathcal{P}$ the set of all probability measures on the finite set of scenarios $r_1, \ldots, r_N$.

This assumption captures a prevailing situation in many practical problems, when one has at his/her disposal N samples of the uncertain vector $\tilde{r}$, in the most of cases obtained from historical data. Then, from Theorem 2.3.3, the problem (3.3) is equivalent to the following one:

$$\inf \gamma \tag{3.4}$$

$$s.t. \sup_{q \in \mathcal{P}} \left\{ -\sum_{i=1}^{N} q_i r_i^T x - \alpha(q) \right\} \leq \gamma \tag{3.5}$$

$$\mu^T x \geq R \tag{3.6}$$

$$x \in X. \tag{3.7}$$

The formulation (3.4)-(3.7) describes the family of models on which we will work. According to the observations in Section 3.1, constraint (3.5) captures a relaxed notion of robustness. Such a constraint states that the weight of the probability q depends on a penalty function $\alpha(q)$ suitably defined, which, in turn, can be interpreted as a kind of distance between q and a reference probability, say $p \in \mathcal{P}$.

Mathematically, the notion of distance is often tied to the concept of norm. Based on this, we will study the special case of (3.4)-(3.7) where the penalty function $\alpha(q)$ is defined in terms of an arbitrary norm $|| \cdot ||$. The aim is to define relaxed robust models that can be computed in a very efficient way. Indeed, this choice of $\alpha(q)$ will lead to the investigation of special convex risk measures in formulation (3.4)-(3.7) that allows to obtain linear programming problem (LP) or second order cone programming (SOCP) problem, i.e., tractable models. The risk measures under investigation will be called the norm-risk measures, while the related models will be called the *norm-portfolio* models.

The most general formulation of *norm-portfolio* is therefore the following:

$$\begin{aligned} \inf \quad & \gamma \\ s.t. \ & \sup_{q \in \mathcal{P}} \left\{ -\sum_{i=1}^{N} q_i r_i^T x - \lambda \|\cdot\| \right\} \leq \gamma \\ & \mu^T x \geq R \\ & x \in X \end{aligned} \tag{3.8}$$

where $\|\cdot\|$ describes an arbitrary norm that measures the distance between a generic probability q and the reference probability p. The non-negative scalar λ is used to gauge this distance, as we will see better below.

In this work, the following norms will be addressed: the $||\cdot||_\infty$ norm, the $||\cdot||_1$ norm, the D-norm [15] and the Euclidean norm. In the first three cases, we will prove that the *norm-portfolio* models (3.8) can be formulated in terms of linear programming problems (LP); whereas, in the last case, i.e., when the penalty function is described in terms of the Euclidean norm, the model is reduced to a second order cone programming problem (SOCP).

Finally, in Chapter 4 we will test several models on real-world portfolio optimization data by performing a comparison with other robust models in literature.

3.2.1 The $||\cdot||_\infty$ norm case

Let $p \in \mathbb{R}^N$ and $q \in \mathbb{R}^N$ denote the reference probability and a generic probability, respectively with $p_i = \frac{1}{N}$ for each i. Let us consider the vector $(p-q) \in \mathbb{R}^N$ and its *infinity norm* (indicated with $||\cdot||_\infty$):

$$||p-q||_\infty = \max_i |p_i - q_i| \qquad \forall\, i = 1,\dots,N. \tag{3.9}$$

The first class of *norm-portfolio* models is obtained by setting the penalty function $\alpha(q)$ as follows:

$$\alpha(q) = \lambda||p-q||_\infty = \lambda \max_i |p_i - q_i| \tag{3.10}$$

where, as previously mentioned, the parameter λ is a non-negative scalar used to gauge the distance between the probabilities. In other words, we use the λ parameter to give different weights to probability q on the basis of its "distance" from the reference probability p.

Let us now replace the generic penalty function $\alpha(q)$ in the formulation (3.4)-(3.7) with the (3.10), obtaining the following problem

$$\begin{aligned} &\inf\ \gamma \\ s.t. \quad &\sup_{\substack{r=\sum_{i=1}^N r_i q_i \\ \sum_{i=1}^N q_i = 1 \\ q_i \geq 0, i=1,\dots,N}} \left\{-r^T x - \lambda \max_i |p_i - q_i|\right\} \leq \gamma \\ &\mu^T x \geq R, \quad x \in X. \end{aligned} \tag{3.11}$$

Let us analyze the role of parameter λ in the problem (3.11). $\lambda = 0$ is the case in which the investor gives the same weights to the probability measures q without considering the distance from the reference probability p; increasing the value of λ, the investor gives less weight to probabilities that are far away from the reference probabilities. The extreme case $\lambda \to \infty$ is the case in which the investor considers only $q = p$.

In terms of robustness, the case $\lambda = 0$ describes a model more conservative and hence more robust. The model becomes less robust as the parameter λ increases.

Returning on the problem (3.11), let us note that it is not a linear programming problem. However, it is well known that the norm $||\cdot||_\infty$ can be linearized. Therefore:

Theorem 3.2.1. *Under the norm $||\cdot||_\infty$, the norm-risk measure leads to a LP model.*

Proof. Let us consider the inner problem of (3.11), i.e.:

$$\begin{aligned} &\sup \left\{ -r^T x - \lambda \max_i |p_i - q_i| \right\} \\ s.t \quad & r = \sum_{i=1}^N r_i q_i \\ & \sum_{i=1}^N q_i = 1 \\ & q_i \geq 0, \quad i = 1, \dots, N. \end{aligned} \tag{3.12}$$

The objective function of the problem (3.12) can be easily linearized by introducing auxiliary variables z_i which bound from above the N absolute values, and an auxiliary variable z that models the maximum among the absolute values:

$$\begin{aligned} \sup \quad & \left\{ -\sum_{i=1}^N q_i r_i^T x - \lambda z \right\} \\ s.t. \quad & \sum_{i=1}^N q_i = 1 \\ & p_i - q_i \leq z_i, && i = 1, \dots, N \\ & -p_i + q_i \leq z_i, && i = 1, \dots, N \\ & z \geq z_i, && i = 1, \dots, N \\ & q_i \geq 0, && i = 1, \dots, N. \end{aligned}$$

Observe that now we can replace the *sup* operator by the *max* operator. Indeed the optimal value of a linear function on a polytope is always attained, i.e.:

$$\begin{aligned}
\max \quad & \left\{ -\sum_{i=1}^{N} q_i r_i^T x - \lambda z \right\} \\
s.t. \sum_{i=1}^{N} q_i = 1 & & \rightarrow \mathbf{(u)} \\
-q_i - z_i \leq -p_i, & \quad i = 1, \dots, N & \rightarrow (w_i^+) \\
q_i - z_i \leq p_i, & \quad i = 1, \dots, N & \rightarrow (w_i^-) \\
-z + z_i \leq 0, & \quad i = 1, \dots, N & \rightarrow (v_i) \\
q_i \geq 0, & \quad i = 1, \dots, N.
\end{aligned}$$

where the variables are q_i and z_i, $i = 1, \dots, N$, plus z. Since the above linear programming problem is non-empty, and its objective function is bounded from above using the strong duality we can replace it with its dual:

$$\begin{aligned}
\min \quad & u - \sum_{i=1}^{N} \left(w_i^+\right) p_i + \sum_{i=1}^{N} \left(w_i^-\right) p_i \\
s.t. \quad & u - (w_1^+) + (w_1^-) \geq -r_1^T x \\
& \vdots \\
& u - (w_N^+) + (w_N^-) \geq -r_N^T x \\
& - w_1^+ - w_1^- + v_1 = 0 \\
& \vdots \\
& - w_N^+ - w_N^- + v_N = 0 \\
& - v_1 - v_2 - \dots - v_N = -\lambda \\
& w_i^+, w_i^-, v_i \geq 0 \quad i = 1, \dots, N.
\end{aligned}$$

Equivalently:

$$\begin{aligned}
\min \quad & u - \sum_{i=1}^{N} (w_i^+ - w_i^-) p_i \\
s.t. \quad & u - w_i^+ + w_i^- \geq -r_i^T x \quad i = 1, \dots, N \\
& - w_i^+ - w_i^- + v_i = 0 \quad i = 1, \dots, N \\
& - \sum_{i=1}^{N} v_i = -\lambda \\
& w_i^+, w_i^-, v_i \geq 0 \quad i = 1, \dots, N.
\end{aligned}$$

The problem (3.11) can be then formulated in the following compact way where the *inf* operator has been replaced again by the *min* operator since the optimal value is always attained:

$$
\begin{aligned}
& \min \gamma \\
s.t. \quad & u - \sum_{i=1}^{N} (w_i^+ - w_i^-) p_i \leq \gamma \\
& u - w_i^+ + w_i^- \geq - \sum_{j=1}^{n} (r_{ij} x_j) \quad i = 1, \ldots, N \\
& - w_i^+ - w_i^- + v_i = 0 \quad i = 1, \ldots, N \\
& \sum_{i=1}^{N} v_i = \lambda \\
& w_i^+, w_i^-, v_i \geq 0 \quad i = 1, \ldots, N \\
& \mu^T x \geq R \\
& \sum_{j=1}^{n} x_j = 1 \\
& x_j \geq 0 \quad j = 1, \ldots, n.
\end{aligned}
\tag{3.13}
$$

This is a Linear Programming problem. The result follows. □

3.2.2 The $|| \cdot ||_1$ norm case

Let $p \in \mathbb{R}^N$ and $q \in \mathbb{R}^N$ denote the reference and a generic probability respectively. As before, $p_i = \frac{1}{N}$ for each i. Let us consider the vector $(p-q) \in \mathbb{R}^N$ and its L_1 *norm* (indicated with $\|\cdot\|_1$):

$$
||p - q||_1 = \sum_{i=1}^{N} |p_i - q_i|. \tag{3.14}
$$

The second family of *norm-portfolio* models is obtained by setting the penalty function $\alpha(q)$ as follows:

$$
\alpha(q) = \lambda ||p - q||_1 = \lambda \sum_{i=1}^{N} |p_i - q_i| \tag{3.15}
$$

where the parameter $\lambda \in \mathbb{R}^+$ is a non-negative scalar used to gauge the distance between the probabilities as previously mentioned.

Let us now replace the generic penalty function $\alpha(q)$ in the formulation (3.4)-(3.7) with the (3.15). The following theorem holds true:

Theorem 3.2.2. *Under the L_1 norm, the norm-risk measure leads to a LP problem:*

Proof. Let us consider the inner problem:

$$\begin{aligned} \sup & -\sum_{i=1}^{N} q_i r_i^T x - \lambda \sum_{i=1}^{N} |p_i - q_i| \\ s.t. & \sum_{i=1}^{N} q_i = 1 \\ & q_i \geq 0, i = 1, \dots, N. \end{aligned}$$

As before, let us introduce the auxiliary variables z_i, for $i = 1, \dots, N$, which bound from above the absolute values, and rewrite the problem in the following way:

$$\begin{aligned} \sup & -\sum_{i=1}^{N} q_i r_i^T x - \lambda \sum_{i=1}^{N} z_i & \\ s.t. & \sum_{i=1}^{N} q_i = 1 & \\ & p_i - q_i \leq z_i & i = 1, \dots, N \\ & -p_i + q_i \leq z_i & i = 1, \dots, N \\ & q_i \geq 0 & i = 1, \dots, N. \end{aligned}$$

Now we can replace the *sup* operator by the *max* operator. In fact, the optimal value of a linear function on a polytope is always attained:

$$\begin{aligned} \max & -\sum_{i=1}^{N} q_i r_i^T x - \lambda \sum_{i=1}^{N} z_i & & \\ s.t. & \sum_{i=1}^{N} q_i = 1 & & \to (u) \\ & -z_i - q_i \leq -p_i & i = 1, \dots, N & \to (w_i^+) \\ & -z_i + q_i \leq p_i, & i = 1, \dots, N & \to (w_i^-) \\ & q_i \geq 0 & i = 1, \dots, N. & \end{aligned}$$

Since this linear programming problem is a maximization problem over a non-empty polytope and the objective function is bounded from above we can replace it with its dual within the overall formulation by referring to the strong duality theory:

$$
\begin{aligned}
&\min\ u - \sum_{i=1}^{N} p_i w_i^+ + \sum_{i=1}^{N} p_i w_i^- && \\
&s.t.\ u - w_i^+ + w_i^- \geq -r_i^T x, && i = 1, \dots N, \\
&\quad -w_i^+ - w_i^- = -\lambda, && i = 1, \dots N, \\
&\quad\ w_i^+, w_i^- \geq 0, && i = 1, \dots N.
\end{aligned}
$$

The overall formulation is therefore:

$$
\begin{aligned}
&\min \quad \gamma && \\
&s.t.\ u - \sum_{i=1}^{N} \left(w_i^+ - w_i^-\right) p_i \leq \gamma && \\
&\quad u - (w_i^+ - w_i^-) \geq - \sum_{j=1}^{n} r_{ij} x_j, && i = 1, \dots N, \\
&\quad w_i^+ + w_i^- = \lambda, && i = 1, \dots N, \\
&\quad w_i^+, w_i^- \geq 0 && i = 1, \dots N, \\
&\quad \mu^T x \geq R,\ x \in X &&
\end{aligned}
$$

which is a LP problem. □

3.2.3 The D-norm case

In this section we address a particular norm, called the D-norm, which has been introduced by Bertsimas et al. in [15]. We show that also under this scenario the *norm-portfolio* model can be reduced to a LP problem. Firstly, let us introduce the following definition:

Definition 3.1. Given a non-negative integer m ($m \leq N$), let us define the D-norm of vector $(p - q) \in \mathbb{R}^N$ in the following way:

$$
\|p - q\|_m = \max_{\substack{S \subseteq \{1,\dots,N\} \\ |S| \leq m}} \left\{ \sum_{i \in S} |p_i - q_i| \right\}.
$$

In other words, the D-norm can be defined as the sum of the m largest absolute values of the entries of a vector $p - q$.

Notice that, when $m = 1$, then this norm coincides with the L_∞ norm and when $m = N$ then the D-norm coincide with the L_1 norm, i.e., L_1 and L_∞ norms are special cases of the D-norm.

Let us consider the *norm-portfolio* model, and set

$$\alpha(q) = \lambda||p - q||_m, \quad \lambda \geq 0 \tag{3.16}$$

within the model. The analogous of Theorem 3.2.1 and Theorem 3.2.2 is the following one:

Theorem 3.2.3. *Under the D-norm, the norm-risk measure leads to a LP model.*

Proof. Let us consider the inner problem:

$$\begin{aligned} &\sup -r^T x - \lambda||p - q||_m \\ &s.t.\ r = \sum_{i=1}^{N} r_i q_i \\ &\sum_{i=1}^{N} q_i = 1 \\ &q_i \geq 0, i = 1, \ldots, N. \end{aligned} \tag{3.17}$$

Introducing an auxiliary variable z in order to bound the D-norm from above, we have the following problem:

$$\begin{aligned} &\sup -\sum_{i=1}^{N} q_i r_i^T x - \lambda z \\ &s.t.\quad \|p - q\|_m \leq z \\ &\sum_{i=1}^{N} q_i = 1 \\ &q_i \geq 0 \quad i = 1, \ldots N. \end{aligned} \tag{3.18}$$

By exploiting then the definition of D-norm the following problem is obtained:

$$
\begin{aligned}
& \sup -\sum_{i=1}^{N} q_i r_i^T x - \lambda z \\
s.t. \quad & \max_{\substack{S \subseteq \{1,\dots,N\} \\ |S| \leq m}} \left\{ \sum_{i \in S} |p_i - q_i| \right\} \leq z \\
& \sum_{i=1}^{N} q_i = 1 \\
& q_i \geq 0 \quad i = 1, \dots, N.
\end{aligned}
\tag{3.19}
$$

Let us introduce additional variables φ_i in order to model the absolute values $|p_i - q_i|$, so the problem becomes:

$$
\begin{aligned}
& \sup -\sum_{i=1}^{N} q_i r_i^T x - \lambda z \\
s.t. \quad & \max_{\substack{S \subseteq \{1,\dots,N\} \\ |S| \leq m}} \sum_{i \in S} \varphi_i \leq z \\
& p_i - q_i \leq \varphi_i, \qquad i = 1, \dots, N \\
& -(p_i - q_i) \leq \varphi_i, \qquad i = 1, \dots, N \\
& \sum_{i=1}^{N} q_i = 1 \\
& q_i \geq 0 \quad i = 1, \dots, N.
\end{aligned}
\tag{3.20}
$$

Constraint

$$
\max_{\substack{S \subseteq \{1,\dots,N\} \\ |S| \leq m}} \sum_{i \in S} \varphi_i \leq z \tag{3.21}
$$

includes the second inner problem of the given problem, that is:

$$
\max_{\substack{S \subseteq \{1,\dots,N\} \\ |S| \leq m}} \sum_{i \in S} \varphi_i \tag{3.22}
$$

where $\{\varphi_i\}$ for $i = 1, \dots, N$ are considered as constant values.

Let us observe that the second inner problem (3.22) can be formulated as the following knapsack problem:

$$\begin{aligned} \max & \sum_{i=1}^{N} \varphi_i y_i \\ s.t. & \sum_{i=1}^{N} y_i \le m \\ & y_i \in \{0,1\}, \quad i = 1,\dots,N. \end{aligned} \tag{3.23}$$

Let us consider the linear relaxation of the problem (3.23):

$$\begin{aligned} \max & \sum_{i=1}^{N} \varphi_i y_i & \\ s.t. & \sum_{i=1}^{N} y_i \le m & \rightarrow (\pi) \\ & y_i \le 1 \quad i = 1,\dots,N & \rightarrow (\pi_i) \\ & y_i \ge 0 \quad i = 1,\dots,N. & \end{aligned} \tag{3.24}$$

In such a special case, the linear relaxation provides the optimum objective function value of (3.23). Since the feasible set of the linear relaxation is bounded and not empty, using the strong duality we can replace it with its dual:

$$\begin{aligned} \min\ & m \cdot \pi + \sum_{i=1}^{N} \pi_i \\ s.t.\ & \pi + \pi_i \ge \varphi_i \quad i = 1,\dots,N \\ & \pi, \pi_i \ge 0 \quad i = 1,\dots,N. \end{aligned}$$

Therefore, the inner problem (3.17) can be equivalently rewritten as :

$$
\begin{aligned}
& \sup -\sum_{i=1}^{N} q_i r_i^T x - \lambda z \\
& s.t.\ m\pi + \sum_{i=1}^{N} \pi_i - z \leq 0 && \to (\delta) \\
& \quad -\pi - \pi_i + \varphi_i \leq 0, \quad i = 1, \ldots, N && \to (v_i) \\
& \quad \pi \geq 0,\ \pi_i \geq 0, \quad i = 1, \ldots, N && \\
& \quad -q_i - \varphi_i \leq -p_i, \quad i = 1, \ldots, N && \to (w_i^+) \\
& \quad q_i - \varphi_i \leq p_i, \quad i = 1, \ldots, N && \to (w_i^-) \\
& \quad \sum_{i=1}^{N} q_i = 1 && \to (u) \\
& \quad q_i \geq 0,\ i = 1, \ldots, N
\end{aligned}
\tag{3.25}
$$

where the variables are $(q_i, z, \pi, \pi_i, \varphi_i)$.

By replacing now the *sup* operator by the *max* operator since the optimal value of a linear function on a polytope is always attained and by considering that the inner problem (3.25) is not empty as well as its objective function is bounded from above, we replace it by its dual problem in order to get the following linear programming model:

$$
\begin{aligned}
\min\ & -\sum_{i=1}^{N}(w_i^+)p_i + \sum_{i=1}^{N}(w_i^-)p_i + u \\
s.t.\ & -w_i^+ + w_i^- + u \geq -r_i^T x && i = 1, \ldots, N \\
& -\delta = -\lambda \\
& m\delta - \sum_{i=1}^{N} v_i \geq 0 \\
& \delta - v_i \geq 0 && i = 1, \ldots, N \\
& v_i - w_i^+ - w_i^- = 0 && i = 1, \ldots, N \\
& w_i^+, w_i^-, v_i, \delta \geq 0 && i = 1, \ldots, N.
\end{aligned}
$$

Rewriting the variables v_i in terms of w_i^+ and w_i^- and the variable δ in terms of λ, we can simplify the problem in the following way:

$$\begin{aligned}
\min\ & -\sum_{i=1}^{N}(w_i^+)p_i + \sum_{i=1}^{N}(w_i^-)p_i + u \\
s.t.\ & -w_i^+ + w_i^- + u \geq -r_i^T x && i = 1, \ldots, N \\
& m\lambda - \sum_{i=1}^{N}(w_i^+ + w_i^-) \geq 0 && i = 1, \ldots, N \\
& \lambda - w_i^+ - w_i^- \geq 0 && i = 1, \ldots, N \\
& w_i^+, w_i^- \geq 0 && i = 1, \ldots, N.
\end{aligned}$$

The complete problem is the following:

$$\begin{aligned}
& \min\ \gamma \\
s.t.\ & -\sum_{i=1}^{N}(w_i^+ - w_i^-)p_i + u \leq \gamma \\
& -w_i^+ + w_i^- + u \geq -\sum_{j=1}^{n} r_{ij}x_j && i = 1, \ldots, N \\
& m\lambda - \sum_{i=1}^{N}(w_i^+ + w_i^-) \geq 0 && i = 1, \ldots, N \\
& \lambda - w_i^+ - w_i^- \geq 0 && i = 1, \ldots, N \\
& v_i = w_i^+ + w_i^- && i = 1, \ldots, N \\
& w_i^+, w_i^- \geq 0 && i = 1, \ldots, N \\
& \mu^T x \geq R \\
& x \in X
\end{aligned}$$

which is a Linear Programming problem. □

3.2.4 The Euclidean norm case

Finally, let us consider the case in which the penalty function $\alpha(q)$ is described by the Euclidean norm. In this case we prove that the *norm-portfolio* model can be formulated in terms of a Second Order Cone Programming (SOCP) problem.

Let us specialize the problem as follows:

$$
\begin{aligned}
& \inf \gamma \\
s.t. \quad & \sup_{\substack{r=\sum_{i=1}^{N} r_i q_i \\ \sum_{i=1}^{N} q_i = 1 \\ q_i \geq 0, i=1,\ldots,N}} \left\{ -r^T x - \lambda \|p - q\|_2 \right\} \leq \gamma \\
& \mu^T x \geq R, \quad x \in X
\end{aligned}
$$

where $\lambda \geq 0$ is a given value and $p_i = \frac{1}{N}$ for each i denotes the reference probability.

Theorem 3.2.4. *Under the Euclidean norm, the norm-risk measure leads to a SOCP model.*

Proof. Let us consider the following inner problem:

$$
\begin{aligned}
& \sup - \sum_{i=1}^{N} q_i {r_i}^T x - \lambda \|p - q\|_2 \\
s.t. \quad & \sum_{i=1}^{N} q_i = 1 \\
& q_i \geq 0, \quad i = 1, \ldots, N.
\end{aligned}
$$

By introducing an auxiliary variable z, the problem is equivalent to:

$$
\begin{aligned}
& \sup - \sum_{i=1}^{N} q_i {r_i}^T x - \lambda z \\
s.t. \quad & \sum_{i=1}^{N} q_i = 1 \\
& q_i \geq 0, \quad i = 1, \ldots, N \\
& \|p - q\|_2 \leq z
\end{aligned}
$$

Now, we replace the *sup* operator by the *max* operator, in fact the optimal value of a linear function (hence a continuous function) on a compact set is always attained and we set $y = p - q$. So, let us rewrite the formulation above replacing $q = p - y$, recalling then that $p_i = \frac{1}{N}$ for each i represents the reference probability, we obtain:

$$\begin{aligned}
&\max -\sum_{i=1}^{N}(\frac{1}{N}-y_i){r_i}^T x-\lambda z\\
&s.t.\ \sum_{i=1}^{N}\left(\frac{1}{N}-y_i\right)=1\\
&\qquad \frac{1}{N}-y_i\geq 0,\quad i=1,\ldots,N\\
&\qquad \|y\|_2\leq z.
\end{aligned}$$

The last expression of the above problem is equivalent to say $(z,y)\in C_q$, where C_q is a second-order cone.

Introducing additional auxiliary non-negative variables, we have:

$$\begin{aligned}
&-\frac{1}{N}\sum_{i=1}^{N}r_i^T x\ +\ \max\sum_{i=1}^{N}y_i r_i^T x-\lambda z\\
s.t.\ &\sum_{i=1}^{N}y_i=0 && \rightarrow (w_0)\\
&y_i+s_i=\frac{1}{N}\quad i=1,\ldots,N && \rightarrow (w_i)\\
&\|y\|_2\leq z\qquad (\equiv (z,y)\in C_q)\\
&s_i\geq 0\quad i=1,\ldots,N\qquad (\equiv s\in C_l\ \textit{nonnegative orthant})
\end{aligned}$$

Since the problem is feasible ($y=0, z=0, s_i=\frac{1}{N}\ \forall i$ is a feasible solution), from the Conic duality theorem [6] the optimum objective function value is equal to the one of its dual:

$$\begin{aligned}
&-\frac{1}{N}\sum_{i=1}^{N}r_i^T x\ +\ \min\frac{1}{N}\sum_{i=1}^{N}w_i\\
s.t.\quad &w_0+w_i\underset{C_q^*}{\geq} r_i^T x\quad i=1,\ldots,N\\
&0\underset{C_q^*}{\geq}-\lambda\\
&w_i\underset{C_l^*}{\geq}0\qquad i=1,\ldots,N,
\end{aligned}$$

where C_q^* and C_l^* denote the dual cones of C_q and C_l respectively.

The first two groups of constraints (i.e., the ones related to the dual cone C_q^*) are equivalent to $((w_0+w_1-r_1^T x),\ldots,(w_0+w_N-r_N^T x),\lambda)\in C_q^*$. Therefore, since the second order cone is equal to its dual cone, these constraints can be equivalently rewritten as:

$$\left\|(w_0 + w_1 - r_1^T x), \ldots, (w_0 + w_N - r_N^T x)\right\|_2 \le \lambda.$$

Moreover, since the dual cone of the non-negative orthant is the non-negative orthant, constraints $w_i \underset{C_i^*}{\ge} 0, i = 1, \ldots, N$ are equivalent to constraints $w_i \ge 0, i = 1, \ldots, N$. Therefore, the overall model is equivalent to:

$$\begin{aligned}
&\min \gamma \\
s.t. \ & -\frac{1}{N}\sum_{i=1}^{N} r_i^T x \ + \frac{1}{N}\sum_{i=1}^{N} w_i \le \gamma \\
&\left\|(w_0 + w_1 - r_1^T x), \ldots, (w_0 + w_N - r_N^T x)\right\|_2 \le \lambda \\
&w_i \ge 0 \qquad i = 1, \ldots, N \\
&\mu^T x \ge R, \quad x \in X
\end{aligned}$$

which is a SOCP model. □

3.3 Coherent variant of the norm-portfolio models

The family of the *norm-portfolio* models described in the Section 3.2 has a common feature: the norm-risk measure used within each model is a *convex* risk measure based on an arbitrary norm. In this section we study the variant of norm-portfolio models where the considered risk measure is also *coherent*.

As proved in [36] and reviewed in Section 2.3, a coherent risk measure arises from some families $\mathcal{Q}$ of probability measures by computing the expected loss under $q \in \mathcal{Q}$ and then by taking the worst result as q varies over $\mathcal{Q}$, i.e.:

$$\rho(Y) = \sup_{q \in \mathcal{Q}} E_q[-Y], \quad Y \in \Phi. \tag{3.26}$$

By exploiting such a characterization, the coherent version of the *norm-portfolio* family (3.8) can therefore be defined as follows:

$$
\begin{array}{ll}
 & \inf \gamma \\
s.t. & \sup\limits_{\substack{r=\sum_{i=1}^{N} r_i q_i \\ \sum_{i=1}^{N} q_i=1 \\ q_i \geq 0,\ i=1,\ldots,N \\ ||\cdot|| \leq \pi,}} -r^T x \leq \gamma \\
 & \mu^T x \geq R, \quad x \in X
\end{array}
\tag{3.27}
$$

where $||\cdot||$ is a generic norm and the parameter π represents an upper bound on the distance measure.

Let us now specialize the family (3.27) using the infinity norm in the following way (analogously we can specialize the family considered using the other three norms defined in the previous sections):

$$
\begin{array}{ll}
 & \inf \gamma \\
s.t. & \sup\limits_{\substack{r=\sum_{i=1}^{N} r_i q_i \\ \sum_{i=1}^{N} q_i=1 \\ q_i \geq 0,\ i=1,\ldots,N \\ ||p-q||_\infty \leq \pi,}} -r^T x \leq \gamma \\
 & \mu^T x \geq R, \quad x \in X.
\end{array}
\tag{3.28}
$$

As previously mentioned parameter π represents an upper bound on the distance measure between the probability q and the reference probability p. Consequently, the bound π belongs to the interval $[0, 1]$ where the extreme values describe the case in which the distance between the probabilities is null, i.e., the case in which the probability q coincides with the reference probability p, and the case in which one considers the set of all probabilities on the given set of scenarios respectively.

In terms of robustness the case $\pi = 0$ is therefore the least conservative and hence the least robust one.

We prove that, under the norm $||\cdot||_\infty$, problem (3.27) can be reduced to a Linear Programming Problem.

Let us consider the inner problem of the (3.28):

$$\begin{array}{lll} \max & -\sum_{i=1}^{N} q_i r_i^T x & \\ s.t. & \sum_{i=1}^{N} q_i = 1 & \\ & (p_i - q_i) \leq \pi, & i = 1, \dots, N \\ & (-p_i + q_i) \leq \pi, & i = 1, \dots, N \\ & q_i \geq 0 & i = 1, \dots, N \end{array}$$

in which we replace the *sup* operator by the *max* operator (since the maximum of a linear function over a polytope set is achieved).

Equivalently:

$$\begin{array}{llll} \max & -\sum_{i=1}^{N} q_i r_i^T x & & \\ s.t. & \sum_{i=1}^{N} q_i = 1 & & \rightarrow \mathbf{(u)} \\ & -q_i \leq \pi - p_i, & i = 1, \dots, N & \rightarrow (w_i^+) \\ & q_i \leq \pi + p_i, & i = 1, \dots, N & \rightarrow (w_i^-) \\ & q_i \geq 0, & i = 1, \dots, N. & \end{array}$$

Now, let us write the dual of the above problem (DP):

$$\begin{array}{ll} \min & u + \sum_{i=1}^{N} w_i^+ (\pi - p_i) + \sum_{i=1}^{N} w_i^- (\pi + p_i) \\ s.t. & u - (w_1^+) + (w_1^-) \geq -r_1^T x \\ & \vdots \\ & u - (w_N^+) + (w_N^-) \geq -r_N^T x \\ & w_i^+, w_i^- \geq 0 \qquad i = 1, \dots, N. \end{array}$$

Equivalently:

$$\begin{array}{lll} \min & u + \sum_{i=1}^{N} (w_i^+ + w_i^-)\pi - \sum_{i=1}^{N} (w_i^+ - w_i^-) p_i & \\ s.t. & u - w_i^+ + w_i^- \geq -r_i^T x & i = 1, \dots, N \\ & w_i^+, w_i^- \geq 0 & i = 1, \dots, N. \end{array}$$

The complete problem is the following:

$$
\begin{aligned}
&\min \gamma \\
s.t.\quad & u + \sum_{i=1}^{N} (w_i^+ + w_i^-)\pi - \sum_{i=1}^{N} (w_i^+ - w_i^-)p_i \le \gamma \\
& u - w_i^+ + w_i^- \ge - \sum_{j=1}^{n} (r_{ij} x_j) \qquad i = 1, \dots, N \\
& w_i^+, w_i^- \ge 0 \qquad i = 1, \dots, N \\
& \mu^T x \ge R \\
& \sum_{j=1}^{n} x_j = 1 \\
& x_j \ge 0 \quad j = 1, \dots, n
\end{aligned}
\tag{3.29}
$$

that is a Linear Programming Problem. Let us observe that also the *coherent variant* (3.29) of norm-portfolio models implements a relaxed form of robustness under a particular penalty function that assumes value 0 for all probability q distant from the reference probability p at most π, and infinity otherwise.

3.3.1 Some considerations about the bound π

Among the values that the parameter π can assume in the model (3.27), there is one, denoted by π^*, that allows us to describe an interesting relation between the coherent risk measure introduced in Section 3.3, based on the norm $|| \cdot ||_\infty$, and the most note CVaR measure (that is also coherent).

As before, let N denote the number of the considered samples. In addition, let α be the confidence level chosen to define the $CVaR_\alpha$ risk measure. Let us assume that $(1-\alpha) \ge \frac{1}{N}$ and let η and τ be the quotient and the rest respectively obtained by dividing the quantity $(1-\alpha)$ by $\frac{1}{N}$. Let us denote by p^* the following value:

$$
p^* = \frac{p_i}{1-\alpha} + \frac{\tau}{\eta(1-\alpha)} = \frac{1}{N(1-\alpha)} + \frac{\tau}{\eta(1-\alpha)}. \tag{3.30}
$$

Let us assume $\eta < N$ and define

$$
\pi^* = \max \left\{ \frac{1}{N}, \left| \frac{1}{N} - p^* \right| \right\}. \tag{3.31}
$$

Then:

Proposition 3.3.1. *If a discrete probability distribution is associated with the random events, and all scenarios have the same probability, then setting $\pi = \pi^*$ the coherent variant (based on the norm $||\cdot||_\infty$) is equivalent to the $CVaR_\alpha$ model (in terms of optimal portfolio).*

Proof. W.l.o.g. consider the case in which the loss function of a feasible portfolio x is $-\tilde{r}^T x$ where, as previously introduced, $\tilde{r}$ denotes the vector of the random returns. Let us recall the definition of VaR_α:

$$VaR_\alpha(x) = \min\left\{\gamma : Prob(-\tilde{r}^T x \geq \gamma) \leq 1 - \alpha\right\}. \tag{3.32}$$

Assume now that a discrete probability distribution is associated with the random return $\tilde{r}$. In particular, let $r_1, \ldots, r_N$ be the samples related to $\tilde{r}$, and let $p_i = \frac{1}{N}$ denote the probability of the sample i, $i = 1, \ldots, N$. Under these assumptions, the CVaR of portfolio x, $CVaR_\alpha(x)$, can be calculated as follows [26]:

$$CVaR_\alpha(x) = \frac{1}{1-\alpha} \sum_{i:-r_i^T x \geq VaR_\alpha(x)} \frac{1}{N}(-r_i^T x). \tag{3.33}$$

Let us distinguish two cases:

Case $\tau = 0$

Let us assign the probability $\tilde{p}_i = p^* = \frac{1}{N(1-\alpha)}$ to the η scenarios with greatest loss, and $\tilde{p}_i = 0$ otherwise (remember that $\eta < N$ by assumption). Now, if we choose a bound $\pi \geq \pi^*$ within the coherent variant of the norm-portfolio models, then such a probability is feasible for the coherent variant; in fact the following constraints are satisfied:

1) $\tilde{p}_i \geq 0$ for $i = 1, \ldots, N$

2) $\sum_{i=1}^{N} \tilde{p}_i = 1 \Leftrightarrow$

$$\begin{aligned}
&\Leftrightarrow \sum_{i=1}^{N-\eta} \tilde{p}_i + \sum_{i=N-\eta+1}^{N} \tilde{p}_i = 1 \\
&\Leftrightarrow 0 + (\eta)\frac{1}{N(1-\alpha)} = 1 \\
&\Leftrightarrow N(1-\alpha)\frac{1}{N(1-\alpha)} = 1
\end{aligned}$$

3) $\max_i |p_i - \tilde{p}_i| \leq \pi^*$
To prove the above condition, let us distinguish the following two sub-cases:
a) $\tilde{p}_i = 0$:

$$\max_i |p_i - 0| \leq \pi^* \Leftrightarrow \frac{1}{N} \leq \max\left\{\frac{1}{N}, \left|\frac{1}{N} - p^*\right|\right\}$$

that is always satisfied;
b) $\tilde{p}_i \neq 0$:

$$\max_i |p_i - \tilde{p}_i| \leq \max\left\{\frac{1}{N}, \left|\frac{1}{N} - p^*\right|\right\} \Leftrightarrow$$

$$\Leftrightarrow \left|\frac{1}{N} - \frac{1}{N(1-\alpha)}\right| \leq \max\left\{\frac{1}{N}, \left|\frac{1}{N} - \frac{1}{N(1-\alpha)}\right|\right\}$$

where the last inequality is always satisfied.

In addition, setting $\pi = \pi^*$ the probability $\{\tilde{p}_i\}$ is also the probability that maximizes the objective function of the inner problem (3.27). Indeed, the maximum possible increment of the probability (with respect to the reference probability $p_i = \frac{1}{N}$ for each scenario i) is given by the scenarios with greatest loss.

But this is exactly the probability addressed in (3.33). Hence, we can state that an optimal portfolio for the coherent variant is optimal also for the $CVaR_\alpha$ model. The result follows.

Case $\tau \neq 0$

The proof is the same of the case $\tau = 0$. In this case, we assign the probability $\tilde{p}_i = p^* = \frac{1}{N(1-\alpha)} + \frac{\tau}{\eta(1-\alpha)}$ to the η scenarios with greatest loss, and $\tilde{p}_i = 0$ otherwise. Again, setting $\pi = \pi^*$ the probability $\{\tilde{p}_i\}$ is such that maximizes the objective function of the inner problem (3.27). In fact, the maximum possible increment of the probability (with respect to the reference probability $p_i = \frac{1}{N}$ for each scenario i) is given by the scenarios with the highest loss. However, the optimum value of the objective function does not necessarily coincide with the optimum value returned by the $CVaR_\alpha$ model.

□

Corollary 3.3.2. *For each $\pi \geq \pi^*$, the coherent variant of the $||\cdot||_\infty$-portfolio models is "more robust" than the $CVaR_\alpha$ model, in the sense that the considered probability set includes the one (implicitly) addressed by the $CVaR_\alpha$ model.*

Chapter 4

Computational analysis

In this chapter we use real market data set in order to describe the performance of some models proposed in Chapter 3. The analysis could be ideally divide in two parts. In the first one we focus on the *norm-portfolio* family with norm $\| \|_\infty$ and on its coherent variant in order to shed light on their "actual" performances; in the second part we compare these two families with a classical CVaR model, a slight variant of Tütüncü-Koenig model in [79] and the *entropic* one in [5].

The models have been chosen as benchmark to provide a preliminary comparison among the related concepts of robustness: the relaxed robustness (characterizing the *norm-portfolio* models and their coherent variant), a classical robustness based on uncertainty sets for the covariance matrix (characterizing Tütüncü-Koenig model) and the soft robustness (characterizing the *entropic* model). In addition, we compare these approaches to an approach based on CVaR that is related to a relaxed robustness of the *coherent* variant, as proved in Chapter 3. Hence, the second part of our analysis aims at examining how the different concepts of robustness influence the optimal value and its regularity in the long term.

The three data sets for conducting the numerical experiments are provided by Tütüncü (the first two) and by Byrne, and by them used in [79] and [18] respectively.

Each case study consists of a twofold analysis: an *in-sample* analysis, whose aim is to determine suitable values of parameters that describe the models, and an *out-of-sample* analysis, in which we utilize the scenario information based on past history in producing a portfolio strategy.

4.1 Plan of the experiments

4.1.1 Input Data

Before proceeding to a detailed description of the in-sample and the out-of-sample analysis, we provide the input data of the models we are studying (*norm-portfolio* models, *Coherent* variant, CVaR model, *Tütüncü-Koenig model*, *entropic* model). Let us now distinguish the data common to all models and those specific one for each model.

The common input data are the following:

- a matrix D of data whose elements r_{ij} with $i = 1, \ldots, N$ and $j = 1, \ldots, n$ describe the monthly returns of each asset j;

- the vector μ calculated as the mean of the columns of the data matrix D, i.e., μ_j - the j^{th} component of vector μ - denotes the mean return of security j, that is $\mu_j = \frac{\sum_{i=1}^{N} r_{ij}}{N}$ for $j = 1, \ldots, n$.

- the upper bound R in the constraint $\mu^T x \geq R$, calculated as the mean value of the vector μ.

Let us observe that the parameter R represents the minimum return that the investor would be willing to receive. To describe a realistic behaviour, we always take a non-negative value of R (in the cases in which R results a negative value, we set $R = 0$).

The specific input data of each family of models are the following:

- the reference probability $p_i = \frac{1}{N}$ for $i = 1, \ldots, N$ used in the *norm-portfolio* models, in their *Coherent* variant and in the *entropic model*;

- the confidence level $\alpha = 0.9$ used for the *CVaR* and the *entropic* models (this value of α is one of the most used values in CVaR optimization problems).

- the parameter $\delta = \gamma^{-1} log\left(\frac{1}{\alpha}\right)$ used in the *entropic* model which is described in a detailed way below (this choice has been suggested in [5]);

- the upper bound matrix Q^U in *Tütüncü-Koenig* model, recalled below, which denotes an upper bound of the covariance matrix Q generated through a method based on quantiles like in [79].

Let us recall the following robust counterpart of the Markowitz's problem (2.3) described in Chapter 2:

$$\begin{aligned} &\min \max_{Q \in \mathcal{U}_Q} x^T Q x \\ &s.t. \ \min_{\mu \in \mathcal{U}_\mu} \mu^T x \geq R \\ &\quad x \in X \end{aligned} \tag{4.1}$$

where the covariance matrix Q and the expected vector μ belong to the uncertain sets $\mathcal{U}_Q = \{Q : Q^L \leq Q \leq Q^U, Q \succeq 0\}$ and $\mathcal{U}_\mu = \{\mu : \mu^L \leq \mu \leq \mu^U\}$ respectively.

One solving method for (4.1) is to study the following formulation:

$$\begin{aligned} &\min x^T Q^U x \\ &s.t. \ (\mu^L)^T x \geq R \\ &\quad x \in X \end{aligned} \tag{4.2}$$

that it is correct when Q^U is a positive semidefinite matrix. In this way, the problem (4.2) becomes a standard quadratic problem (QP).

In our work, we investigate a slight variant of (4.1) (and hence of (4.2)) in which the uncertainty is described only in terms of covariance matrix. The problem we study is the following one:

$$\begin{aligned} &\min x^T Q^U x \\ &s.t. \ \mu^T x \geq R \\ &\quad x \in X. \end{aligned} \tag{4.3}$$

To generate the upper bound matrix Q^U we choose moving windows of four years and compute the covariance matrix in each such window. Let us then consider the first covariance matrix and let us extract the component corresponding to the first raw and the first column; we repeat this procedure for all covariance matrices putting all extracted elements in a unique vector b_{11} (where the subscript indicates the position of the components, i.e., the components in the first raw and in the first column of each submatrix). In analogous way, we extract all components of the covariance matrices obtaining n^2 vectors b_{ij} for $i, j = 1, \ldots, n$.

For each vector b_{ij}, we compute the 95 percentile and then we construct a matrix $\tilde{Q}$ that contains all such percentile's values. The obtained matrix thus represents an upper bound for Q (because Q corresponds to the 50 percentile). Unfortunately, nothing assures that the matrix $\tilde{Q}$ is a positive semidefinite (and hence a covariance) matrix. So, in the cases this property does not hold, we solved the following subproblem that enables to compute a covariance matrix "nearest" to $\tilde{Q}$, which bounds $\tilde{Q}$ from above. Given a $n \times n$ matrix $\tilde{Q}$, a general formulation of the problem we solve is the following:

$$\begin{aligned} &\min \sum_{i=1}^{n}\sum_{j=1}^{n} q_{ij}^U \\ s.t.\ & Q^U \succeq 0 \\ & q_{ij}^U \geq \tilde{q}_{ij} \quad i = 1,\ldots,n \quad j = 1,\ldots,n. \end{aligned} \tag{4.4}$$

The problem (4.4) is a semidefinite programming problem that has been implemented in a MatLab 7.7 (R2008b) environment with Yalmip toolbox.

Concerning the soft robust approach, the model used in this experiment is a slight variant of that one described by Ben Tal et al. in [5]. Following the standard formulation of the portfolio asset allocation problem until now used, we minimize a risk measure subject to constraints on the return. Like penalty function we choose the ϕ-divergence measure at level $\gamma = 1$, i.e.:

$$\alpha(q) = log\mathbb{E}_p\left[\phi\left(\frac{dq}{dp}\right)\right]$$

and like loss function the expression $-\tilde{r}^T x$, in which $\tilde{r}$ is the random return vector [1].

Following then the equivalence (2.48), we obtain the problem below:

$$\begin{aligned} &\min_{c,x}\ c\ \delta + (c+1)log\left[\mathbb{E}_p\left(e^{\frac{-\tilde{r}^T x}{c+1}}\right)\right] \\ &s.t.\ \mu^T x \geq R \\ &\quad x \in X \end{aligned} \tag{4.5}$$

where $X = \left\{x \in \mathbb{R}^n | \sum_{j=1}^{n} x_j = 1,\ x_j \geq 0, j = 1,\ldots,n\right\}$ is the feasible portfolio set and $\delta = log\left(\frac{1}{\alpha}\right)$.

The problem (4.5) is a convex programming problem in variables (c,x) that we solve by performing a binary research on the variable c: we look

[1]The choice of γ reflects the degree of the investor's risk-aversion. As observed in [5], γ could be interpreted as the reciprocal of the risk tolerance for a CARA utility. A low value of γ corresponds to a high risk aversion. In this preliminary comparison it has been chosen $\gamma = 1$ in order to describe a more conservative approach, but in the future researches it could be interesting observing as the performance of the model changes increasing γ.

for a rounded value of c that minimizes the objective function of the (4.5). Setting then this value, we compute the optimal portfolio x that solves the problem (4.5). To implement the problem, the MatLab system cvx 6.1 for convex optimization is used.

4.1.2 In-sample analysis

As it has been said, the *in-sample* analysis lies in defining the most suitable parameters values of the models we investigate in this work. Let us start from the family of the *norm-portfolio* models (3.11), described in Section 3.2; as we have shown, the family (3.11) can be formulated as a λ-parametric linear programming problem family. In order to choose the most suitable values of λ, we proceed in the following way: at first, by fixing all other parameters we solve the *norm-portfolio* models for all integer λ values belonging to the interval $[0, 120]$, plus a further λ value that describes the performance of the models at infinity (we choose $\lambda = 10^7$). Observing then the composition of the portfolio, its risk and its return we choose the suitable values for the out-of-sample experiment. This procedure is adopted for all three considered data sets.

Regarding the second family of models, i.e., the *coherent* variant of the *norm-portfolio* models, we test the model for some values of parameter $\pi \in [0, 1]$ (by fixing the other input data) and we choose again the suitable values of π in terms of composition of portfolio, risk and return. Among the considered values of π, there is one that is of particular interest, denoted by π^*, that describes the relation between the coherent variant of the norm-portfolio models and the *CVaR* model, as proved in Chapter 3, i.e.,:

$$\pi^* = \max\left\{\frac{1}{N}, \left|\frac{1}{N} - p^*\right|\right\} \tag{4.6}$$

where N is the number of periods considered and p^* is defined in the (3.30).

4.1.3 Out-of-sample analysis

The *out-of-sample* procedure enables to observe how the different models utilize the scenario information in producing a successful portfolio. To compare the different models we use a sort of moving windows method; the length of each window is $T = 12$ which for monthly data correspond to 1 year (i.e., we subdivide the whole time period in subperiods 1 year long). Using the monthly return data input in the windows, we work out the mathematical models described in Chapter 3: we compute the composition of the optimal portfolio and its value obtained choosing the return of the following month as return vector.

Like example (concerning the first data set that we will see in the next section) let us take the 12 monthly returns from January 1979 to December 1979 as the initial historical data for constructing the first portfolio to invest in, and calculate the portfolio's realized value by observing the historical returns for the following month, i.e., for January 1980. Then, we move a month and we consider the second subperiod, i.e., from February 1979 to January 1980; we observe the historical return for February 1980 and calculate the value of the portfolio obtained. We continue doing this until the end of data set is reached and we repeat the procedure for each family of the considered models (*norm-portfolio* models, *coherent* variant, *CVaR* model, *Tütüncü-Koenig* model and *entropic* model) and for each value of the parameters chosen in the *in-sample* analysis. At the end of this process we generate $N - T$ portfolio vectors x_t for $t = T, \dots, N-1$ with N the total numbers of samples in considered data set.

Finally, we evaluate the out-of-sample performance of each model according to the following statistics: mean-realized returns, variance and Sharpe Ratio of realized return and the portfolio turnover. In other words, holding the portfolio x_t for one period gives the following out-of-sample *realized return* at time $t+1$: $\hat{r}_{t+1} = x_t^T r_{t+1}$ where r_{t+1} denotes the historical return at that time. After collecting the time series associated to $N - T$ realized returns $\hat{r}_t$ we evaluate the out-of-sample mean $\hat{\mu}$, the out-of-sample variance $\hat{\sigma}^2$, the out-of-sample Sharpe Ratio $\hat{SR}$ and the portfolio turnover defined as follows:

$$\begin{aligned}
\hat{\mu} &= \frac{1}{N-T} \sum_{t=T}^{N-1} x_t^T r_{t+1} \\
(\hat{\sigma})^2 &= \frac{1}{N-T-1} \sum_{t=T}^{N-1} (x_t r_{t+1} - \hat{\mu})^2 \\
\hat{SR} &= \frac{\hat{\mu}}{\hat{\sigma}} \\
Turnover &= \frac{1}{N-T-1} \sum_{t=T}^{N-1} \sum_{j=1}^{n} |x_{j,t+1} - x_{j,t}|
\end{aligned}$$

where $x_{j,t+1}$ and $x_{j,t}$ are the portfolio weight in asset j at time $t+1$ and at t respectively.

The out-of-sample mean evaluates the mean of realized returns in each subperiod; the out-of-sample variance is a measure of sample variation of realized returns with respect to the mean value.

The out-of-sample Sharpe Ratio measures the realized return per unit of risk. Finally, the portfolio turnover is a measure of variability in the portfolio holdings and can indirectly indicate the magnitude of the transaction costs associated to each strategy.

In the following sections we present a detailed description of the *out-of-sample* analysis for each data set considered and comment the final results.

The tested linear programming problems, i.e., the *norm-portfolio* models, the *coherent* variant and the *CVaR* model have been implemented both in the *in-sample* and in the *out-of-sample* analysis using Tomlab/Cplex v11.2 within MatLab 7.7 (R2008b).

The experiments related to *Tütüncü-Koenig* model (i.e. a quadratic programming problem) have been conducted using Yalmip MatLab's toolbox.

Finally, the *entropic* soft robust model, as we have said, has been solved with MatLab system cvx 6.1.

4.2 The first computational test

In this first experiment we use a universe of five asset classes: large and small cap growth stocks, large and small cap value stocks and fixed income securities. Each class is represented through the monthly log-return time series of corresponding market indices: Russel 1000 growth, Russell 1000 value, Russell 2000 growth, Russell 2000 value and Lehman Brothers U.S. Government/Credit Bond, that span the period January 1979 to July 2002, i.e., a total of $N = 283$ months. For each index we have the monthly returns (in percentages), so the complete data consist of a matrix D (283×5) in which the raws and the columns correspond to the time periods and the market indices respectively.

In this first test, the suitable values of the parameter λ chosen using the *in-sample* analysis are the following: $\lambda = 0$, $\lambda = 5$, $\lambda = 10$, $\lambda = 15$, $\lambda = 20$, $\lambda = 25$, $\lambda = 30$, $\lambda = 35$ and $\lambda = 10^7$, where this last value (as it has been said) describes the performance of the model when λ goes to infinity.

For the coherent variant (3.27) described in Section 3.3, the parameter we have to choose is π that represents, as it has been said, a sort of distance between the reference probability p and the generic probability q. The parameter π assumes values between 0 and 1. Again, using the *in-sample* analysis we choose the following values: $\pi = 0$ and $\pi = 1$ as extreme values, and $\pi = 0.25$, $\pi = 0.5$ and $\pi = \pi^*$ like intermediate values, where π^* is obtained as in (4.6).

Following now the *moving-windows* procedure, we subdivide the whole time period in subperiods 1 year long obtaining in this case a total of $N - T = 271$ subintervals. For each subinterval (to which corresponds a submatrix of data) we compute the composition of the optimal portfolio x_t and its realized return $\hat{r}_{t+1}$ calculated as $\hat{r}_{t+1} = x_t r_{t+1}$ for $t = T, \dots, N-1$.

We start applying this procedure to the *norm-portfolio* models described in the (3.11). The figures 4.2-4.4 illustrate the historical trajectories of optimal portfolio value (realized return) calculated under different values of the parameter λ; let us observe that we chose to plot the performance of the *norm-portfolio* models only for a subset of the parameters λ computed in the in-sample analysis, i.e., the most significant ones. We divide the whole time period in four subperiods to provide a clear representation of this performances.

Let us note that on the Y axis we report the portfolio's "realized" value in percentages relatively to the initial portfolio value and on the X axis the corresponding time period that goes from $t = 13$ (January 1980) to $t = 283$ (July 2002).

All performances are quite similar approximately, but let us note that increasing λ leads to a slight increase of the portfolio value, but also it produces severe drops. In other words, increasing the value of this parameter, in some cases enables to attain higher returns (such as at $t = 254$ that corresponds to February 2000), but also may produce drops such as at $t = 263$ (November 2000).

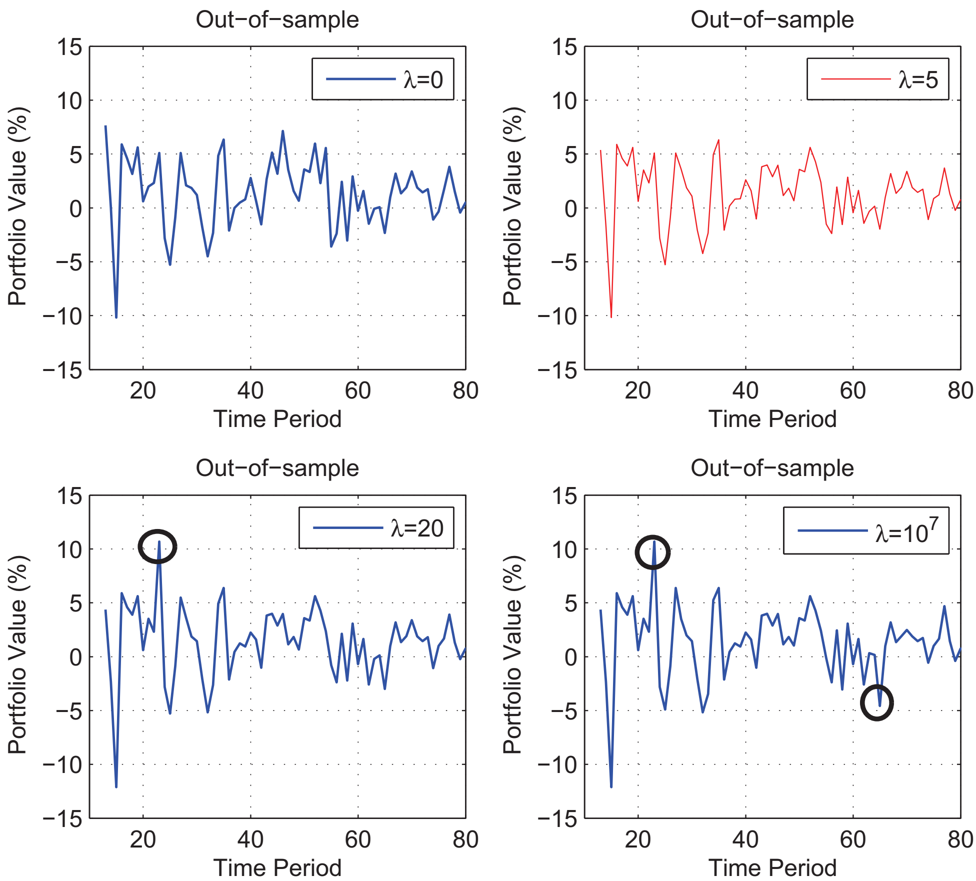


Figure 4.1: *Evolution of realized return related to the norm-portfolio models for different λ values in the first subperiod*

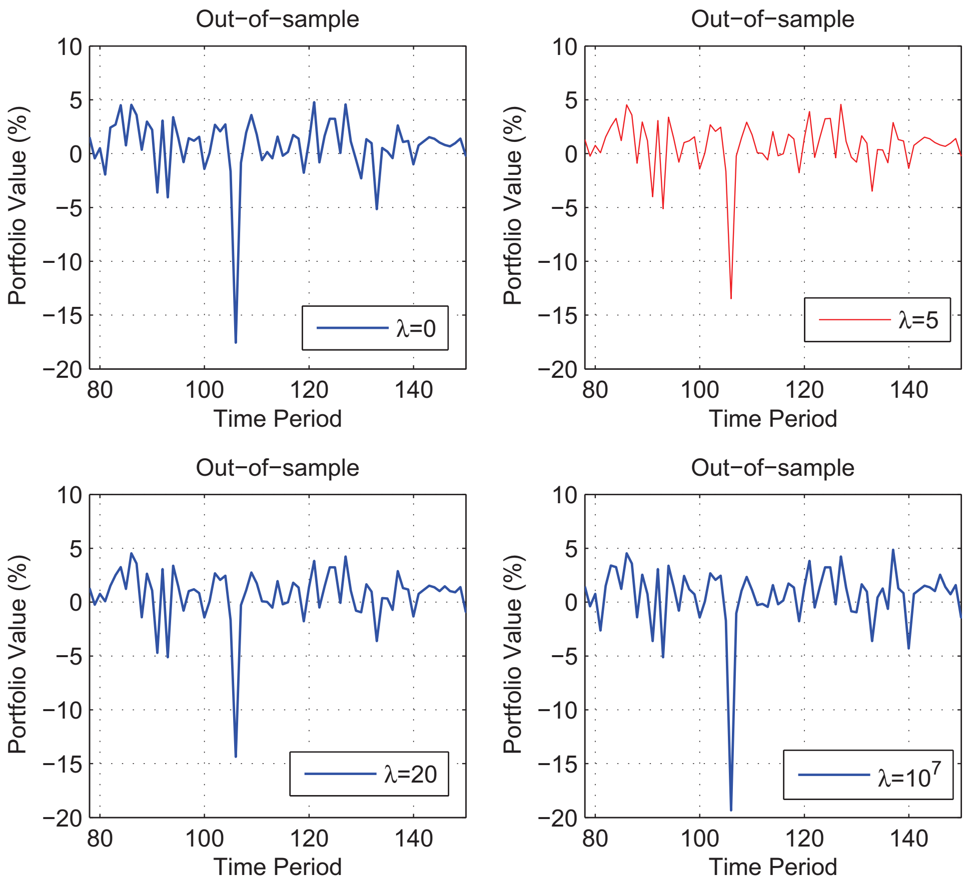


Figure 4.2: *Evolution of realized return related to the norm-portfolio models for different λ values in the second subperiod*

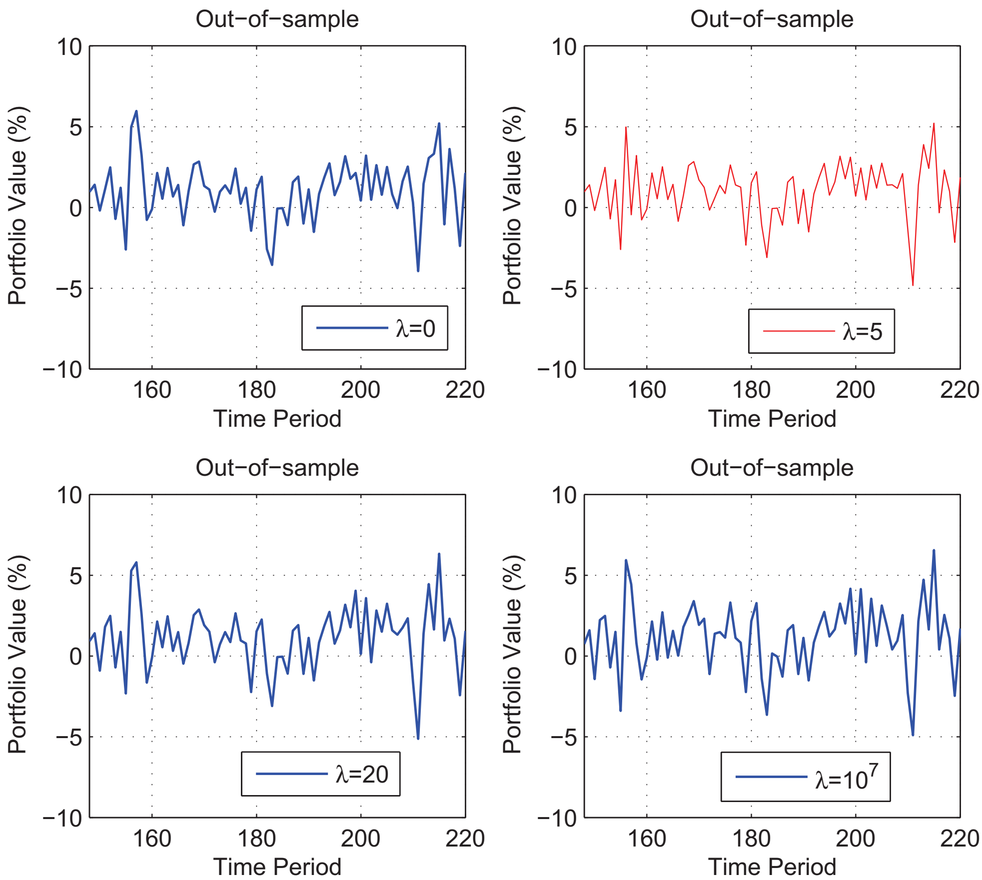
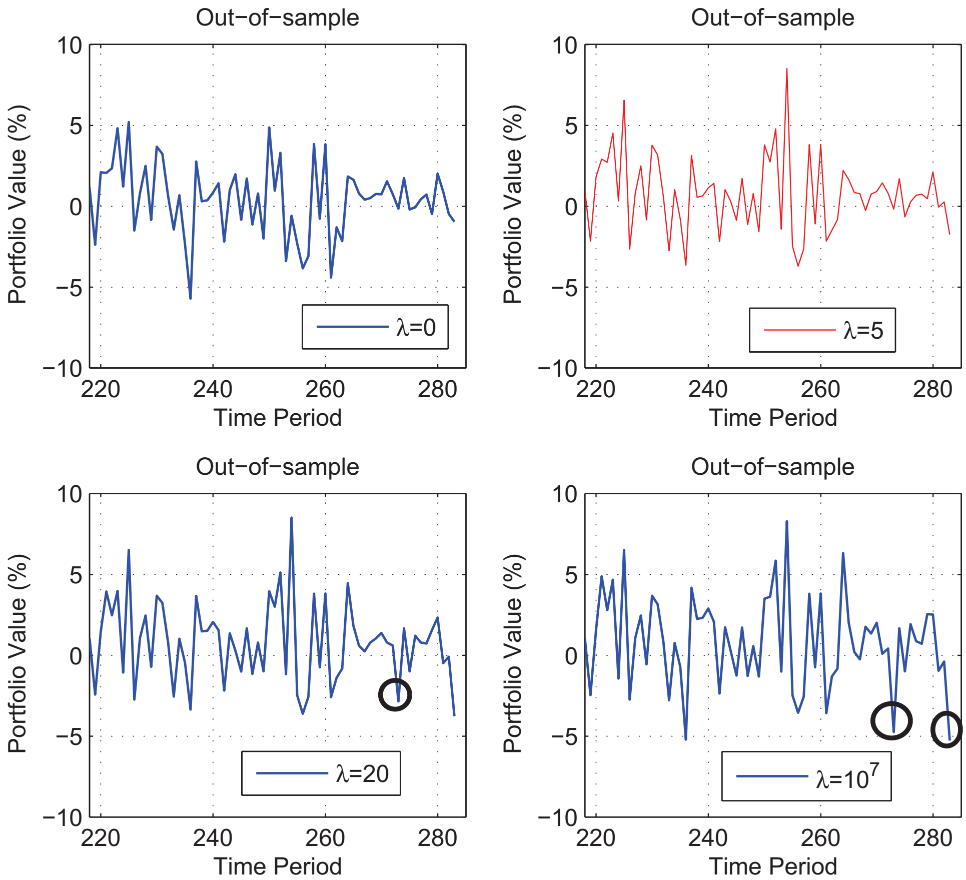


Figure 4.3: *Evolution of realized return related to the norm-portfolio models for different λ values in the third subperiod*

Figure 4.4: *Evolution of realized return related to the norm-portfolio models for different* λ *values in the fourth subperiod*



This straightforward observation is confirmed by the statistics reported in the table below:

Norm-portfolio models	**Mean** ($\hat{\mu}$)	**Variance** ($\hat{\sigma}^2$)	**Sharpe Ratio** ($\hat{SR}$)	**Turnover**
$\lambda = 0$	0.8951	6.8783	0.3413	0.3649
$\lambda = 5$	0.9582	5.8995	0.3945	0.3440
$\lambda = 20$	1.0501	10.4485	0.3249	0.3854
$\lambda = 10^7$	1.2134	29.6181	0.2230	0.4831

Table 4.1: *Out-of-sample mean, variance, Sharpe Ratio and portfolio turnover.*

Table 4.1 reports the out-of-sample mean, variance, Sharpe Ratio and portfolio turnover related to the *norm-portfolio* models for different values of parameter λ.

Our first observation is that the out-of-sample variance degrades significantly with the increase in the out-of-sample mean; indeed, variance increases from approximately 6 to approximately 30 (or in percentages from 2.4% to 5.4%, by calculating the square root of variance), whereas mean increases from 0.9 to 1.21 approximately.

The second observation is that no model performs better than others among all four the out-of-sample criteria (mean, variance, Sharpe Ratio and turnover). For example, the case $\lambda = 10^7$ shows the best performance in terms of the out-of-sample mean, but the worst one in terms of the out-of-sample variance; instead the case $\lambda = 5$ performs better in terms of the out-of-sample variance, Sharpe Ratio and portfolio turnover.

As it has been said, from a theoretical point of view when $\lambda = 0$ the investor considers all probabilities q and not only the nearest to the reference probability measure p. Although this choice provides the lowest value of the out-of-sample mean, it reacts in a better way in terms of the out-of-sample variance, producing a low level of variability; in addition, in terms of the last two statistics, it produces good results even if not the best one. In other words, in this case the investor tends to overprotect himself.

Increasing the values of parameter λ, the investor gives less weight to probabilities that are far away from the reference probability, so he focuses on a subset of probabilities. This choice produces positive results but only for values of λ not too high; in fact, the case $\lambda = 5$, as it has been noted, is the best one in terms of the out-of-sample variance, Sharpe Ratio and portfolio turnover.

Increasing farther on the value of λ improves only the out-of-sample mean, showing instead the worst results in terms of the remaining statistics.

In accordance to that, for small values of λ the portfolio's realized values show a more regular behaviour, at least in this first set of experiments. On the other hand, increasing the value of parameter λ, the models result less conservative.

Let us now plot the coherent model, i.e., the variant of the *norm-portfolio* family where the risk measure verifies also the property of positive homogeneity as we have described in Section 3.3. Figures 4.5-4.8 depict the behaviour of the *coherent* variant for the different values of the chosen parameter π. From a theoretical point of view, the parameter π measures the distance between the generic probability q and the reference one p. Let us observe that a value of π equal to zero describes the case in which the distance between the probabilities is null; in other words, the investor considers only the case $p = q$. Increasing π, the investor considers all probabilities measures q such that the absolute value $|p_i - q_i|$ is at most π for all i. The case $\pi = 0$ is theoretically the less conservative case.

The values of parameter π that we represent in figures 4.5-4.8 are the following: $\pi = 0.25$, $\pi = 0.5$, $\pi = 1$ and $\pi = \pi^*$, where π^* is calculated like in (4.6); by rounding up the theoretical value, we obtain $\pi^* = 0.92$. Moreover, we exclude the trivial case $\pi = 0$ because it does not fit a realistic behaviour of the investor.

Also in this case, we divide the whole time period in four subperiods to provide a clear representation of the model.

Table 4.2 provides the result of statistics for the *coherent* variant.

Coherent variant	**Mean** ($\hat{\mu}$)	**Variance** ($\hat{\sigma}^2$)	**Sharpe Ratio** ($\hat{SR}$)	**Turnover**
$\pi = 0.25$	0.9835	6.1545	0.3964	0.3207
$\pi = 0.5$	0.9240	6.2106	0.3708	0.3384
$\pi = \pi^* = 0.92$	0.8951	6.8783	0.3413	0.3649
$\pi = 1$	0.8951	6.8783	0.3413	0.3649

Table 4.2: *Out-of-sample mean, variance, Sharpe Ratio and portfolio turnover.*

Let us note that increasing π, the values of corresponding statistics do not vary significantly. Different from the *norm-portfolio* models just analyzed, the empirical behaviour of their *coherent* variant does not seem to confirm the theoretical expectation in terms of robustness. This would seem to reveal

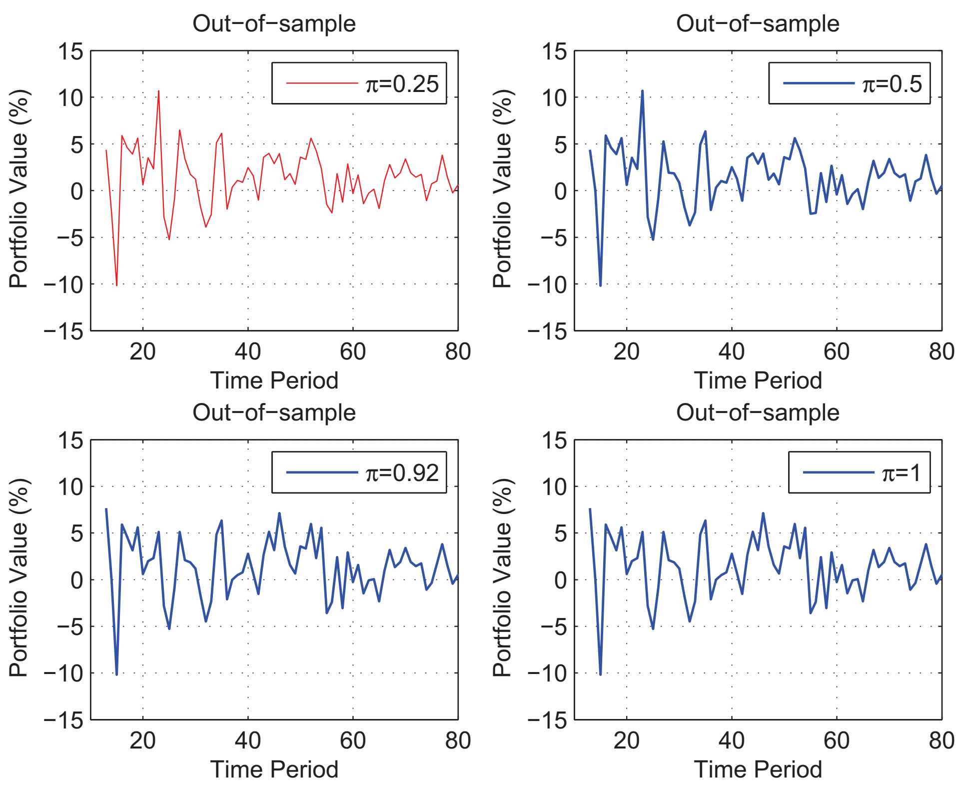


Figure 4.5: *Evolution of realized return related to coherent variant of the norm-portfolio models for different* π *values in the first subperiod*

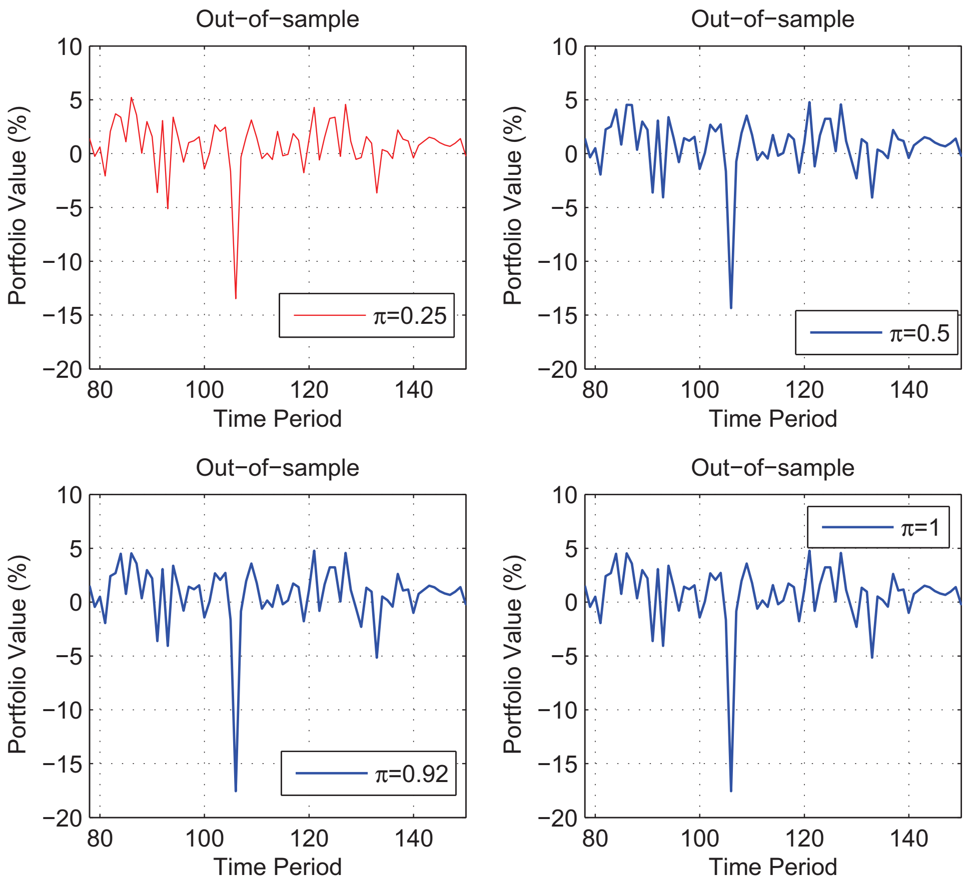


Figure 4.6: *Evolution of realized return related to coherent variant of the norm-portfolio models for different π values in the second subperiod*

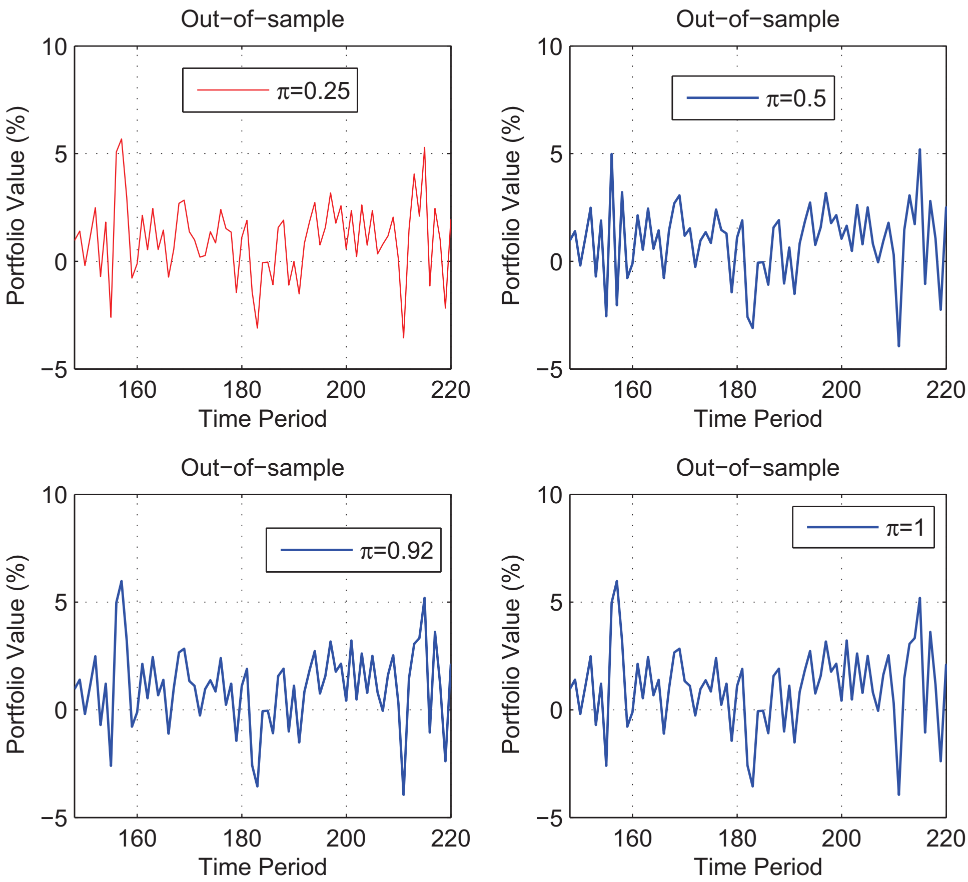


Figure 4.7: *Evolution of realized return related to coherent variant of the norm-portfolio models for different π values in the third subperiod*

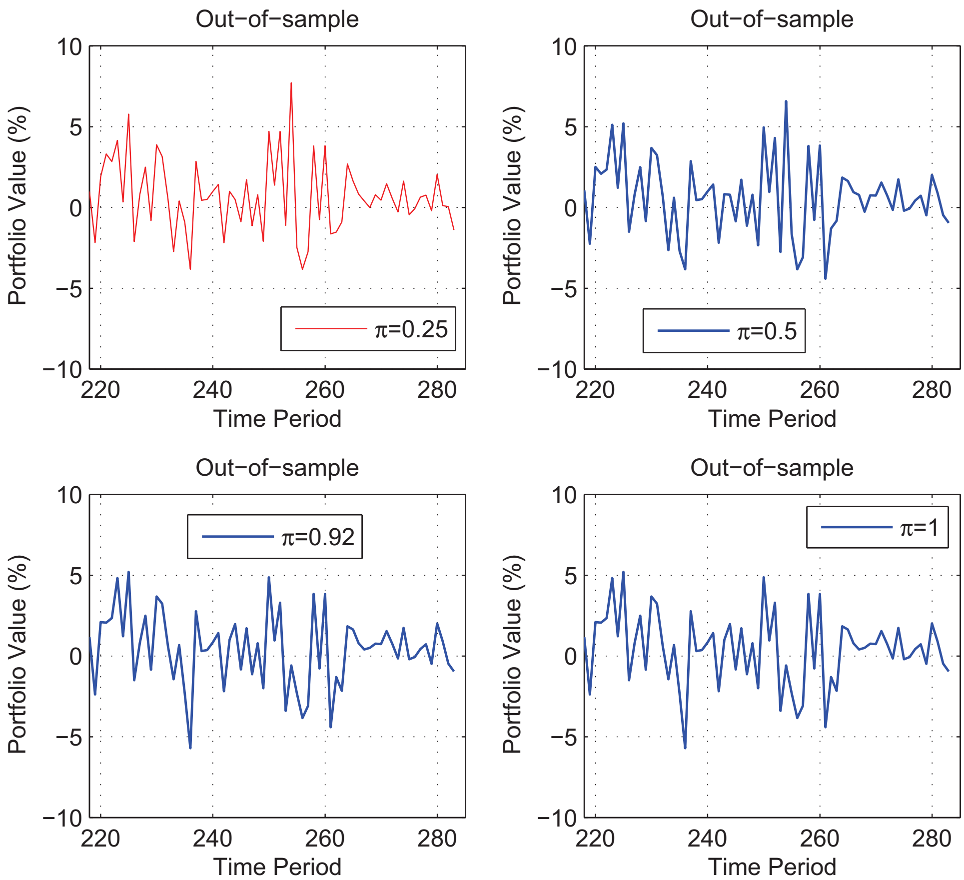


Figure 4.8: *Evolution of realized return related to coherent variant of the norm-portfolio models for different π values in the fourth subperiod*

that, from a practical point of view, assigning the same weight to all the considered probabilities might influence the regularity of the robust optimal portfolio value.

From the experiment it results that among all four statistics, the best behaviour is described by the case $\pi = 0.25$ in which one has the highest out-of-sample mean and Sharpe Ratio and the lowest out-of-sample variance and turnover.

Then, let us note that the cases $\pi = \pi^*$ and $\pi = 1$ are identical each others.

Figures 4.9-4.11 depict the behaviour of the CVaR, the *Tütüncü-Koenig* and the *entropic* models respectively. To solve CVaR model, as it has been anticipated in the previous sections, we model random event by a finite set of scenarios, obtaining a linear programming problem like in [70]. For $\pi = \pi^*$, let us note that the *coherent* variant is not equivalent to the CVaR model in terms of the optimal portfolio value; the slight difference might depend on the rounding error generated in calculating the empirical value of π^* (so, the empirical value of π^* is just an approximation of the theoretical value (3.31) in the subsection 3.3.1).

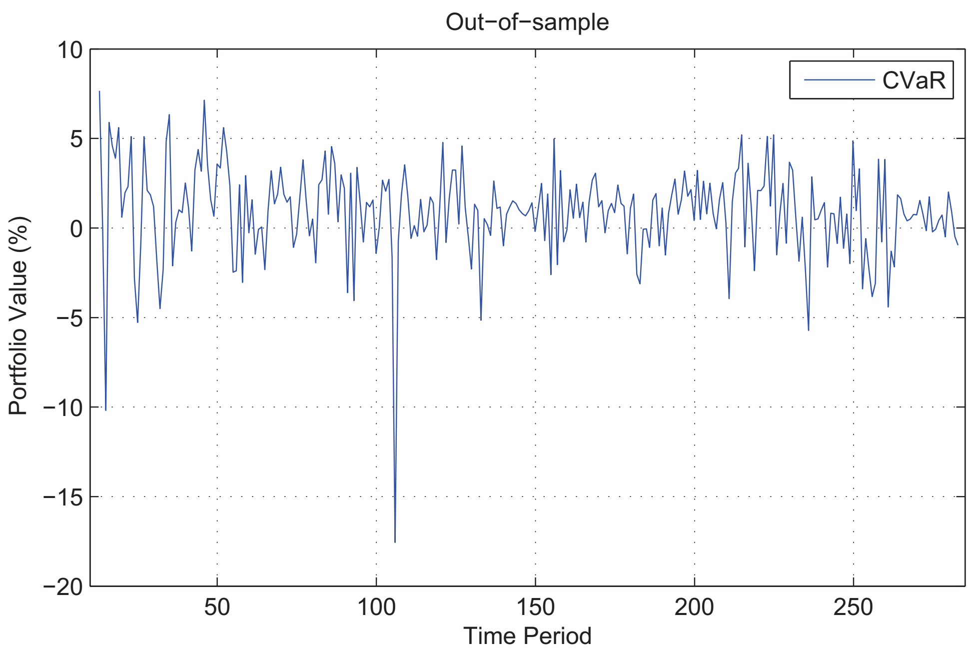


Figure 4.9: *Out-of-sample performance for the CVaR model*

As it has been said, the *Tütüncü-Koenig* model we use in these experiments is a slight variant of the model described in [79] where the robustness appears only at the covariance matrix level. Like for the previous models, we work on annual subintervals and solve problem (4.3) for each of them. At each iteration, the procedure that generates the upper bound Q^U of the covariance matrix Q is described in Section 4.1.1; let us note that working on annual subintervals we choose moving windows of sixt months in order to generate Q^U.

Figure 4.10 is a plot of the performance of the *Tütüncü-Koenig* model which incorporates a standard robustness type. Notice that at $t = 15$ as well as at $t = 106$ there is a precipitous drop in the relative performance of the standard robust model. These drops are slightly higher (in modulus) with respect to drops achieved in the previous models.

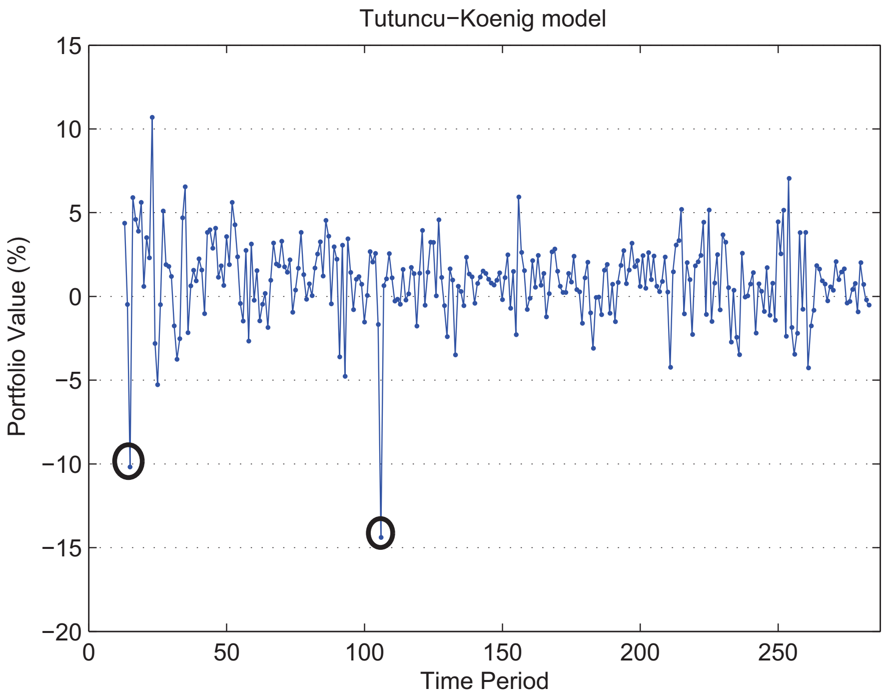


Figure 4.10: *Performance of the Tütüncü-Koenig model*

Figure 4.11 depicts the trajectory of the *entropic* soft robust model described in subsection 4.1.3. Chosen the parameter $\delta = log\left(\frac{1}{\alpha}\right)$ with α the confidence level of CVaR model, we effect a binary search like described in subsection 4.1.1 in order to find an approximate value of variable c that minimizes the objective function; then, we set c and compute the optimal portfolio x and its corresponding value in terms of return.

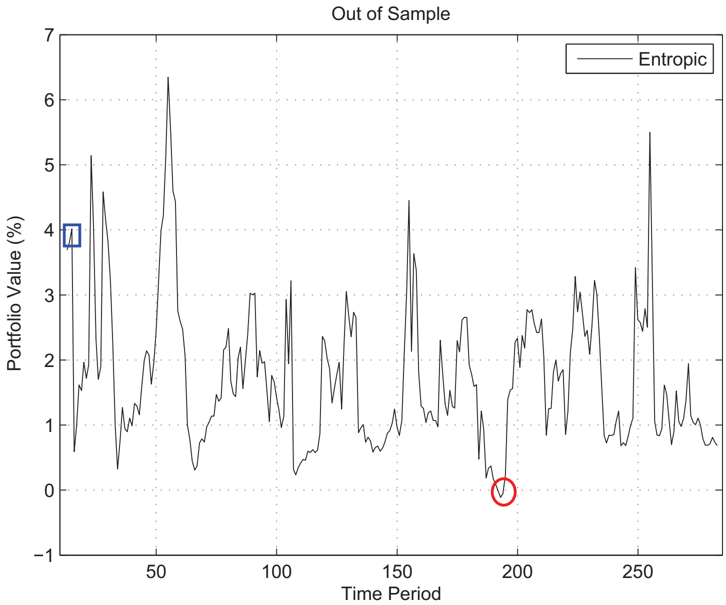


Figure 4.11: *Performance of the entropic model*

Let us observe that with respect to the previous models the *entropic* one never attains severe drops, but rather it always reaches a positive value of the portfolio (there are only two exceptions corresponding to the cases $t = 193$ and $t = 194$ as signed in the figure 4.11 with a red circle). This major regularity is confirmed also by other cases like at $t = 15$ where all models until now analyzed attain severe drops, whereas *entropic* model states up the zero (exactly around the value 4%) signed in the figure with a blue rectangle.

As next step we compare the performance of all considered models to evaluate how the different concepts of robustness influence the optimal portfolio value and its regularity in the long term. First of all, we choose the value of the parameters λ and π that produce the best performance of the *norm-portfolio* models and the *coherent* variant respectively.

The "best value" chosen from Table 4.1 is $\lambda = 5$; although it does not produce the higher mean of returns, however it results the best in terms of variance, Sharpe Ratio and portfolio turnover; in analogous way, from Table 4.2, we choose the case $\pi = 0.25$ that performs better among all four criteria.

Table 4.3 gives the out-of-sample results for all considered models (and hence for all robustness types).

Let us observe that Table 4.3 above reports the values of the statistics

Models	**Mean** ($\hat{\mu}$)	**Variance** ($\hat{\sigma}^2$)	**Sharpe Ratio** ($\hat{SR}$)	**Turnover**	**Time**
Norm-portfolio ($\lambda = 5$)	0.9582	5.8995	0.3945	0.3440	0.0235
Coherent variant ($\pi = 0.25$)	0.9835	6.1545	0.3964	0.3207	0.0626
CVaR	0.8797	6.7174	0.3394	0.3415	0.1953
T-K	0.9358	7.8912	0.3331	0.3116	0.1328
Entropic	1.7223	1.2034	1.5700	0.4495	91.1330

Table 4.3: *Out-of-sample mean, variance, Sharpe Ratio, portfolio turnover and computational time for all models chosen.*

for each chosen model. In addition to the statistics defined in Subsection 4.1.3, we report also the computational time, i.e., the average cpu-time (in seconds) needed to produce the optimal solution of each problem.

Observing the cpu-time shown in Table 4.3, it is quite evident that the entropic model is computationally expensive; indeed, it requires a computational time almost $4 \cdot 10^3$ times larger than all other models. The main source of time complexity could be the nature of the problem that is a convex programming problem solved with cvx software using a successive approximation method.

About the other statistics, the following considerations could be drawn:

- Between a traditional robustness incorporated by *Tütüncü-Koenig* model and a more flexible robust approach (described through the *norm-portfolio*, its *coherent* variant and the *entropic* models) let us note that the second approach always outperforms the first one among almost all considered criteria (exactly mean, variance and Sharpe Ratio). In the other words, by relaxing the robustness constraints in a flexible way, one can potentially gain out-of-sample performance for not too high of a price.

- Without doubts, the *entropic model* results the best approach in terms of the out-of-sample mean, variance and Sharpe Ratio even if it results the most expensive in terms of variability of the assets in the portfolio and hence, in terms of the transaction costs as well as in terms of computational cost as it has been said above.

- The best approach in terms of turnover is the traditional one described by *Tütüncü-Koenig* model. Anyway, it was to be expected that relaxing robust constraint it could have required higher costs, but the supported major costs have been compensated for the major mean return and Sharpe Ratio value and for the minor variance value.

- Finally, let us observe that the *norm-portfolio* model and its *coherent* variant show a quite similar behaviour with a slight better performance of the coherent one in terms of the out-of-sample mean, Sharpe Ratio and turnover; the lowest variance that marks the *norm-portfolio* model is not enough to provide a better performance of it. In addition, the *coherent* variant results the least expensive immediately after the *Tütüncü-Koenig* model.

From this first computational test, it results that a flexible robust approach can potentially outperforms a traditional one.

4.3 The second computational test

For the second experiment we used a wider set of market indices [79]: growth and value stocks in large-cap, mid and small-cap categories, intermediate term fixed-income securities, international stocks, real estate securities and high-yield corporate bonds. To represent each category we used Wilshire Target indices, Lehman Brothers Intermediate Government/Credit index, MSCI EAFE (Europe, Australasia, Far East) index, Wilshire Real Estate Securities index and Lehman Brothers High-Yield Bond index. Time series data covers the period July 1983-July 2002 for a total of N=229 months.

In this second experiment, we use the same procedure described in the first computational test. Once fixed all other parameters, we solve the *norm-portfolio* model for each integer value of λ chosen into a specified interval (also in this case $[0, 120]$). The selected values of λ are the following: $\lambda = 0$, $\lambda = 5$, $\lambda = 10$, $\lambda = 15$, $\lambda = 20$, $\lambda = 25$, $\lambda = 30$, $\lambda = 35$ and $\lambda = 10^7$. About the π value, the same considerations of the first experiment hold. We choose the following values: $\pi = 0$, $\pi = 0.25$, $\pi = 0.5$, $\pi = \pi^*$ and $\pi = 1$, where as before π^* is a rounding value of the theoretical one calculated like in (4.6).

Let us begin to present the results of the *out-of-sample* analysis for this second data set.

Figures 4.12-4.14 plot the time evolution of the relative performance of the *norm-portfolio* models for various values of λ; in particular for the following ones: $\lambda = 0$, $\lambda = 5$, $\lambda = 20$ and $\lambda = 10^7$.

In order to show a clear performance of the models, we subdivide the whole time period in three subperiods. Also in this case, it is verified an analogous behaviour in terms of regularity when λ increases: the models show high peaks and severe drops. For example, the performance described by the *norm-portfolio* models attains one of the highest peaks at $t = 102$ (Figure 4.13, signed with a circle) and falls headlong at $t = 182$ as well as at $t = 229$ (Figure 4.14, signed with circles). Let us observe that all these three extreme cases correspond to values of parameter λ very high, confirming the role of this parameter in terms of robustness.

Figure 4.12: *Second data set: evolution of realized return related to the norm-portfolio models for different values of λ in the first subperiod.*

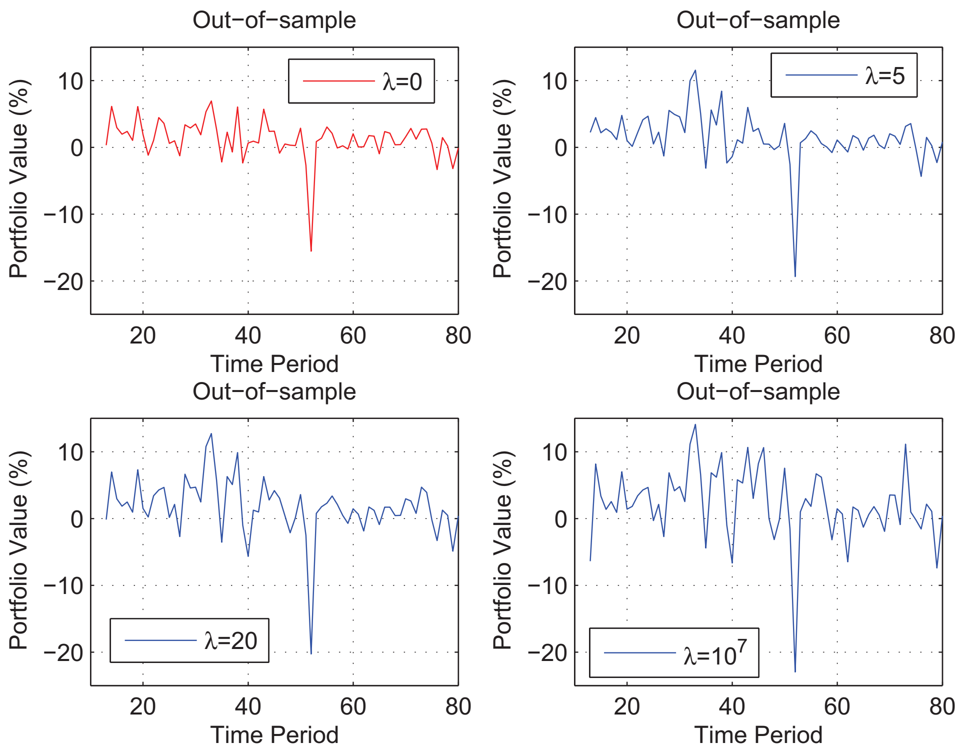


Figure 4.13: *Second data set: evolution of realized return related to the norm-portfolio models for different values of λ in the second subperiod.*

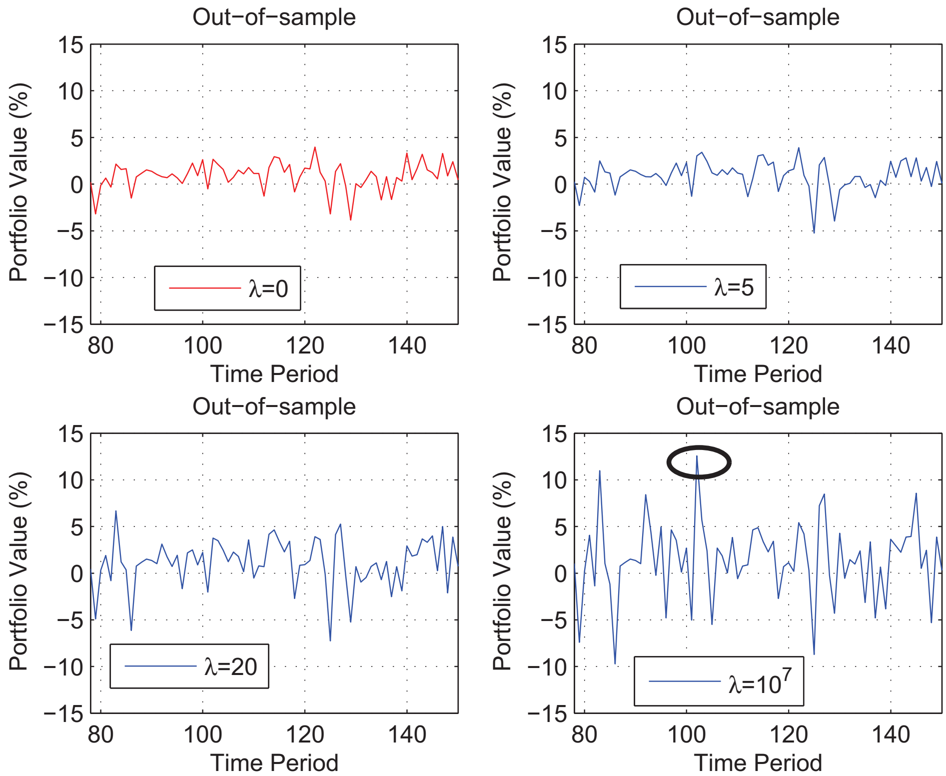
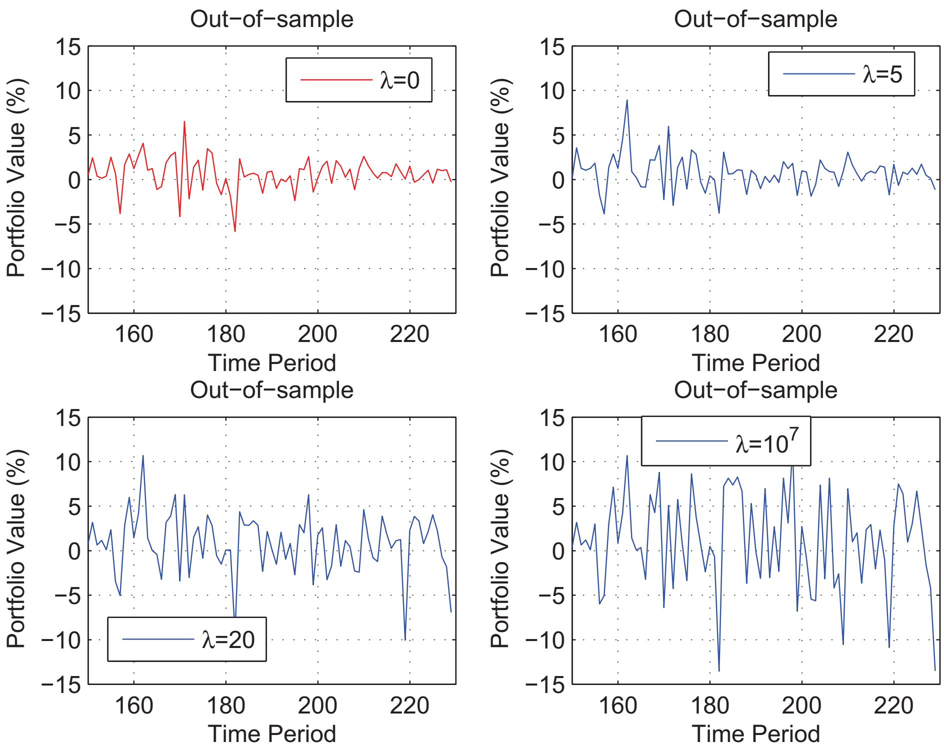


Figure 4.14: *Second data set: evolution of realized return related to the norm-portfolio models for different values of* λ *in the third subperiod.*



Let us now report in Table 4.4 the out-of-sample results of the *norm-portfolio* models in terms of the three criteria chosen (out-of-sample mean, variance, Sharpe Ratio and portfolio turnover).

Norm-portfolio models	**Mean** ($\hat{\mu}$)	**Variance** ($\hat{\sigma}^2$)	**Sharpe Ratio** ($\hat{SR}$)	**Turnover**
$\lambda = 0$	0.9250	4.6469	0.4291	0.4028
$\lambda = 5$	1.0492	6.6965	0.4054	0.4351
$\lambda = 20$	1.2324	12.0267	0.3554	0.5136
$\lambda = 10^7$	1.5798	25.3376	0.3138	0.4815

Table 4.4: *Out-of-sample mean, variance, Sharpe Ratio and portfolio turnover.*

As just shown in the relative figures, let us observe that by increasing the values of the parameter λ, the out-of-sample variance degrades significantly with the increase of the out-of-sample mean (as in the first experiment). The case $\lambda = 10^7$ shows the worst result in terms of the variance and the best one in terms of the mean; the opposite extreme case, i.e., $\lambda = 0$, shows the best relation risk-return providing the highest value of the Sharpe Ratio and yielding the lowest cost evaluated in terms of portfolio turnover. In other words, in this second experiment, an approach more conservative provides the best results among almost all adopted criteria.

If one increases little more the value of λ (case $\lambda = 5$) however it is possible to obtain higher mean return at not too high of a price (notice that the difference in terms of turnover between the case $\lambda = 0$ and the case $\lambda = 5$ is not too high).

Figures 4.15-4.17 depict the performance of the coherent variant. Again, we subdivide the whole time period in three subperiods and for each of them we report the following cases: $\pi = 0.25$, $\pi = 0.5$, $\pi = \pi^*$ and $\pi = 1$. Let us observe that last two cases ($\pi = \pi^*$ and $\pi = 1$) are identical each others as confirmed by the results of statistics reported in Table 4.5.

Furthermore, notice that increasing the value of parameter π, the out-of-sample variance tends to decrease, even if not in a significant way; in other words there is not a big difference among values. However, different from the first computational test, the theoretical expectation in terms of robustness is here confirmed, by increasing the value of π in fact the model becomes more conservative (as shown by corresponding Sharpe Ratios). In addition, the models with highest value of π are also the least expensive in terms of portfolio turnover and hence in terms of the transaction costs.

Coherent variant	Mean ($\hat{\mu}$)	Variance ($\hat{\sigma}^2$)	Sharpe Ratio ($\hat{SR}$)	Turnover
$\pi = 0.25$	0.9745	5.5252	0.4146	0.4632
$\pi = 0.5$	0.8910	4.4600	0.4219	0.4360
$\pi = \pi^* = 0.92$	0.9250	4.6469	0.4291	0.4328
$\pi = 1$	0.9250	4.6469	0.4291	0.4328

Table 4.5: *Out-of-sample mean, variance, Sharpe Ratio and portfolio turnover.*

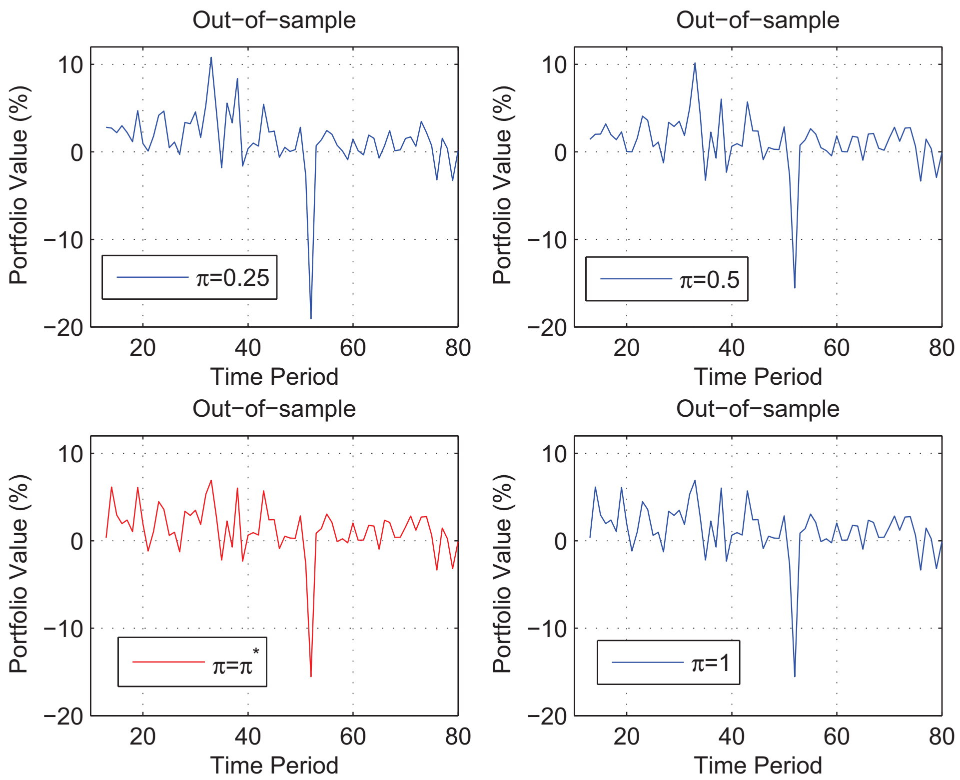


Figure 4.15: *Evolution of realized return related to Coherent variant of the norm-portfolio models for different π values in the first subperiod*

Figure 4.16: *Evolution of realized return related to Coherent variant of the norm-portfolio models for different π values in the second subperiod*

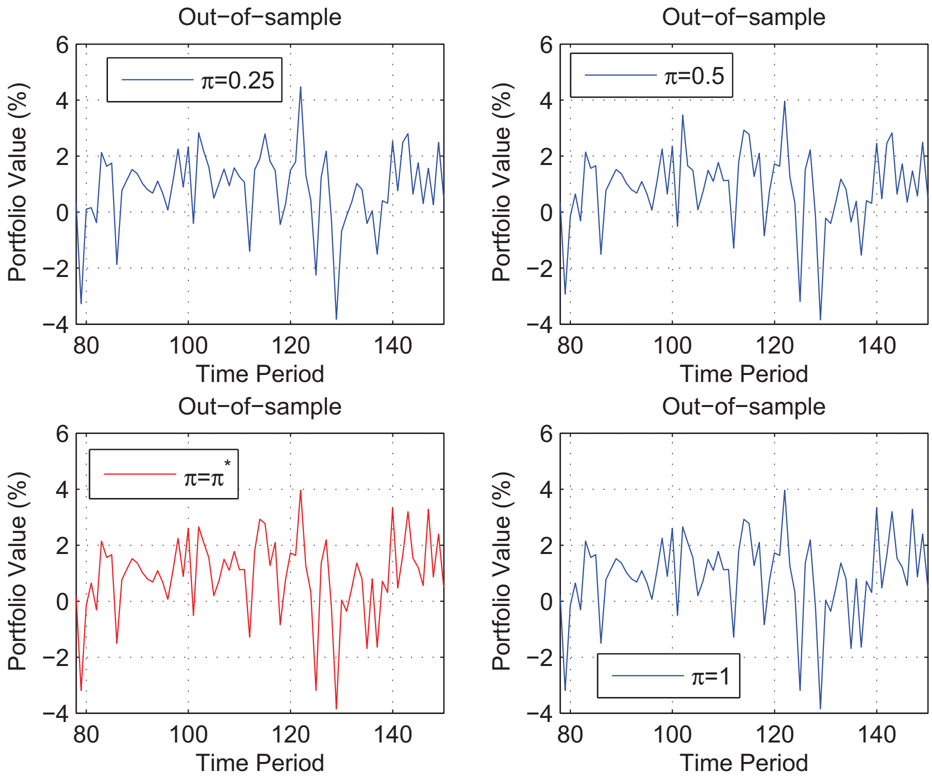
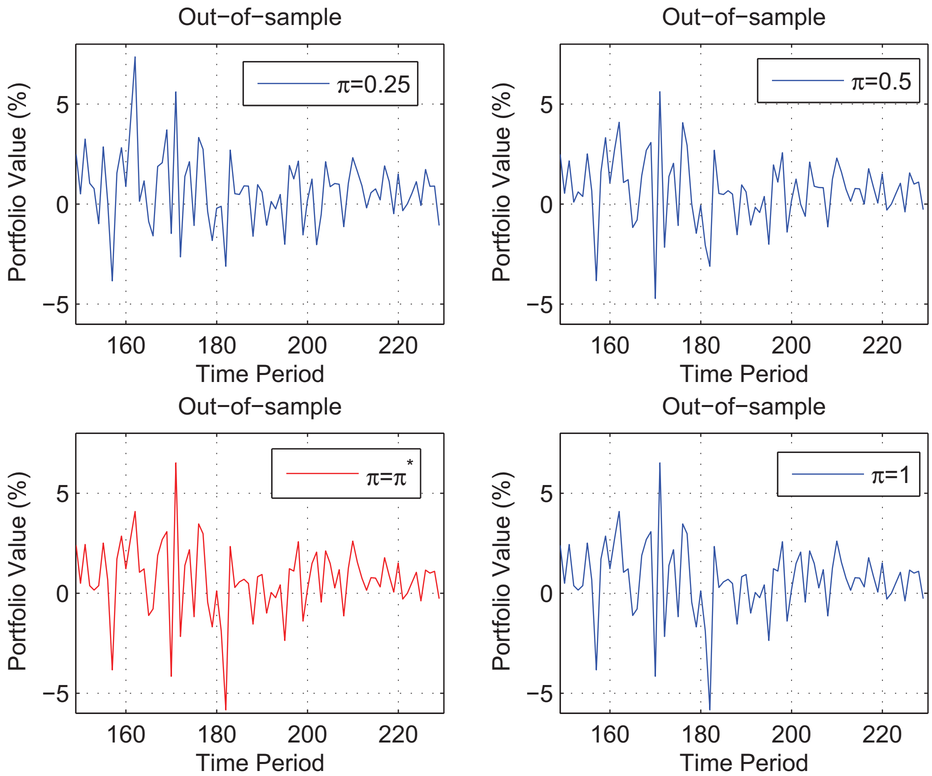


Figure 4.17: *Evolution of realized return related to Coherent variant of the norm-portfolio models for different π values in the third subperiod*



Figures 4.18-4.20 depict the performance (in terms of the portfolio value) of the CVaR, the *Tütüncü-Koenig* and the *entropic* models respectively. The behaviour of the CVaR model and of the *coherent* variant are quite similar, even if there is no a complete overlapping of them (considering the case $\pi = \pi^*$). The slight difference could depend on the rounding errors between the theoretical value of π and the empirical one.

The *Tütüncü-Koenig* performance (figure 4.19) appears quite regular by considering the values of the portfolio, even if the gap (in terms of return's values) between a time period and the following one is often high.

In addition, the absolute value of drops attained in this model is very lower than this one generated by the others models; for example, let us observe that at $t = 52$ the performance of *Tütüncü-Koenig* model shows the worst drop, but its absolute value (about 8%) is lower than the value produced by the other analyzed models (about 15%).

Figure 4.18: *Out-of-sample behaviour of CVaR model.*

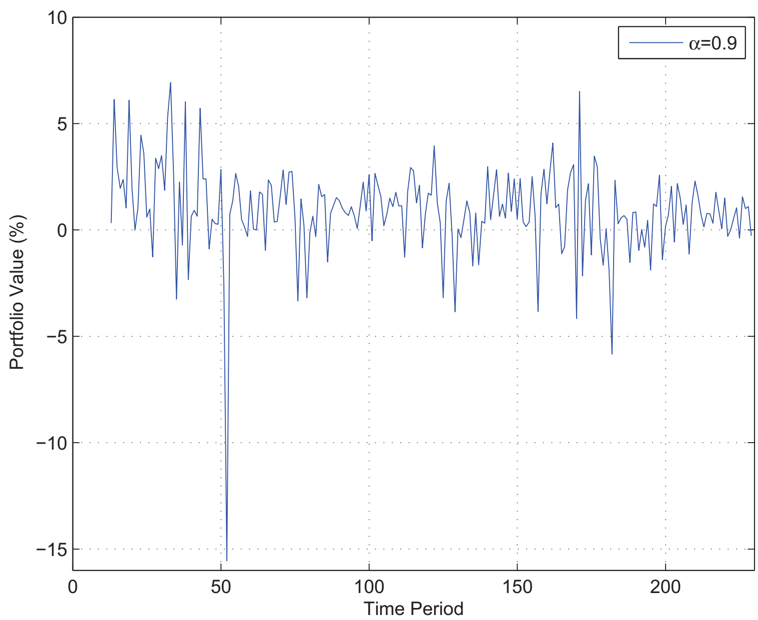
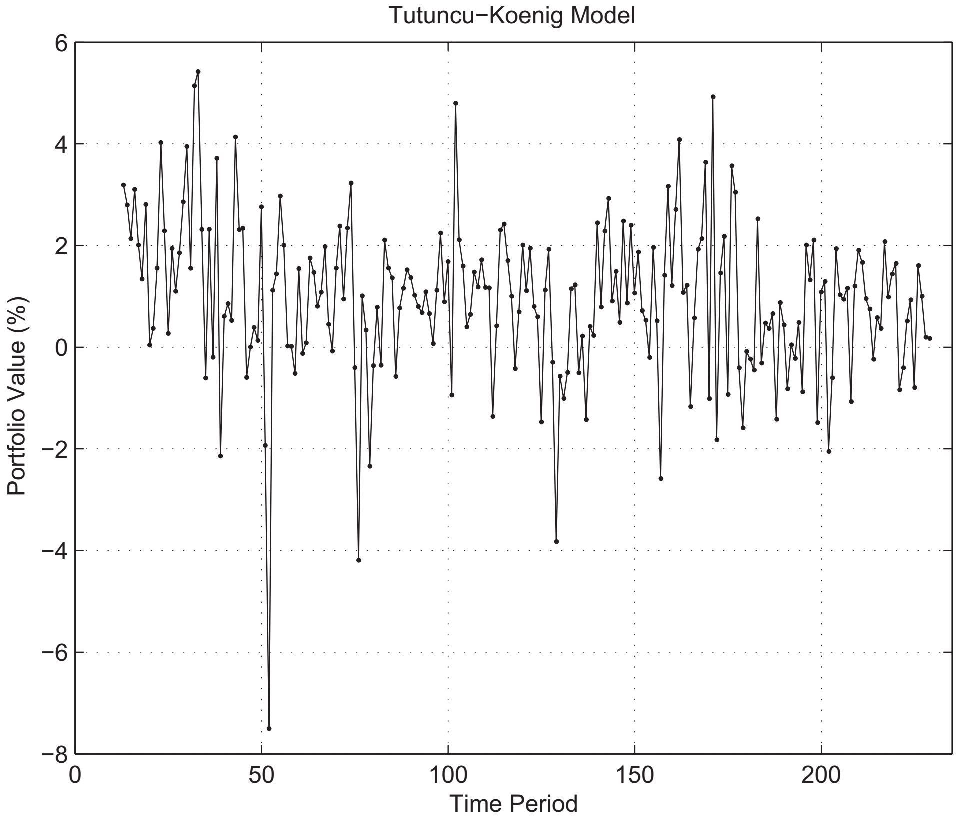


Figure 4.19: *Out-of-sample behaviour of Tütüncü-Koenig model.*



The *entropic* approach (figure 4.20) produces the most regular performance among all studied models; the value of the portfolio very often states up to zero. In addition, sometimes the trend of the *entropic* model seems to go against the stream: for example at $t = 52$ all described models attain a very low value of the portfolio, whereas the *entropic* model reaches a positive value of the portfolio (between 3% and 4% with respect to initial portfolio value). At $t = 140$ an opposite case is verified: the *entropic* model attains its lowest value of the portfolio, whereas all other models accomplish values between 2.5% and 3.5%.

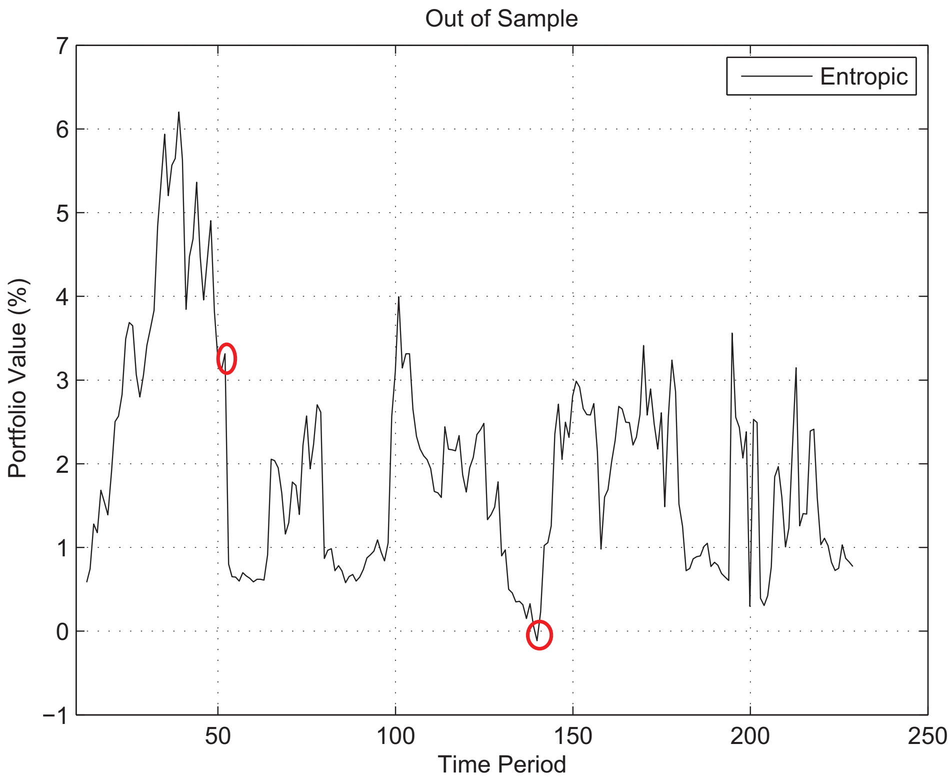


Figure 4.20: *Out-of-sample behaviour of the entropic model*

Table 4.6 reports the complete results of the out-of-sample analysis for all models (among the indicators, also in this case we introduce the computational time).

Models	**Mean** ($\hat{\mu}$)	**Variance** ($\hat{\sigma}^2$)	**Sharpe Ratio** ($\hat{SR}$)	**Turnover**	**Time**
Norm-portfolio ($\lambda = 0$)	0.9250	4.6469	0.4291	0.4028	0.0469
Coherent variant ($\pi^* = 0.92$)	0.9250	4.6469	0.4291	0.4328	0.0156
CVaR	0.9085	4.6196	0.4227	0.4621	0.0156
T-K	0.9317	2.5412	0.5844	0.3924	0.1875
Entropic	1.9583	1.6420	1.5282	0.4827	106.9952

Table 4.6: *Out-of-sample mean, variance, Sharpe Ratio, portfolio turnover and computational time.*

The "best performance" of the *norm-portfolio* models and their *coherent* variant is represented by the cases $\lambda = 0$ and $\pi = \pi^*$ respectively as reported in Table 4.6.

From this second computational test the following observations are drawn:

- The *entropic* model produces again the best performance in terms of the out-of-sample mean, variance and Sharpe Ratio but, as in the first experiment, it requests the highest cost in terms of portfolio turnover and computational time. Let us note that also in this case the computational time needed to solve the *entropic* problem is very large, indeed it is more than $2 \cdot 10^3$ times larger than the time needed to solve all other models.

- After the *entropic* model, the best one results the *Tütüncü-Koenig* model. Indeed, it shows the best performance in terms of the out-of-sample variance, Sharpe Ratio and portfolio turnover; whereas, in terms of the out-of-sample mean it is overcame also by the *norm-portfolio* model.

- The *norm-portfolio* model, its coherent variant and CVaR are quite similar in terms of almost all adopted criteria except for portfolio turnover; let us note that the *norm-portfolio* model shows the best performance in terms of the portfolio turnover (in fact, the value of this statistic is almost close to the result of *Tütüncü-Koenig* model that, as it has been above said, is the least expensive); whereas, CVaR model results the most expensive approach immediately after the entropic one.

In general we can conclude that the standard robustness incorporated by the *Tütüncü-Koenig* model shows a better performance than relaxed one incorporated by the *norm-portfolio* model and the *coherent* variant; but with respect to the *entropic* model, it produces a worse behaviour.

4.4 The third computational test

The data used in this last experiment are used by Byrne in [18]. They represent the total monthly returns for 10 market segment indices: Standard Retail Southeast (SRSE), Standard Retail Rest of UK (SRRUK), Shopping Centres (SHC), Retail Warehouse (RW), Offices in the City of London (OCITY), Offices in the West End (OWE), Offices Rest of Southeast (ORSE), Offices Rest of UK (ORUK), Industrials Southern and Eastern (ISE) and Industrials Rest of UK (IRUK) that span the period December 1987 to January 2002 for a total of $N = 181$ monthly returns. As before, in the *in-sample* analysis we choose the values of the parameter λ and the values of π. The chosen λ values are the following: $\lambda = 0$, $\lambda = 5$, $\lambda = 10$, $\lambda = 25$, $\lambda = 50$, $\lambda = 75$, $\lambda = 90$ and, like in the previous experiments, $\lambda = 10^7$. About the parameter π, we use the values of the previous two cases, i.e., $\pi = 0$, $\pi = 0.25$, $\pi = 0.5$, $\pi = 1$ and $\pi = \pi^*$ calculated following the (4.6).

In the *out-of-sample* analysis, we apply the moving windows method and calculate the portfolio value in each subperiod and for each family of described models.

In figures 4.21 - 4.22 we plot the behaviour of the *norm-portfolio* models for some values of λ, i.e.: $\lambda = 0$, $\lambda = 5$, $\lambda = 10$, $\lambda = 25$, $\lambda = 10^7$.

In this last computational test, when λ increases there are not precipitous drops or high peaks, instead there are only minimal variations in the optimal portfolio value. Like example let us note that by increasing the parameter λ, the value of the portfolio at the starting time and at the ending time increases.

Furthermore, in terms of optimal portfolio, the values of positive peaks are almost four times greater than those generated in the first two tests; and in terms of drops, the performance of the *norm-portfolio* models for different λ values does not almost always decrease under the zero threshold.

Figure 4.21: *Behaviour of the norm-portfolio model for different values of* λ *in the first subperiod*

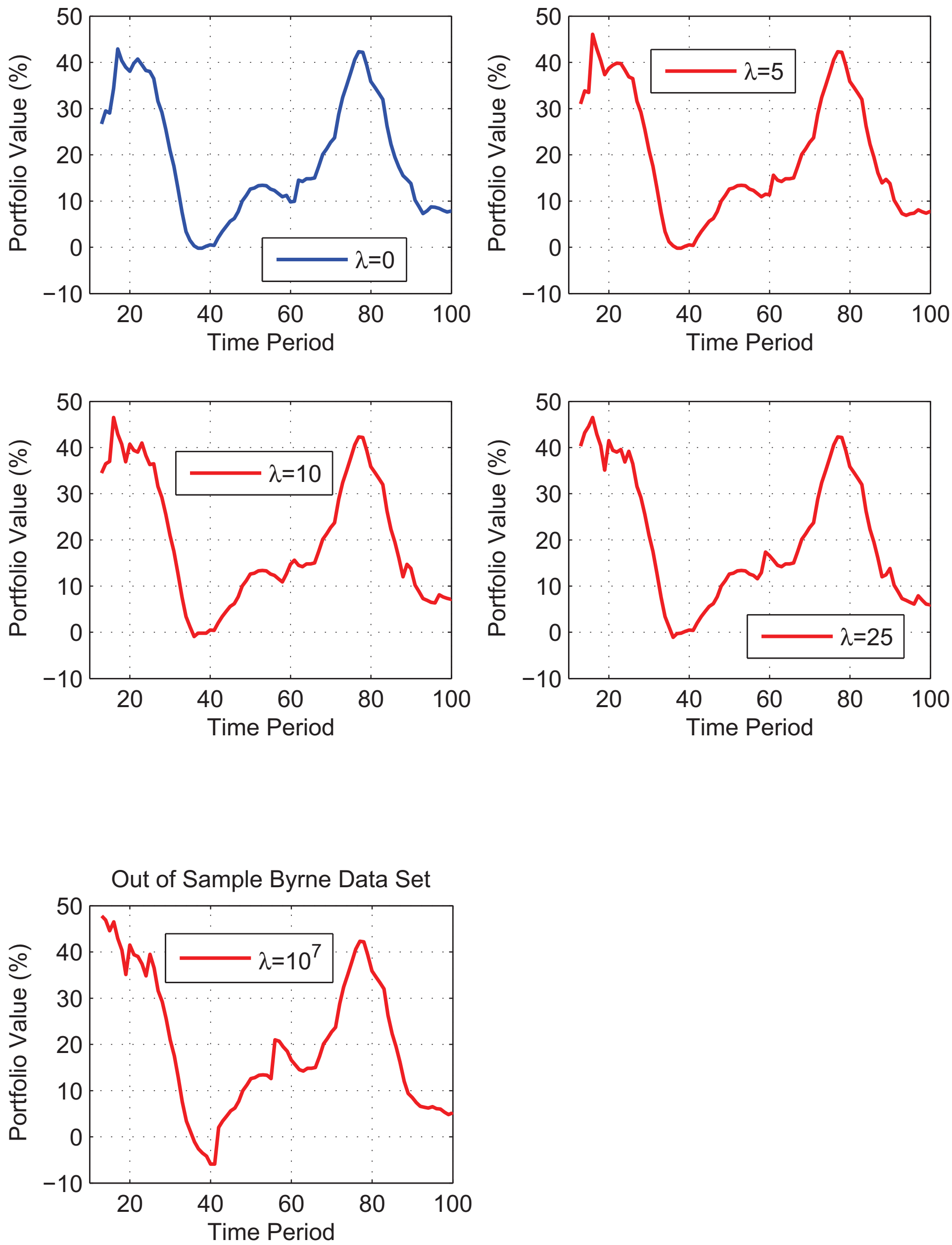
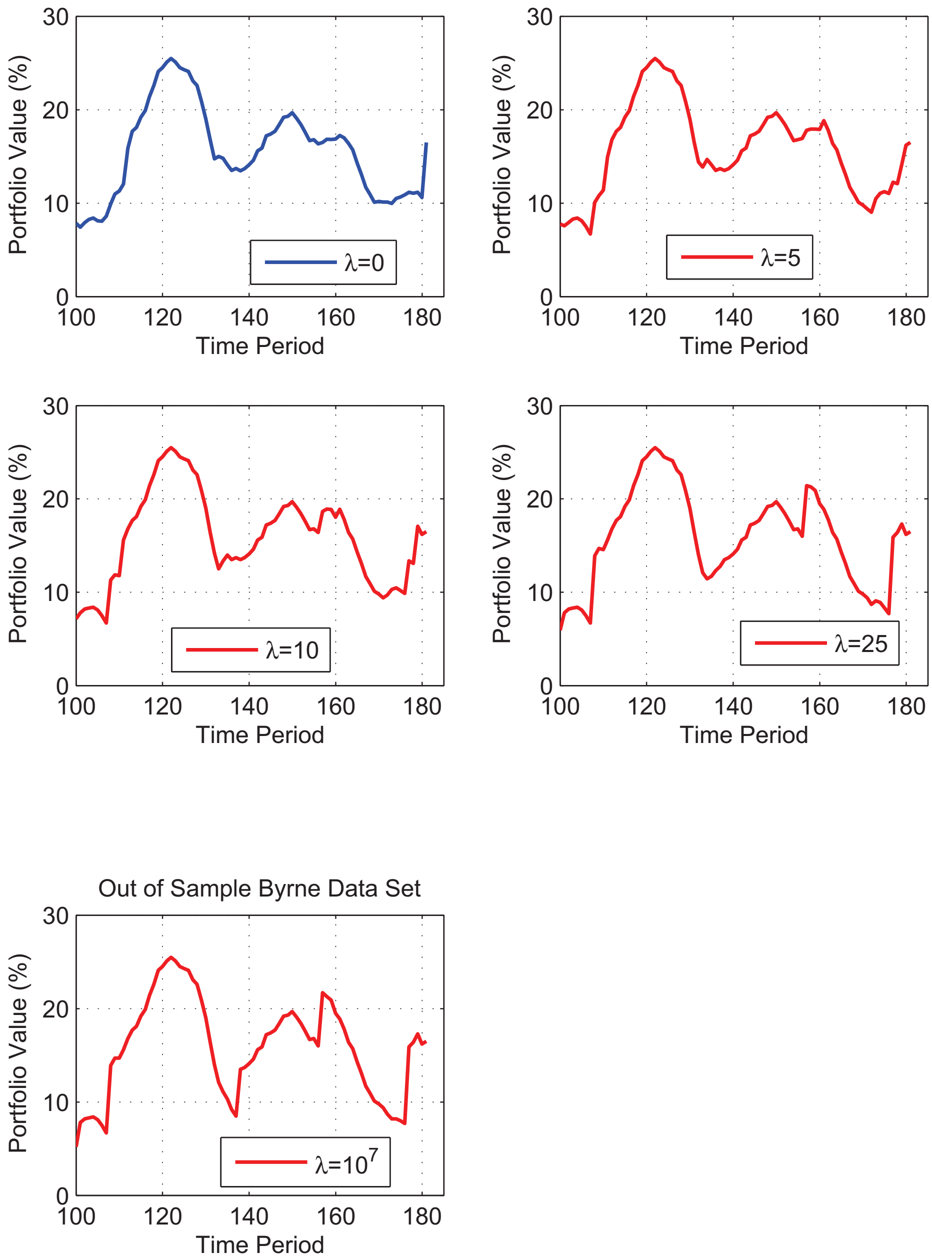


Figure 4.22: *Behaviour of the norm-portfolio model for different values of λ in the second subperiod*



The similar performance in all chosen cases is confirmed by the results of the statistics reported in Table 4.7 below.

Norm-portfolio models	**Mean** ($\hat{\mu}$)	**Variance** ($\hat{\sigma}^2$)	**Sharpe Ratio** ($\hat{SR}$)	**Turnover**
$\lambda = 0$	17.4711	103.1393	1.7203	0.1482
$\lambda = 5$	17.7795	108.2132	1.7091	0.1364
$\lambda = 10$	17.7975	110.275	1.6948	0.1443
$\lambda = 25$	17.972	116.1047	1.6679	0.1424
$\lambda = 10^7$	17.8479	126.9886	1.5838	0.1190

Table 4.7: *Out-of-sample mean, variance, Sharpe Ratio and portfolio turnover.*

The values of the statistics do not change too much increasing the value of λ; let us observe that the out-of-sample variance does not degrade significantly with the increase in the out-of-sample mean, i.e., the out-of-sample variance increases, but not in a significant way (differently from what it has happened in the previous experiments).

In terms of Sharpe Ratio, increasing the λ value, the return in excess per unit of risk decreases; for this reason the case $\lambda = 0$ describes the better result.

In terms of the transaction costs, the best model is represented by the opposite extreme case, i.e., the case $\lambda = 10^7$; of course, the minor costs observed in this case depend on the weights assigned to the assets. Indeed, the weights are not well distributed in the optimal portfolios, but rather they tend to focus on a single asset, in other words the portfolios are not well diversified.

Let us now describe the performance of the coherent variant (of the *norm-portfolio* problems) and the CVaR model as depicted in figures 4.4-4.4. Since the cases $\pi = \pi^*$ and $\pi = 1$ produce an identical behaviour in terms of portfolio value, on the plot only one of them is reported; in addition, as confirmed by the theory, the results of the coherent variant in the case $\pi = \pi^*$ and the CVaR model coincide almost everywhere. For this, the top figures plot together the first two cases of the coherent variant ($\pi = 0.25$ and $\pi = 0.5$) and the bottom ones plot together the case $\pi = \pi^*$ with the CVaR model. Like before we do not report the performance of the case $\pi = 0$.

Let us observe that all models appear enough similar each others. In fact, the trajectories overlap in the most of cases with some exceptions signed with a circle in figures 4.4-4.4; a slight greater fluctuation for $\pi = 0.25$ (according to the theoretical results) is showed. So, increasing the value of the parameter π the investor does not make different his behaviour.

To better show the performances related to the optimal portfolio value as well as to compare the coherent models easily, we subdivide the whole time period in two subperiods like reported in the following figures.

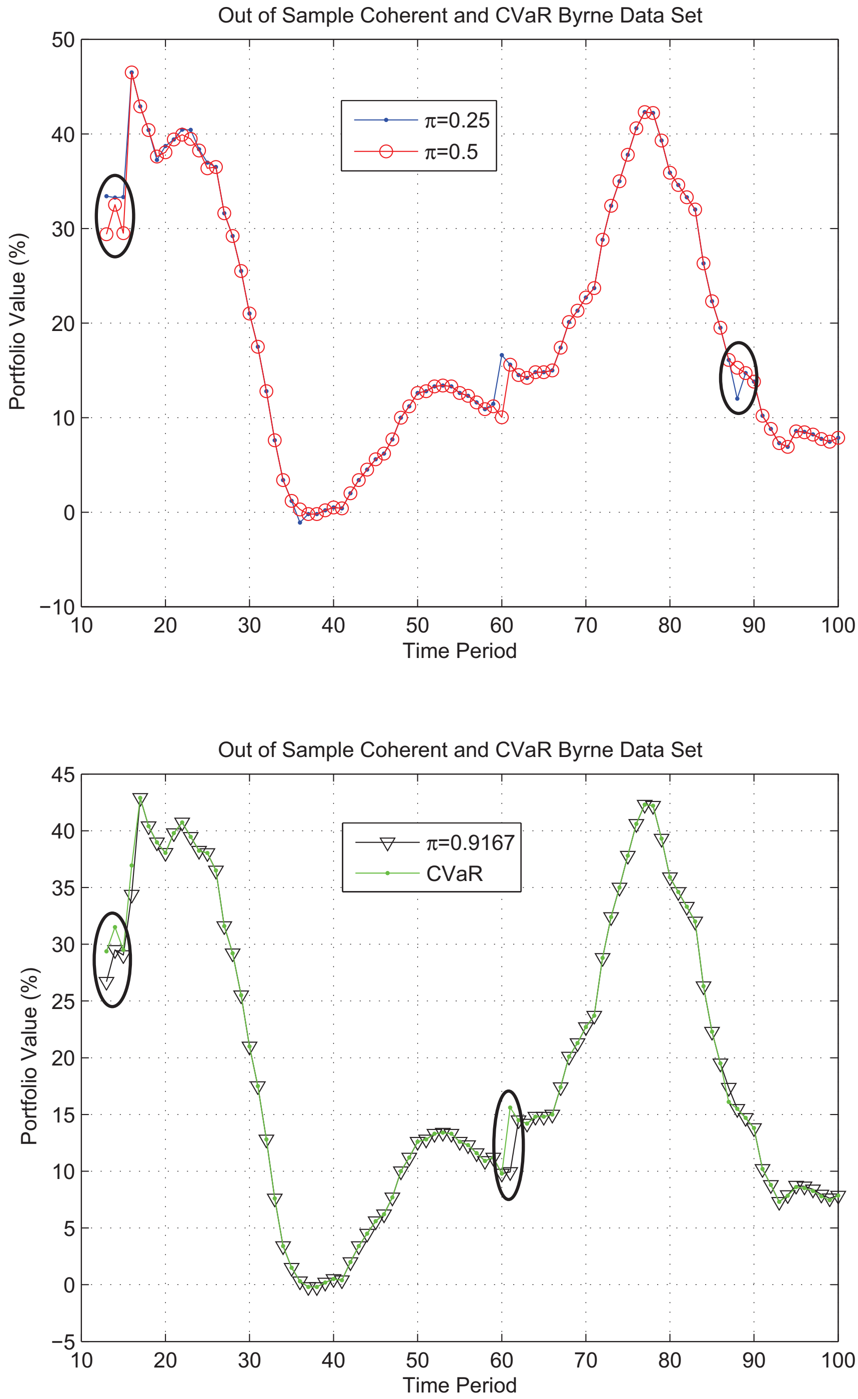


Figure 4.23: *Behaviour of Coherent models for each value of* π *and CVaR model in the first subperiod*

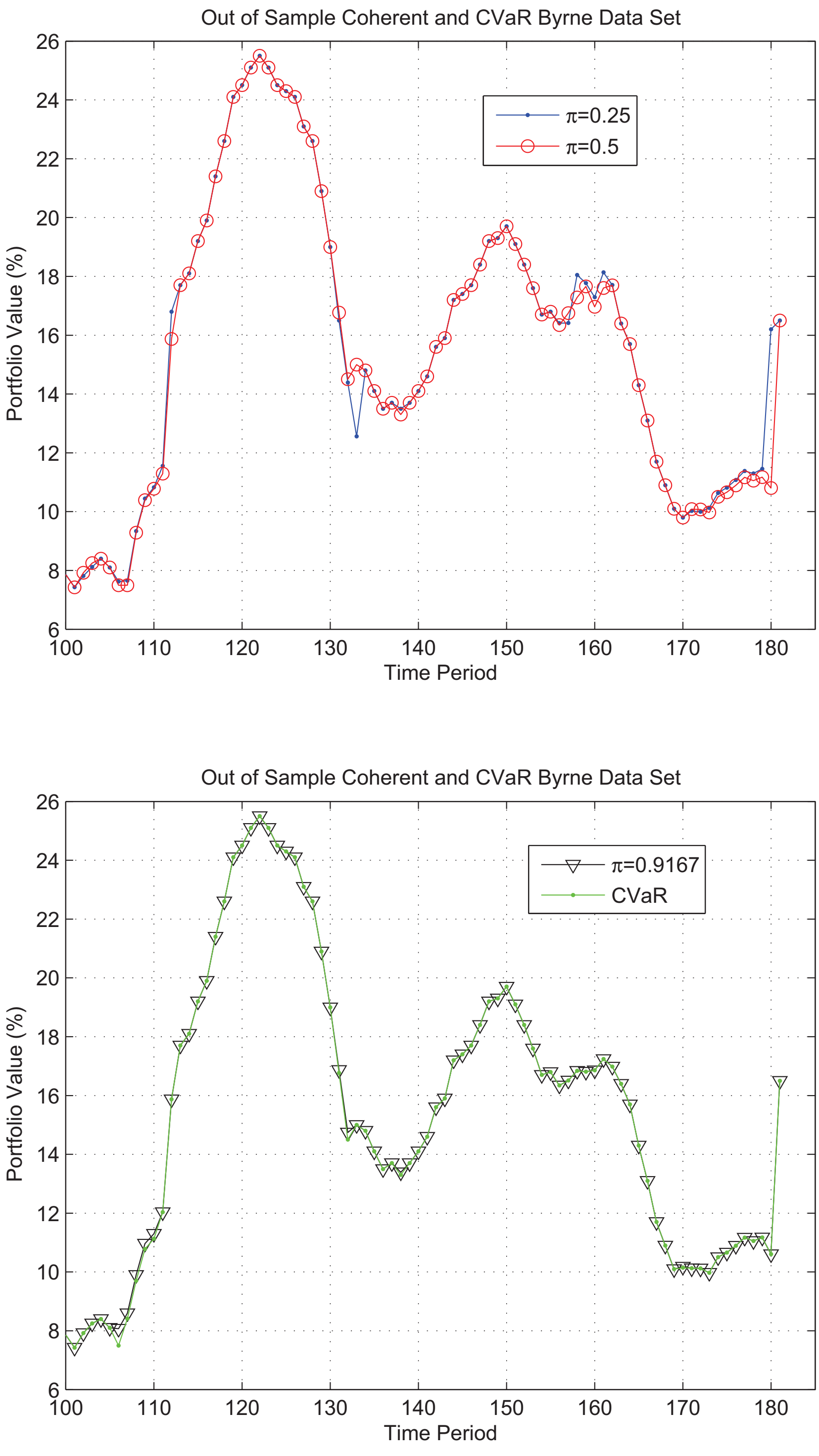


Figure 4.24: *Behaviour of Coherent models for each value of* π *and CVaR model in the second subperiod*

All previous considerations are confirmed by the results related to the chosen statistics that are resumed in the following table:

Coherent variant	**Mean** ($\hat{\mu}$)	**Variance** ($\hat{\sigma}^2$)	**Sharpe Ratio** ($\hat{SR}$)	**Turnover**
$\pi = 0.25$	17.6732	108.5106	1.6997	0.1332
$\pi = 0.5$	17.5536	106.6609	1.5193	0.1518
$\pi = \pi^* = 0.92$	17.4706	103.1427	1.7202	0.1482
$\pi = 1$	17.4708	103.1414	1.7203	0.1489
CVaR	17.5269	104.3457	1.7158	0.1590

Table 4.8: *Out-of-sample mean, variance, Sharpe Ratio and portfolio turnover.*

Table 4.8 indeed reports the performances of the considered models in terms of the out-of-sample mean, variance, Sharpe Ratio and the portfolio turnover. Let us note that the results described by adopted statistics are quite similar each others. By increasing the value of π, the out-of-sample mean and variance decrease, but not in a significant way.

The values of the Sharpe Ratio attain almost the same value (around 1.7) except for the case $\pi = 0.5$ that shows the lowest value of the statistic.

The cases $\pi = \pi^*$ and $\pi = 1$ related to the *coherent* variant reveal an identical behaviour that it is enough similar also to the CVaR model. As confirmed by the theory as well as shown in the figures the coherent variant is equivalent to the CVaR model in terms of optimal portfolio value (the slight difference between coherent cases and the CVaR model could depend on rounding errors in calculating π^* value).

In terms of the portfolio turnover, increasing the value of π, the statistic also increases with exception of the case $\pi = 0.5$ in which an unexpected peak is observed.

Figure 4.25 depicts the performance of the *Tütüncü-Koenig* and the *entropic* models. Let us observe that in terms of positive peaks and negative drops, the entropic model is more similar to the *norm-portfolio* model, whereas the *Tütüncü-Koenig* model seems to show the worst performance. In fact, *Tütüncü-Koenig* model attains lower positive peaks and higher (in absolute value) drops than all other models.

If we plot all models together, we obtain a picture like that one described by figures 4.4-4.27 where we show the performance of the *norm-portfolio* model in the case $\lambda = 0$, the *coherent* variant in the case $\pi = \pi^*$ and the CVaR, *Tütüncü-Koenig* and the *entropic* models.

For the particular data sequence used in this experiment, the traditional robust strategy incorporated by the *Tütüncü-Koenig* model results clearly inferior with respect to the other models: it almost always generates the lowest values of the portfolio.

The performances of the *norm-portfolio* model, its *coherent* variant and CVaR approach are quite close; in the most of cases the trajectories of these three models overlap.

Finally, let us observe the performance of the the *norm-portfolio* model with respect to the *entropic* approach. Except for a few cases in which both models generate close values of the portfolio (such as $t = 43$, $t = 104$, $t = 153$), in the rest of the cases it is not easy to evaluate the performance of the models because there is not a clear superiority of one model with respect to the other one in terms of portfolio value. For example in the period that spans from $t = 13$ to $t = 43$ the *entropic* model produce the highest value of the portfolio while from $t = 108$ to $t = 126$ an opposite behaviour is plotted with the *norm-portfolio* model that overcomes the *entropic* one in terms of portfolio value. However, the final portfolio generated by the *norm-portfolio* model produce a higher value than that generated by the *entropic* portfolio.

The results obtained in this last experiment are resumed in Table 4.9.

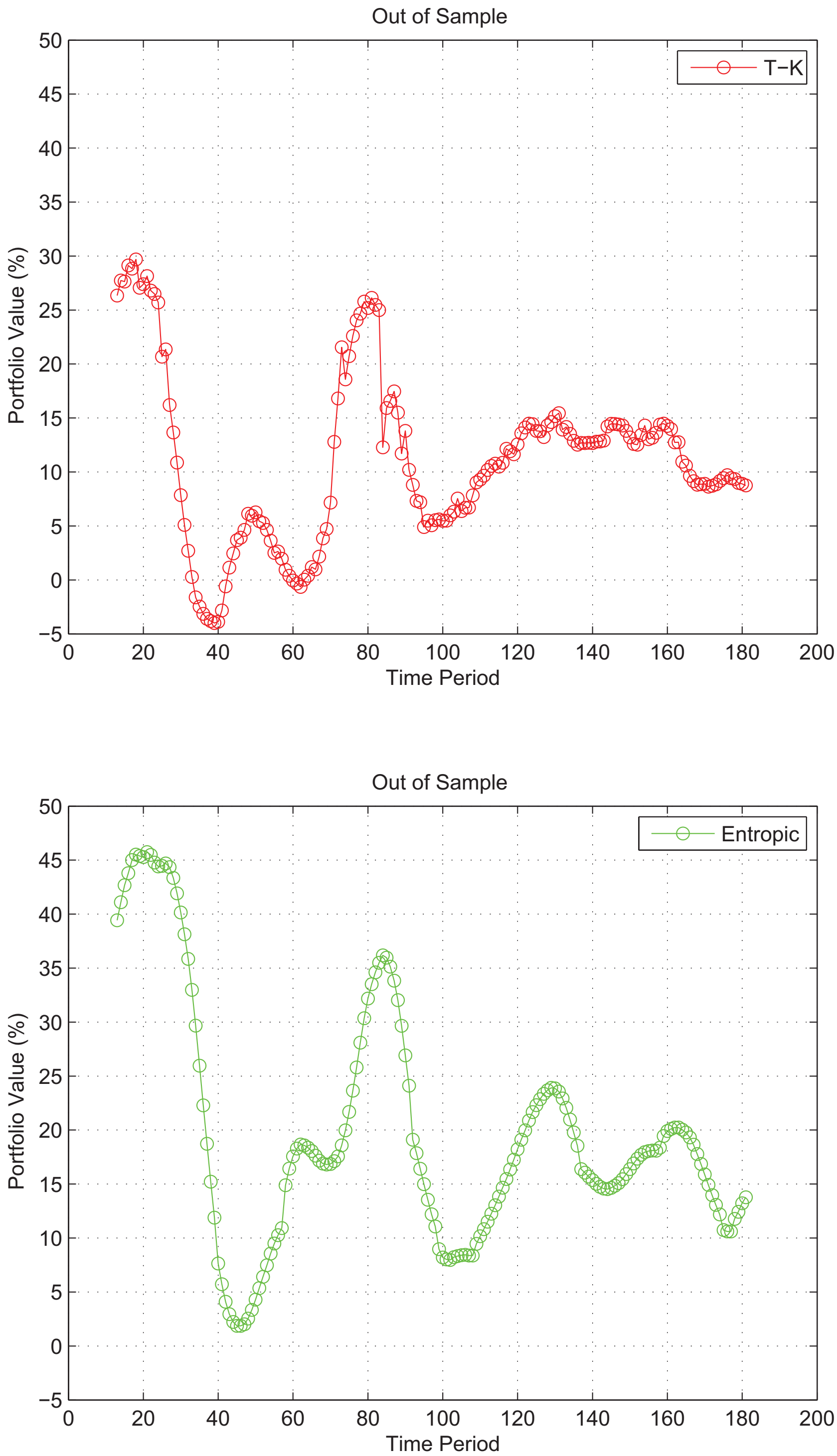


Figure 4.25: *Performances of Tütüncü-Koenig and entropic models*

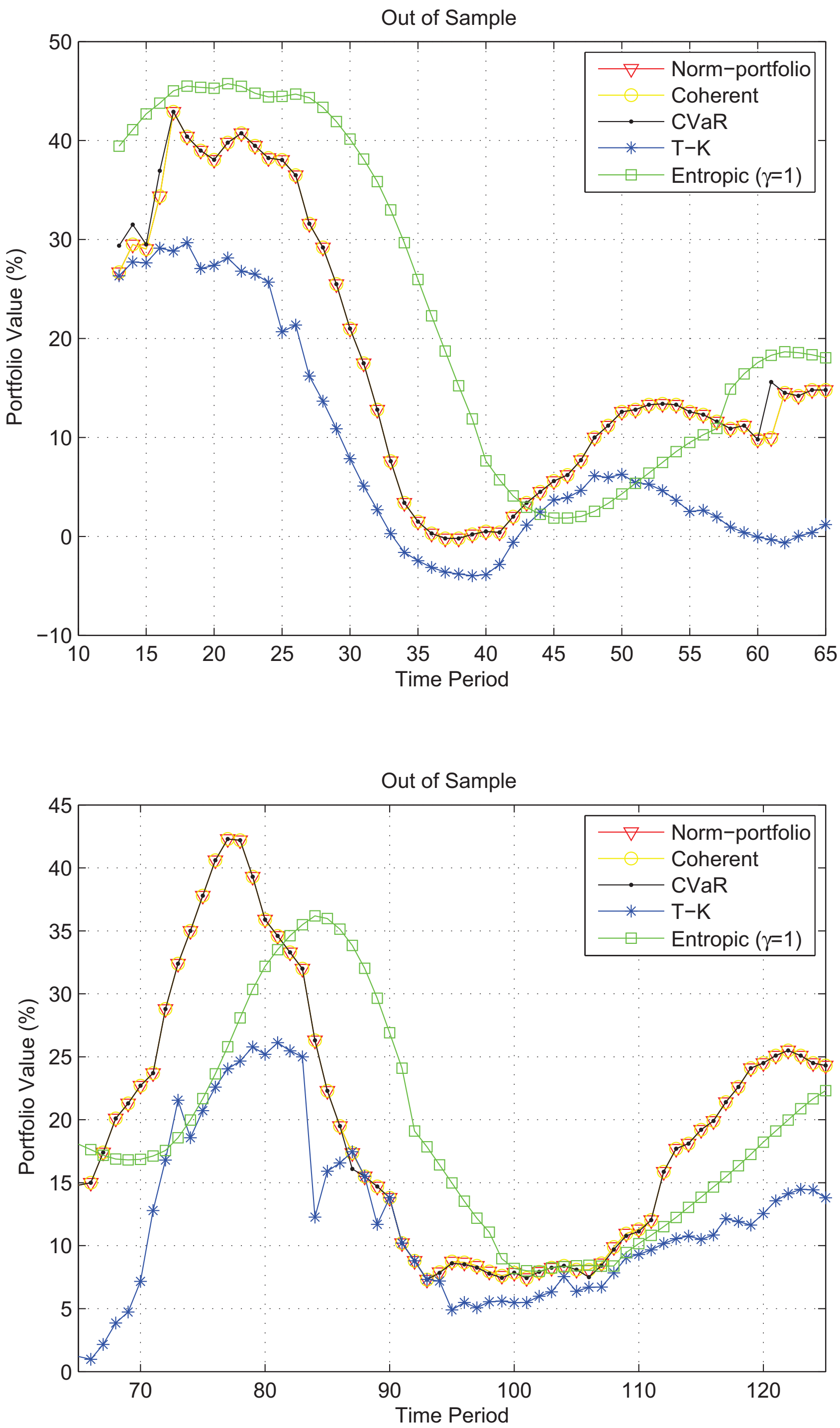
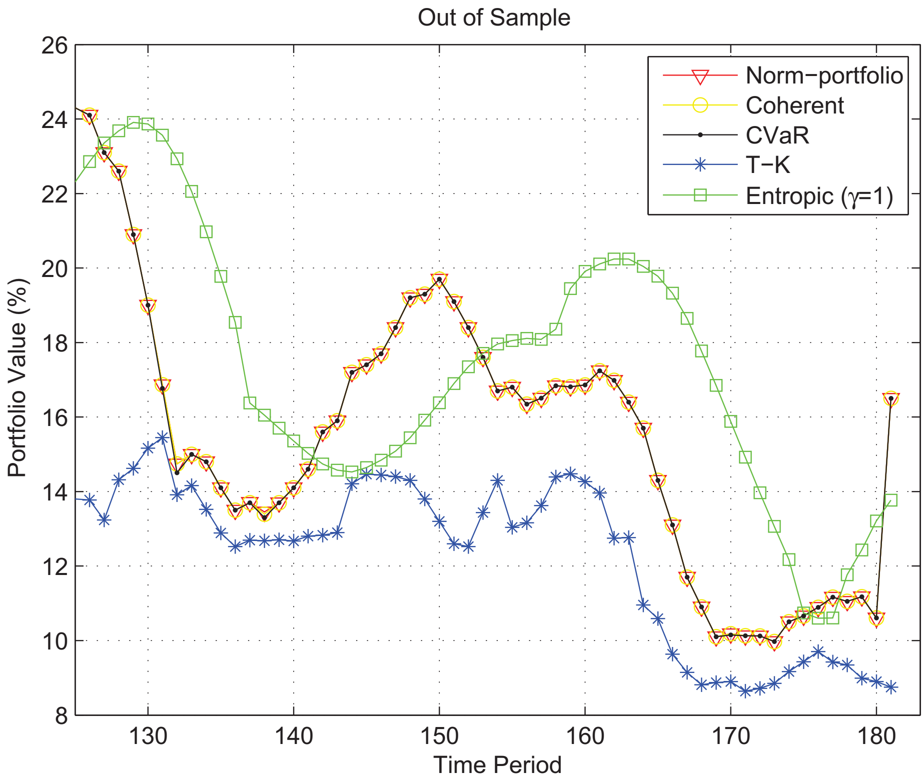


Figure 4.26: *Comparison among the chosen models*

Figure 4.27: *Comparison among the chosen models*



Models	Mean ($\hat{\mu}$)	Variance ($\hat{\sigma}^2$)	Sharpe Ratio ($\hat{SR}$)	Turnover	Time
Norm-portfolio ($\lambda = 0$)	17.4711	103.1393	1.7203	0.1482	0.025
Coherent variant ($\pi = \pi^*$)	17.4706	103.1427	1.7202	0.1482	0.0258
CVaR	17.5269	104.3457	1.7158	0.1590	0.0257
T-K	11.1195	61.0785	1.4228	0.4472	0.1915
Entropic	19.8666	124.1378	1.7831	0.1191	617.1563

Table 4.9: *Out-of-sample mean, variance, Sharpe Ratio, portfolio turnover and computational time.*

Without doubts, as observed by the figures, the *Tütüncü-Koenig* model shows the worst performance in terms of all adopted criteria (except for the out-of-sample variance). It provides the lowest value of the out-of-sample mean and Sharpe Ratio and the the highest value of turnover.

Again, also in terms of statistics adopted is confirmed what we have just observed in the figures, i.e., the *norm-portfolio* model, its *coherent* variant and CVaR model present a similar behaviour.

Finally, from Table 4.9, let us observe that the *entropic* model show a better behaviour in terms of the out-of-sample mean, Sharpe Ratio and turnover, but huge computational time is needed to solve it; let us observe that it is required a mean time about $2.5 \cdot 10^4$ times larger than that one required to solve all other models. As it has been said, this bad result could be depend in part on the nature of the model that it has been solved using approximation methods.

In concluding, the experimental results on this last data set show that the relaxed robustness incorporated by the *norm-portfolio* model, its *coherent* variant and the *entropic* approach provide a better performance than the traditional one incorporated by the *Tütüncü-Koenig* model not only in terms of portfolio value, but also in terms of costs; in fact, the *Tütüncü-Koenig* model produce the worst performance at the highest cost in terms of portfolio turnover. The bad performance of the *Tütüncü-Koenig* model could partly depend on the nature of the data used in this experiment. Indeed, this last data set consists of returns for different segments of the real estate market. Coming from the same industry, the returns on these indices are likely to be more correlated than arbitrary asset returns and they show a non-normal distribution. In such cases portfolio-optimization strategies using the variance as risk measure of risk could result inappropriate [50].

Considering then the results related to the performance of the *entropic* model and the *norm-portfolio* problem we can conclude that the *entropic* model produce better results in terms of the out-of-sample mean, Sharpe Ratio and turnover, but it is not able to provide the optimal solution in a reasonable amount of CPU-time. In addition, if we focus on a long term investment, the *norm-portfolio* model produce the best final portfolio value at the lowest computational cost.

Chapter 5

Conclusions and future researches

In many optimization models, the inputs to the problem are not known at the time the problem must be solved, are computed inaccurately, or are otherwise uncertain. Robust optimization provides an alternative approach to handle this uncertainty.

In the last decades, many results have been obtained in this field above all in terms of comparison with the classical Markowitz's approaches. The focus of this work has been to handle uncertainty using a more flexible robust approach.

Firstly, we have provided an overview of the development in the field of robust optimization including the innovative recent additions to this literature. Some important results from risk measure theory are then presented and the interesting links between risk measures and robust optimization are pointed out. In addition mathematical models and relative algorithmic approaches have been highlighted.

The main theoretical contributions of the dissertation have been presented in Chapter 3 where the notion of "soft robustness" is presented and a new family of models has been introduced.

We start the chapter by discussing different formulations that represent optimal trade-offs between expected returns and arbitrary convex risk measures. From the observation that the problem of choosing a suitable convex risk measure in risk-return trade-off formulations is equivalent to choosing a suitable penalty function $\alpha(q)$ over the set $\mathcal{P}$ of probability measures on the scenario set, we have chosen a particular family of penalty functions by us called *norm-risk* measures. The family of models corresponding to these risk measures (called *norm-portfolio* models) characterizes a more flexible approach to robust portfolio asset allocation, i.e., an approach in which not only the values of the uncertainty parameters, but also their degree of feasibility are specified.

Moreover, it has been proved that this new class of models includes as special cases linear programming (LP) and second order cone programming (SOCP) problems, i.e., solvable models. In order to obtain models computationally tractable, we have focused on penalty functions based on tractable-like measures, such as general norms.

A variant of *norm-portfolio* models based on coherent risk measures has been then proposed as well as an interesting link between this coherent variant and CVaR model has been proved in final session.

Chapter 4 provides a computational analysis of the proposed methodology. We have implemented various robust optimization approaches with real market data including the ones proposed in the dissertation to generate optimal portfolios and we have compared their relative performance. So, we have observed how different risk measures utilize the scenario information based on past history in producing a successful portfolio.

The tested models into computational analysis have been chosen as benchmark for a preliminary comparison among various robustness types: a *relaxed* robustness incorporated by the *norm-portfolio* models and its *coherent* variant, the *soft* robustness described by the *entropic* model and a more *standard* robustness (based on uncertainty sets for the covariance matrix). In addition, a comparison with an approach based on CVaR has been put forward.

In this way, our aim has been to understand how different robustness types influence the optimal portfolio value and its regularity in the long term.

Comparison has been conducted through the evaluation of the following statistics: the out-of-sample mean, variance, Sharpe Ratio of realized return and portfolio turnover. In addition a comparison of computational costs has been conducted through the evaluation of the mean time needed to obtain the optimal solution.

Generally speaking, the computational analysis has proved that a flexible robustness in which not only the values of the uncertain parameter, but also the degree of feasibility are specified produces better results than a traditional one.

In particular in the first two experiments the best approach is resulted the *entropic* one in terms of out-of-sample mean, Sharpe Ratio and variance even if at higher cost in terms of portfolio turnover and mean time needed to solve it. This means that by relaxing the robustness constraints in a flexible way, it is possible gain out-of-sample performance for not too relatively high of a price. About then the *norm-portfolio* model, it has outperformed the traditional one in the first test and it has took up position close (from below) to the traditional approach in the second test.

In the last computational experiment, the performance of soft and relaxed robustness is resulted by far higher than the performance obtained by the traditional robustness. So, in terms of robustness, also in this last case, we can conclude that a general flexible robust approach has gained out-of-sample performance.

However, in all three experiments it is evident as a huge computational cost (calculated in terms of time) is needed to solve the *entropic* model with respect to that needed to solve all other models.

In the field of the relaxed robust portfolio optimization future researches can be developed. Investigating the relation between robustness and convex risk measures by working on alternative penalty functions might be an interesting direction. In addition, a wider computational analysis that takes further robust models into account might provide detailed evidence on the goodness of a more relaxed robust approach to portfolio asset allocation problems.

Bibliography

[1] F. Alizadeh, D. Goldfarb, *Second-order cone programming*, Mathematical programming, Vol.95 (1), pp.3-51, 2003.

[2] P. Artzner, F. Delbaen, J. Eber and D. Heath, *Coherent measures of risk*, Mathematical Finance, Vol.9, pp.203-228, 1999.

[3] D. Bai, T. Carpenter and J. Mulvey, *Making a case for robust optimization models*, Journal of Management Science, Vol.43 (7), 1997.

[4] A. Ben-Tal, M.Teboulle *An old new concept of convex risk measures: the optimized certainty equivalent*, Mathematical Finance, Vol.17 (3), pp.449-476, 2007.

[5] A. Ben-Tal, D. Bertsimas and D.B. Brown, *A soft robust model for optimization under ambiguity*, 2009. To appear in Operation Research.

[6] A. Ben-Tal and A. Nemirovski, *Lectures on Modern Convex Optimization: Analysis, Algorithms, Engineering Applications*, MPS-SIAM Series on Optimization, SIAM, Philadelphia, 2001.

[7] A. Ben-Tal and A. Nemirovski, *Robust convex optimization*, Mathematics of Operation Research, Vol.23 (4), pp. 769-805, 1998.

[8] A. Ben-Tal and A. Nemirovski, *Robust solution of uncertain linear programs*, Operation Research Letters, Vol.25 (1), pp.1-13, 1999.

[9] A. Ben-Tal and A. Nemirovski, *Robust Optimization-methodology and applications*, Mathematical Programming, Series B 92, pp.453-480, 2002

[10] A. Ben-Tal, S. Boyd, A. Nemirovski, *Extending scope of robust optimization:comprehensive robust counterparts of uncertain problems*, Mathematical Programming, Series B 107, pp.63-89, 2006.

[11] S. Benati and R. Rizzi, *A mixed integer linear programming formulation of the optimal man/Value-at-Risk portfolio problem*, European Journal of Optimization Research, Vol.176 (1), pp.423-434, 2007.

[12] D. Bertsimas and D.B. Brown, *Constructing uncertainty sets for robust linear optimization*, Operation Research, Vol.57 (6), pp. 1483-1495, 2009.

[13] D. Bertsimas, D.B. Brown and C. Caramanis, *Theory and applications of Robust Optimization*, 2008, Submitted.

[14] D. Bertsimas and D. Pachamanova, *Robust multiperiod portfolio management in the presence of transaction costs*, Computers and Operation Research, Vol.35 (1), pp. 3-17, 2008.

[15] D. Bertsimas, D. Pachamanova and M.Sim, *Robust Linear Optimization under General Norms*, Operation Research Letters, Vol.35, pp. 510-516, 2004.

[16] D. Bienstock, *Histogram Models for Robust Portfolio Optimization*, The Journal of Computational Finance, Vol.11 (1), 2007.

[17] F. Black and R. Litterman, *Global Portfolio Optimization*, Financial Analysts Journal Vol.48 (5), pp.28-43, 1992.

[18] P. Byrne and S. Lee, *Different Risk Measures: Different Portfolio Compositions?*, Journal of Property Investment and Finance, Vol.2 (6), pp. 501-511, 2004.

[19] R. Byrd, M.E. Hribar, J. Nocedal *An interior point method for large scale nonlinear programming*, SIAM Journal of Optimization Vol.9 (4), pp.877-900, 1999.

[20] M. Broadie, *Computing efficient frontiers using estimated parameters*, Annales of Operation Research Vol. 45, pp.21-58, 1993.

[21] D.B. Brown *Risk and robust optimization*, Ph.D. Dissertation, Massachusetts Institute of Technology, 2006.

[22] F. Cavadini, A. Sbuelz, F. Trojani *A simplified way of incorporating model risk, estimation risk and robustness in mean variance portfolio management* Working Paper, 2001.

[23] S. Ceria and R. Stubbs, *Incorporating Estimation Errors into Portfolio Selection: Robust Portfolio Construction*, Journal of Asset Management, Vol.7 (2), pp.109-127, 2006.

[24] L.K.C. Chan, J. Karceski, J. Lakonishok *On portfolio optimization: Forecasting covariances and choosing the risk model*, Review of Financial Studies Vol.12 (5), pp.937-974, 1999.

[25] V.K. Chopra, W.T. Ziemba *The effect of errors in means, variances, and covariances on optimal portfolio choices*, Journal of Portfolio Management Vol.19 (2), pp.6-11, 1993.

[26] G. Cornuejols and R.H. Tütüncü, *Optimization Methods in Finance*, Cambridge University Press, 2007.

[27] O.L. Costa and A.C. Paiva, *Robust portfolio selection using linear matrix inequalities*, Journal of Economic Dynamics and control, Vol.26 (6), pp. 889-909, 2002.

[28] V. DeMiguel and F.J. Nogales, *Portfolio selection with robust estimation*, Operation Research, Vol.57 (3), pp.560-577, 2009.

[29] D. Denneberg *Non-Additive Measure and Integral*, Kluwer Academic Publishers, Boston, 1994.

[30] J. Dhaene, S. Vanduffel, Q. Tang, M. Goovaerts, R. Kaas, D. Vyncke, *Risk measures and comonotonicity: a review* Stochastic Models Vol.22, pp.573-606, 2006.

[31] M. Diehl, H.G. Bock and E. Kostina, *An approximation technique for robust nonlinear optimization*, Mathematical Programming, Vol.107, pp. 213-230, 2006.

[32] L. El Ghaoui and H. Lebret, *Robust solutions to least-squares problems with uncertain data*, SIAM Journal on Matrix Analysis and Applications, Vol.18 (4), pp. 1035-1064, 1997.

[33] L. El Ghaoui, F. Oustry and H. Lebret, *Robust solutions to uncertain semidefinite programs*, SIAM Journal on Optimization, Vol.9, pp. 33-52, 1998.

[34] L. El Ghaoui, M. Oks and F. Oustry, *Worst Case Value-at-Risk and robust portfolio optimization: A conic programming approach*, Operations Research, Vol.51 (4), pp. 543-556, 2003.

[35] F.J. Fabozzi, Pettern N. Kolm, Dessislava A. Pachamanova, Sergio M. Focardi, *Optimization and Management*, John Wiley and Sons, Inc, 2007.

[36] H. Föllmer and A. Schied, *Convex measures of risk and trading constraints*, Finance and stochastic, Vol.6, pp.429-447, 2002.

[37] H. Föllmer and A. Schied, *Robust preferences and convex risk measures*, Advances in Finance and Stochastic: Essays in Honour of Dieter Sondermann, pp. 39-56, Springer-Verlag, Berlin, 2002.

[38] H. Föllmer and A. Schied, *Stochastic Finance: An Introduction in Discrete Time*, Walter de Gruyter, Berlin 2004.

[39] A. Gaivoronski, G. Pflug *Value-at-Risk in Portfolio Optimization: Properties and Computational Approach* Journal of Risk, Vol.7 (2), pp.1-31, 2005.

[40] Goh J, Sim M. ROME: Robust Optimization Made Easy. Version 1.0.4 (beta), available at http://robustopt.com/

[41] D. Goldfarb and G. Iyengar, *Robust Portfolio Selection*, Mathematics of Operations Research, Vol.28 (1), pp.1-38, 2003.

[42] M. Grant and S. Boyd, *Cvx: Matlab software for disciplined convex programming*, web page:http://stanford.edu/boyd/cvx, 2008.

[43] B.V. Halldórsson and R.H. Tütüncü, *An Interior-Point Method for a Class of Saddle-Point Problems*, Journal of Optimization Theory and Applications, Vol.116 (3), pp.559-590, 2003.

[44] P.J. Huber *Robust Statistics*, John Wiley & Sons, New York, 2004.

[45] R. Jagannathan, T. Ma *Risk reduction in large portfolios: Why imposing the wrong constraints help?* Journal of Finance, Vol.58 (4), pp.1651-1684, 2003.

[46] R.A. Jarrow and A.K. Purnanandam, *A generalized coherent risk measure: the firm's perspective*, Finance Research Letters, Vol.2 (1), pp. 23-29, 2005

[47] P. Krokhmal, J. Palmquist and S. Uryasev, *Portfolio optimization with conditional value-at-risk objective and constraints*, Journal of risk, Vol.4 (2), 2002.

[48] P. Krokhmal, S. Uryasev and G. Zrazhevsky, *Numerical Comparison of CVaR and CDaR Approaches: Application to Hedge Funds*, Report.

[49] G.J. Lauprete, A.M. Samarov, R.E. Welsch *Robust portfolio optimization*, Metrica Vol.25, pp.139-149, 2002.

[50] C. Lizieri and C.W.R. Ward, *The Distribution of Real Estate Returns* In: J. Knight and S. Satchell (eds) Return Distributions in Finance. London: Butterworth-Heinemann, 2000.

[51] M.S. Lobo, L. Vandenberghe, S. Boyd, H. Lebret *Applications of second order cone programming*, Linear Algebra and its Applications Vol.284 (1-3), pp.193-228, 1998.

[52] H.J. Lüthi and J. Doege, *Convex risk measures for portfolio optimization and concepts of flexibility*, Mathematical Programming, Vol.104 (2-3), pp.541-559, 2005.

[53] H. Markowitz, *Portoflio Selection*, Journal of Finance, Vol.7, pp.77-91;

[54] A.J. McNeil, R. Frey, P. Embrechts *Quantitative Risk Management: Concepts, Techniques, Tools Princeton Series in Finance*, Princeton University Press, (2005).

[55] B.V.M. Mendes, C.R.P. Leal, *Robust modeling of multivariate financial data*, International Journal of Managerial Finance Vol.1 (2), pp.95-106.

[56] R.O. Michaud, *The Markowitz optimization enigma: is optimized optimal?*, Financial Analysts Journal Vol.45 (1), pp.31-42, 1989.

[57] J.P. Morgan Inc., *Risk Metrics*^{TM}, technical document (4th edn), New York:Morgan, 1996.

[58] K. Natarajan, D. Pachamanova and M. Sim *Constructing Risk Measures from Uncertainty Sets*, Working paper, 2005.

[59] K. Natarajan, D. Pachamanova, M. Sim *Incorporating Asymmetric Distributional Information in Robust Value at Risk Optimization*, Management science Vol.54 (3), pp.573-585, 2008.

[60] Y. Nesterov, A. Nemirovski, *Interior-point polynomial algorithms in convex programming*, SIAM, Philadelphia, PA 1993.

[61] Y. Nesterov , A. Nemirovski *Interior point polynomial methods in convex programming: Theory and Applications*, SIAM, Philadelphia, PA 1994.

[62] C. Perret-Gentil, M.P. Victoria-Feser *Robust mean-variance portfolio selection*, FAME Research Paper 140. International Center for Financial Asset Management and Engineering, Geneva 2004.

[63] G.Ch. Pflug, *Some remarks on the Value-at-Risk and the Conditional Value-at-Risk*, In. 'Probabilistic Constrained Optimization: Methodology and Applications', Ed. S.Uryasev, Kluver Academic - publishers, 2000.

[64] M.. Pinar and R.H. Tütüncü, *Robust profit opportunities in risky financial portfolios*, Operation Reserach Letters, Vol.33 (4), pp.331-340, 2005.

[65] A.G. Quaranta and A. Zaffaroni, *Robust Optimization of Conditional Value at Risk and Portfolio Selection*, Journal of Banking and Finace, Vol.32 (10), pp.2046-2056, 2008.

[66] R. Recchia and M.G. Scutellà, *Robust portfolio asset allocation and risk measure*, 2010. To appear in 4OR Journal.

[67] R. Recchia and M.G. Scutellà, *Experiments with robust asset allocation strategies: classical versus relaxed robustness*, 2010. Submitted.

[68] R.T. Rockafellar, *Convex Analysis*, Princeton University Press, Princeton, New Jersey, 1970.

[69] R.T. Rockafellar and S.Uryasev, *Conditional value at risk for general loss distributions*, Journal of Banking and Finance, Vol.26 (7), pp.1443-1471, 2002.

[70] R.T. Rockafellar and S.Uryasev, *Optimization of Conditional Value-at-Risk*, Journal of risk, Vol.2 (3), pp.21-41, 2000.

[71] R. Saigal, L. Vandenberghe, H. Wolkowicz *Handbook on Semidefinite Programming and Applications*, Kluwer Academic Publishes, Dordrecht, The Netherlands 2000.

[72] Santos A.A.P., *The out-of-sample performance of robust portfolio optimization*, Technical Report, 2008.

[73] Scherer B., *Can robust portfolio optimisation help to build better portfolios?*, Journal of Asset Management, Vol.7 (69), pp.374-387, 2007.

[74] K. Schöttle and R. Werner, *Towards reliable efficient frontiers*, Journal of Asset Management, Vol.7 (2), 2006.

[75] A.L. Soyster, *Convex programming with set-inclusive constraints and applications to inexact linear programming*, Operations Research, pp.1154-1157, 1973, pp.128.141, 2006.

[76] J. Stürm *Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones*, Optimation methods and software, Vol.11-12, pp.625-653, 1999.

[77] J. Stürm *Using SeDuMi, a MATLAB toolbox for optimization over symmetric cones*, Department of Ecnometrics, Tilburg University, The Netherlands, (2001).

[78] M.J. Todd *Semidefinite optimization*, Acta Numerica Vol.10, pp.515-560, 2001.

[79] R.H. Tütüncü and M. Koenig, *Robust Asset Allocation*, Annals of Operation Research, 132, pp.157-187, 2004.

[80] L. Vandenberghe, S. Boyd, *Semidefinite programming*, SIAM Review Vol.38, pp.49-95, 1996.

[81] R.A. Waltz, *KNITRO User's Manual version 4.0*, Ziena Optimization, Evanston, IL. 2004.

[82] S. Wang, *Premium calculation by transforming the layer premium density*, ASTIN Bullettin Vol.26, pp.71-92, 1996.

[83] R.E. Welsch, X. Zhou, *Application of robust statistics to asset allocation models*, Revstat Statistical Journal, Vol.5 (1), pp.97-114, 2007.

[84] M.E. Yaari, *The dual theory of choice under risk*, Econometrica Vol.55, pp.95-115, 1987.

[85] S.S. Zhu and M. Fukushima, *Worst-Case Conditional Value at Risk with Application to Robust Portfolio Management*, Operation Research Vol.57 (5), pp.1155-1168, 2009.