Università di Pisa
Dipartimento di Informatica

Technical Report: TR-10-04

Experiments with robust asset allocation strategies: classical versus relaxed robustness

Raffaella Recchia
recchia@di.unipi.it
Maria Grazia Scutellà
scut@di.unipi.it

Dipartimento di Informatica, Università di Pisa

February 23, 2010

ADDRESS: Largo B. Pontecorvo 3, 56127 Pisa, Italy. TEL: +39 050 2212700 FAX: +39 050 2212726

Experiments with robust asset allocation strategies: classical versus relaxed robustness

Raffaella Recchia
recchia@di.unipi.it
Maria Grazia Scutellà *
scut@di.unipi.it

Dipartimento di Informatica, Università di Pisa

February 23, 2010

Abstract: Many optimization problems involve parameters which are not known in advance, but can only be forecast or estimated. This is true, for example, in portfolio asset allocation. Such problems fit perfectly into the framework of Robust Optimization that, given optimization problems with uncertain parameters, looks for solutions that will achieve good objective function values for the realization of these parameters in given uncertainty sets.

Aim of this paper is to compare alternative forms of robustness in the context of portfolio asset allocation. Starting with the concept of convex risk measures, a new family of models, called *norm-portfolio* models, is firstly proposed where not only the values of the uncertainty parameters, but also their degree of feasibility are specified. This relaxed form of robustness is obtained by exploiting the link between convex risk measures and classical robustness.

Then, we test some norm-portfolio models, as well as various robust strategies from the literature, with real market data on three different data sets. The objective of the computational study is to compare alternative forms of relaxed robustness - the relaxed robustness characterizing the norm-portfolio models, the so-called soft robustness and the CVaR robustness. In addition, the models above are compared to a more classical robust model from the literature, in order to experiment similarities and dissimilarities between robust models based on convex risk measures and more traditional robust approaches. To the best of our knowledge, this is the first attempt at comparing robust strategies of different kinds in the framework of portfolio asset allocation.

Keywords: portfolio optimization, robustness, convex risk measures, mathematical models, computational experimentation.

*Corresponding author: Dipartimento di Informatica, Università di Pisa, Italy; e-mail: scut@di.unipi.it; fax: +39-050-2212726

1 Introduction

The first pioneering contribution in the field of allocation of financial assets was made by Markowitz in 1952. The so-called *mean-variance models* by him proposed consist in allocating capital over a number of available assets in order to maximize the *return* on the investment (measured by the expected value) while minimizing a certain measure of *risk* (quantified by the variance of the portfolio).

Despite the strong theoretical support provided 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 available securities and, above all, they are often very sensitive to changes in input parameters.

Several methodologies have then been proposed to reduce the sensitivity of Markowitz's models, such as *Robust Optimization*. This framework entails modeling optimization problems with uncertain parameters to obtain a solution that is guaranteed to be 'good' for all possible realizations of the parameters in given uncertainty sets. Uncertainty in the parameters is in fact modelled using *uncertainty sets*, which contain possible realizations for the uncertain parameters.

Robust models and related algorithmic approaches have been recently proposed in the literature to address uncertainty in portfolio asset allocation problems. Some of these models and methods are described in Fabozzi et al. (2007). More recent robust models are presented in Recchia and Scutellà (2009).

This paper is organized as follows. After a quick review of various classical robust models in Section 2.2, in Section 3 we describe some more flexible robust models that, by exploiting the theoretical link between convex risk measures and classical robustness, furnish an innovative approach to robustness. In Section 4, based on the more flexible concept of robustness mentioned above, in which not only the values of the uncertainty parameters, but also their degree of feasibility are specified, we propose a new family of models that we call the *norm-portfolio* models, that include as special cases linear programming models (LP) and second order cone programming models (SOCP), i.e., computationally tractable models. Specifically, we focus on the notion of penalty function in convex risk measures in order to propose models where these penalty functions are defined in terms of general norms. We then study a variant where the risk measure used is also coherent (a subcase of the convex one), and we conclude the section with considerations about some parameters of the models, that describe an interesting link between this coherent variant of the norm-portfolio models and one of the well-known coherent risk measures in the literature, i.e. the Conditional Value at Risk (CVaR).

In Section 5 we report the results of a computational experimentation performed with real market data on three different data sets. The objective of the computational study was to compare alternative forms of relaxed robustness - the relaxed robustness characterizing the norm-portfolio models, the soft robustness characterizing the entropic model (Ben-Tal et al. 2009), and the CVaR robustness - in order to assess how "simple" penalty functions such as the ones incorporated by the norm-portfolio models compare in practice to more sophisticated convex risk measure approaches. In addition, the models above have been compared to (a slight variant of) Tütüncü-Koenig model (Tütüncü and Koenig 2004), in order to compare robust models based on convex risk measures to more traditional robust approaches. To the best of our knowledge, this is the first attempt at comparing robust strategies of different kinds (i.e. standard robust models based on uncertainty sets versus relaxed robust models based on convex risk measures) in the framework of portfolio asset allocation. In the previous experiments, in fact, the objective was to contrast the performance of some classical portfolio selection strategies (usually the mean-variance and the minimum-variance approaches) with some specific robust selection strategies. In a few cases, different robust methods were also compared (DeMiguel and Nogales 2009, Ben Tal et al. 2009). In those cases, however, the methods compared are of the same kind (methods based on robust estimators in the first case, and based on convex risk measures in the second case).

Section 6 concludes the paper, by providing some final comments and directions for future research, while some proofs are reported in the Appendix.

2 Robust asset allocation

2.1 The classical models

Portfolio asset allocation problems can be formulated mathematically as quadratic programming (QP) problems (Markowitz 1952; Tütüncü and Koenig 2004). Some of them can be formulated as convex QP, that refers to minimizing a quadratic function subject to linear constraints.

Let n be the number of the available assets, and X be the non-empty and bounded set of the feasible portfolios. For example

$$X = \left\{ x \in \mathbb{R}^n | \sum_{i=1}^{n} x_i = 1,\ x_i \geq 0, i = 1, ..n \right\} \tag{1}$$

formulates the case where short-sales are not allowed. Furthermore, let μ be the estimated expected return vector of the given assets, while matrix Q be the covariance matrix of these returns.

The classical *mean-variance optimization (MVO) models* of Markowitz can then 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) minimize the variance subject to a lower limit on the expected return:

$$\begin{aligned} \min \quad & x^T Q x \\ s.t. \quad & \mu^T x \geq R \\ & x \in X; \end{aligned} \tag{3}$$

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

$$\begin{aligned} \max \quad & \mu^T x - \lambda x^T Q x \\ s.t. \quad & x \in X \end{aligned} \tag{4}$$

where $\lambda \in \mathbb{R}$ denotes a risk-aversion parameter.

These three models are parametrized by the variance limit, the expected return limit and the risk-aversion parameter, respectively. Since the variance constraint is a nonlinear constraint, the first formulation can not be classified as a convex QP formulation, while the latter two are convex QP formulations solvable in polynomial time.

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 (Michaud 1989; Black and Litterman 1992; Chopra and Ziemba 1993; Broadie 1993). Moreover, it is commonly accepted that estimation errors in the sample expected return are much larger than in the sample covariance matrix (Chan et. al. 1999; Jagannathan and Ma 2003). 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 only on estimates of the covariance matrix, and therefore usually perform better out of sample.

However, these portfolios are also quite sensitive to estimation errors and have unstable weights that fluctuate substantially over time. The main reason for this 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 particularly relevant for portfolio asset allocations, where extensive evidence shows that the empirical distribution of asset returns usually deviates from the normal distribution (DeMiguel and Nogales 2009).

For practical applications, it is therefore crucial to incorporate the uncertainty regarding the accuracy of the estimates in the portfolio optimization process. One methodology addressing uncertainty in minimum-variance models is *Robust Statistics*. This methodology studies the problem of making estimates that are insensitive to small changes in the assumptions of the models used.

Methods have been proposed which generate robust estimators of the portfolio return characteristics, which enable portfolios with better stability properties to be generated. A different way to achieve robustness is *Robust Optimization*. This framework entails modeling optimization problems with uncertain parameters to obtain a solution that is guaranteed to be “good”for all possible realizations of the parameters in given uncertainty sets. Uncertainty in the parameters is in fact modelled using *uncertainty sets*, which contain possible realizations for the uncertain parameters. Robust mean-variance models have been proposed which define suitable uncertainty sets for μ and Q, and which select the optimal portfolio with respect to the worst data realization within the chosen uncertainty sets. For a survey of models based on robust estimators see Fabozzi et al. (2007) and Recchia and Scutellà (2009). Some robust models and methods belonging to the framework of robust optimization are reviewed in Section 2.2, with an emphasis on those approaches addressed in the computational experiments (see Section 5). For more details and proofs see Fabozzi et al. (2007) and Recchia and Scutellà (2009).

2.2 Robust mean-variance models

Let us assume that the uncertain expected return vector μ and the uncertain covariance matrix Q of the asset returns belong to uncertainty sets with the following form of intervals:

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

where the relation $\leq$ is intended to hold true componentwise (considering both vectors and matrices), while the restriction $Q \succeq 0$ indicates that Q must be a positive semidefinite matrix, since it denotes a covariance matrix.

Based on the uncertainty sets above, Tütüncü and Koenig (2004) formulated robust counterparts of problems (4) and (3) by exploiting formulations previously introduced by Goldfarb and Iyengar (2003) and by Halldórsson and Tütüncü (2003). 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 among the feasible portfolios. In contrast, the latter robust model looks for a feasible portfolio x which guarantees the lower limit R on the expected return also in the worst case, and which minimizes the variance in the worst realization of parameter Q in $\mathcal{U}_Q$. The resulting robust counterparts are therefore:

$$\max_{x \in X} \left\{ \min_{\mu \in \mathcal{U}_\mu, Q \in \mathcal{U}_Q} \mu^T x - \lambda x^T Q x \right\} \tag{5}$$

and

$$\begin{aligned} \min \max_{Q \in \mathcal{U}_Q} \quad & x^T Q x \\ s.t. \min_{\mu \in \mathcal{U}_\mu} \quad & \mu^T x \geq R, \\ & x \in X \end{aligned} \tag{6}$$

Under certain simplifying assumptions, i.e. 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, i.e. μ^L for the uncertain expected return vector μ and Q^U for the uncertain covariance matrix Q, and then solving the resulting QP problems (Tütüncü and Koenig 2004). Without these assumptions, it is not possible to solve the robust asset allocation problems as standard QPs. In the general case, the robust counterparts (5) and (6) can be solved using a nonlinear saddle-point formulation that involves semidefinite constraints (Halldórsson and Tütüncü 2003).

From a computational point of view, Tütüncü and Koenig generated the so-called robust efficient frontier using two market data sets, and compared the robust efficient frontier with the classical efficient frontier, generated solving model (3). The first data set is composed of five asset classes, from January 1979 to July 2002, for a total of 283 months. The second data set addresses eight asset classes from 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. Moreover, the generated robust portfolios show more stability over time, i.e. they remain relatively unchanged over long periods of time.

An alternative method for modeling the uncertainty was proposed by Goldfarb and Iyengar (2003) using the factor model. Under the hypothesis of normality of the random return, and considering suitably defined uncertainty sets for the parameters of the factor model, Goldfarb and Iyengar presented robust optimization models that can be reduced to second order cone programming problems (SOCP). Additional asset allocation problems, called the *Sharpe Ratio problems*, have been studied in the literature, which incorporate assets considered essentially riskless. For a survey of these problems and their robust counterparts we refer to Goldfarb and Iyengar (2003) and Tütüncü and Koenig (2004). See also Recchia and Scutellà (2009).

3 Robustness and risk measures

3.1 Risk measures and related models

In recent years, there has been an increasing interest in defining quantitative methods to assess the risk of financial positions. Basically there are two types of risk measures: *dispersion* and *downside* measures (Fabozzi et al. 2007). Dispersion measures consider both positive and negative deviations from the expected return, and treat these deviations as being 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.

Let us introduce some advanced concepts of downside risk measures, together with related minimum 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. A quantitative *measure of risk* can then be modelled as a mapping ρ from the space of these return functions into the real line. This is the classical definition of the measure of risk as provided by Artzner et al. (1999) in their seminal contribution, where they systematically analysed the concept of risk measures, by formulating certain axioms that should be satisfied by any reasonable measure of risk.

Here the emphasis is on the definitions of monetary risk measures, convex and coherent risk measures, as stated in (Föllmer and Schied 2004). The reason being, from these definitions, Föllmer and Schied (2004) have provided some risk measure characterizations that have been used, quite recently, to model interesting robust optimization models in the framework of asset allocation problems.

Let us denote by Φ the set of all bounded measurable functions on the set of scenarios Ω, and $\mathcal{P}$ the set of all probability measures on Ω. For any $Y, Z \in \Phi$, the shorthand notation $Y \leq Z$ denotes $Y(\omega) \leq Z(\omega)\ \forall \omega \in \Omega$. We define the following (Föllmer and Schied 2004):

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

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

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

Definition 3 *A convex measure of risk ρ is called a coherent risk measure if it satisfies* [1]

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

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

Artzner et al. (1999) provided the following characterization of coherent risk measures (also cited in Föllmer and Schied (2004)):

Theorem 4 *A functional $\rho: \Phi \rightarrow R$ is a coherent measure of risk if and only if there exists a subset $\mathcal{Q} \subseteq \mathcal{P}$ such that*

$$\rho\left(Y\right)=\sup_{q \in \mathcal{Q}} E_{q}\left[-Y\right], \quad Y \in \Phi, \tag{7}$$

where $E_{q}\left[-Y\right]$ 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 (Föllmer and Schied 2002):

Theorem 5 *Suppose that Ω is a finite set* [2]*. Then $\rho: \Phi \rightarrow R$ is a convex measure of risk if and only if a penalty function $\psi: \mathcal{P} \rightarrow R \cup(-\infty,+\infty]$ exists such that*

$$\rho(Y)=\sup_{q \in \mathcal{P}}\left(E_{q}\left[-Y\right]-\psi\left(q\right)\right). \tag{8}$$

The function ψ satisfies $\psi(q) \geq -\rho(0)$ 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 ψ assigns a possibly different weight to the probabilities in $\mathcal{P}$, by penalizing some of them. By choosing $\psi\left(q\right)=0$ for all $q \in \mathcal{Q}$, and $+\infty$ otherwise, the characterization of coherent measures of risk stated by Theorem 4 can be derived as a special case. The characterization expressed by Theorem 5 has been exploited to propose more flexible robust models for asset allocation problems, which will be reviewed in Section 3.2. Next we will briefly review some well-known measures of risk together with the related minimum 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

[1] In Artzner et al. (1999), a mapping $\rho: \Phi \rightarrow \mathbb{R}$ is called a coherent measure of risk if 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.

[2] A generalization of this theorem also exists in the case where Ω is an infinite set.

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). The α-VaR risk measure of x is then defined as follows:

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

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 by Gaivoronski and Pflug (2005). Several other approaches to VaR optimization are used in practice, which show that, in some contexts, VaR can be a suitable measure of risk (Natarajan et al. 2008). However, VaR also has several undesirable properties as a risk measure. Firstly, if studied within the framework of coherent risk measures, it lacks subadditivity (and therefore convexity) (Artzner et al. 1999). 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. Moreover, VaR does not take into account 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. Given a probability level α, the α-CVaR associated with a portfolio x is defined as follows:

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

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

Rockafellar and Uryasev (2000) showed that minimizing CVaR can be achieved by minimizing a more tractable auxiliary function without first predetermining the corresponding VaR. They introduced the following simpler auxiliary function

$$F_\alpha(x, \gamma) = \gamma + \frac{1}{1-\alpha} \int_{f_x(\omega) \geq \gamma} f_x\,(\omega) p\,(\omega)\, d\omega. \tag{9}$$

This formulation can be written in the following equivalent way:

$$F_\alpha(x, \gamma) = \gamma + \frac{1}{1-\alpha} \int (f_x(\omega) - \gamma)^+ p\,(\omega)\, d\omega \tag{10}$$

where $a^+ = \max\{a, 0\}$. Rockafellar and Uryasev showed that $F_\alpha(x, \gamma)$ verifies some interesting properties such that minimizing CVaR is equivalent to minimizing 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 $f_x(\omega)$ 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, provided 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:

$$\tilde{F}_\alpha(x,\gamma) = \gamma + \frac{1}{(1-\alpha)N}\sum_{k=1}^{N}\left(f_x(\omega_k) - \gamma\right)^+. \tag{11}$$

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:

$$\begin{aligned}
&\min_{x,z,\gamma} \gamma + \frac{1}{(1-\alpha)N}\sum_{k=1}^{N} z_k \\
&s.t.\ z_k \geq 0, \quad k = 1,\ldots N \\
&\qquad z_k \geq f_x(\omega_k) - \gamma, \quad k = 1,\ldots N \\
&\qquad x \in X,
\end{aligned} \tag{12}$$

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

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 study opened the door to applying CVaR to financial optimization and risk management in practice.

Several robust models based on VaR and CVaR have been proposed in the literature to address portfolio asset allocation problems. We mention here the robust counterpart of the α-VaR optimization problem, proposed in Goldfarb and Iyengar (2003), where the asset returns are assumed to be normally distributed, and the alternative formulation described by El Ghaoui et al. (2003), where the authors introduce the notion of Worst-case VaR. Such a notion is generalized by Zhu and Fukushima (2006), who define the concept of Worst-case CVaR. We mention also two non standard robust models discussed in Bienstock (2007), which achieve an enhanced risk modeling flexibility by incorporating probability distribution and risk measure elements into the models. For more details see Recchia and Scutellà (2009).

3.2 The soft robust approach

In a recent study, Ben-Tal et al. (2009) proposed another framework for robust optimization which relaxes the classical notion of robustness. They focus on a relaxed approach in which not only the values of the uncertain parameters, but also their degree of feasibility are specified.

Let us formally 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 $x \in X$, 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 modelled here in terms of random returns). Consider the following probabilistic constraint related to the loss of x:

$$-\tilde{r}^T x \leq b, \tag{13}$$

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 (13) should take the form:

$$-r^T x \leq b \quad \forall r \in \mathcal{U} \tag{14}$$

where $\mathcal{U}$ denotes a given uncertainty set for the random return vector. Observe that, if $\mathcal{R}$ indicates the set of all possible return vectors, then constraint (14) could be equivalently rewritten as:

$$-r^T x \leq b + \beta(r) \quad \forall r \in \mathcal{R} \tag{15}$$

where $\beta(r) = 0$ if $r \in \mathcal{U}$ and $\beta(r) = +\infty$ otherwise.

The above introduced function β represents a particularly extreme penalty function (an indicator function of $\mathcal{U}$). The main observation of Ben-Tal et al. (2009) is that milder penalty functions could be used in constraints like (15). Alternative penalty functions would enable not only to control *where* there is feasibility, but would also indicate *how* feasible a particular realization of $\tilde{r}$ is. See also D. Brown's Ph.D. dissertation (2006), where the emphasis is put on the concept of penalty functions.

The use of milder penalty functions, to which correspond different decison-maker's risk preferences, is strictly related to the concept of convex risk measures, reviewed in Section 3.1, based on the following observation. Assume that $\mathcal{R}$ is the convex hull of a set of return vectors $r_1, \ldots r_N$, and so any return vector r can be expressed as $r = \sum\limits_{k=1}^{N} r_k q_k$, $\sum\limits_{k=1}^{N} q_k = 1$, $q_k \geq 0, k = 1, \ldots, N$. In addition, let $\mathcal{P}$ denote the set of all probability measures on the discrete set of scenarios $\Omega = \{r_1, ...r_N\}$, according to the notation introduced in Section 3.1. Then:

Theorem 6 *Let ρ be a convex risk measure, $\psi(q)$ be the penalty function associated with ρ according to Theorem 5 and $\tilde{r}$ be the random 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

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

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

$$\begin{aligned}
\rho(\tilde{r}^T x) \leq b &\Leftrightarrow \sup_{q \in \mathcal{P}} \left\{ E_q \left(-\tilde{r}^T x \right) - \psi\left(q\right) \right\} \leq b \quad \textit{(from Theorem 5)} \\
&\Leftrightarrow E_q \left(-\tilde{r}^T x \right) - \psi\left(q\right) \leq b, \quad \forall q \in \mathcal{P} \\
&\Leftrightarrow -\sum_{i=1}^{N} r_i^T x q_i - \psi\left(q\right) \leq b, \quad \forall q \in \mathcal{P} \\
&\Leftrightarrow -r^T x - \psi(q) \leq b, \quad \forall q \in \mathcal{P}(r),\ \forall r \in \mathcal{R} \\
&\Leftrightarrow -r^T x \leq b + \psi(q), \quad \forall q \in \mathcal{P}(r),\ \forall r \in \mathcal{R} \\
&\Leftrightarrow -r^T x \leq b + \inf_{q \in \mathcal{P}(r)} \psi\left(q\right), \quad \forall r \in \mathcal{R} \\
&\Leftrightarrow -r^T x \leq b + \beta(r), \quad \forall r \in \mathcal{R}.
\end{aligned}$$

■

Theorem 6 thus states that the relaxed notion of robustness introduced above corresponds to defining constraints, of probabilistic type, based on convex risk measures. Ben-Tal et al. (2009) analysed several penalty functions in such convex risk measure constraints, thus deriving several types of relaxed robustness. Among the others, they focused on the following constraints, where $\delta \geq 0$ is given:

$$\sup_{\{q:\psi(q)\leq\delta\}} (E_q(-\tilde{r}^T x) - \psi(q)) \leq b. \tag{16}$$

They proved that the left hand side of (16) is equivalent to:

$$\begin{aligned}
&\min_{c\geq 0} \left\{ c\delta + \sup_{q\in\mathcal{P}} (E_q(-\tilde{r}^T x) - (c+1)\psi(q)) \right\} = \\
&= \min_{c\geq 0} \left\{ c\delta + (c+1)\ \rho\left(\frac{\tilde{r}^T x}{c+1} \right) \right\}
\end{aligned} \tag{17}$$

where ρ denotes the convex risk measure induced by the penalty function ψ.

Moreover, if $b = 0$ then for each $\delta > 0$ any feasible portfolio x satisfying

$$\min_{c\geq 0} \left\{ c\delta + (c+1)\ \rho\left(\frac{\tilde{r}^T x}{c+1} \right) \right\} \leq 0 \tag{18}$$

belongs to the subset X^S_δ defined below, called the *soft robust set*:

$$X^S_\delta = \left\{ x \in X : \inf_{q\in\mathcal{P}(\varepsilon)} E_q \left[\tilde{r}^T x \right] \geq -\varepsilon \quad \forall \varepsilon \in [0, \delta] \right\}, \tag{19}$$

where $\mathcal{P}(\varepsilon) \subset \mathcal{P}$ define a sequence of convex sets of probability measures nondecreasing on $\varepsilon \geq 0$, such that $\mathcal{P}(\varepsilon) = \{q \in \mathcal{P} : \psi(q) \leq \varepsilon\}$ and $\mathcal{P}(0)$ is nonempty, and ρ is required to be a *normalized* convex risk measure. For details and formal proofs see (Ben-Tal et al. 2009).

Ben Tal et al. suggested various forms of *soft robustness*. In particular, they defined the so-called *ϕ-divergence penalty function*:

$$\psi(q) = \begin{cases} \sum_{i=1}^{N} p_i \phi\left(\frac{q_i}{p_i}\right) & \text{if } \sum_{i=1}^{N} p_i \phi\left(\frac{q_i}{p_i}\right) \leq \delta \\ +\infty & \text{otherwise} \end{cases} \tag{20}$$

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 of q with respect to the reference probability p, and it represents a distance-like measure from q to p. The authors considered the ϕ-divergence function $\phi : \phi(t) = t\ log(t) - t + 1$ and, scaling the corresponding penalty function by a positive factor γ, they studied the relative entropy from q to p scaled by γ, i.e. $\psi(q) = \frac{1}{\gamma} \sum_{i=1}^{N} log \frac{q_i}{p_i}$, by showing that the convex risk measure induced by the relative entropy is $\rho_\gamma(\tilde{r}^T x) = \frac{1}{\gamma} log(\sum_{i=1}^{N} p_i e^{-\gamma r_i^T x})$, also known as the *entropic risk measure at level γ*.

By choosing the entropic risk measure at level γ, then the equivalence

$$x \in X_\delta^S \Leftrightarrow \min_{c \geq 0} \left\{ c\delta + (c+1)\rho_\gamma \left(\frac{\tilde{r}^T x}{c+1} \right) \right\} \leq 0 \tag{21}$$

stated before, becomes:

$$x \in X_\delta^S \Leftrightarrow \min_{c \geq 0} \left\{ c\delta + (c+1)\ \frac{1}{\gamma} log\ \mathbb{E}_p \left(e^{\frac{-\gamma \tilde{r}^T x}{c+1}} \right) \right\} \leq 0 \tag{22}$$

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

Ben Tal et al. (2009) experimented some asset allocation problems related to the above notion of soft robustness. The formulation used for the computational analysis consists of maximizing the expected return of a portfolio x by imposing that the convex risk measure induced by a ϕ-divergence function is not greater than zero (or equivalently, according to (22), that x belongs to a specified soft robust set), i.e. :

$$\begin{aligned} &\max\ \mu^T x \\ &s.t.\ \rho(\tilde{r}^T x) \leq 0 \\ &\quad x \in X. \end{aligned} \tag{23}$$

or, equivalently:

$$\max \ \mu^T x$$
$$s.t.\ x \in X^S_\delta.$$

The authors considered the entropic risk measure and solved problem (23) by using monthly historical data related to 11 publicly traded asset classes, from April 1981 to February 2006. They compared the portfolios generated via the soft robust approach to the portfolios obtained by a variant of the soft approach, called the "comprehensive soft robust approach" and to portfolios generated by imposing that CVaR is less than or equal to zero. The main result is that, on the dataset considered, 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. This is expressed in terms of the realized return, for not too high a price in the increased downside risk, expressed in terms of realized CVaR.

4 A new family of models

Starting with the relationship between a relaxed form of robustness and convex risk measures, expressed by Theorem 6, we propose a new family of robust asset allocation models by defining penalty functions which are based on general norms. The aim is to define relaxed robust models that can be computed in a very efficient way. The proposed family includes in fact, as special cases, linear programming models and second order cone programming problems (SOCP). We also study a variant of the models for the coherent subcase, i.e. the case where the considered risk measures are also coherent.

Some of the proposed models have been tested with real market data, and compared to classical robust approaches from the literature as well as to the soft robust approach in Ben Tal et al. (2009). This will be the subject of Section 5.

4.1 The norm-portfolio models

Let us consider an environment with n risky financial securities. As in previous sections, let $x \in \mathbb{R}^n$ represent a portfolio of n securities, and x_j be the amount invested in security j. In addition, let μ denote the expected return vector, while $\tilde{r}$ be the vector of the random returns (so, $\tilde{r}^T x$ is the random return of portfolio x). Let us consider the following portfolio selection problem, where ρ denotes a convex measure of risk:

$$\begin{aligned} \inf \quad & \rho\left(\tilde{r}^T x\right) \\ s.t. \quad & \mu^T x \geq R \\ & x \in X. \end{aligned} \tag{24}$$

Let us reformulate problem (24) by introducing an auxiliary variable ζ:

$$\begin{aligned} &\inf\ \zeta \\ &s.t.\ \rho\left(\tilde{r}^T x\right) \leq \zeta \\ &\mu^T x \geq R \\ &x \in X. \end{aligned} \tag{25}$$

We can then exploit the relationship between the probabilistic constraint $\rho\left(\tilde{r}^T x\right) \leq \zeta$ and convex risk measures, expressed by Theorem 6. Assume to know a set $\{r_1, \ldots, r_N\}$ of return vectors, and $\mathcal{P}$ denote the set of all probability measures on this finite set of scenarios. This assumption captures a prevailing situation in many practical applications, when one has at his/her disposal N samples of the uncertain vector $\tilde{r}$, usually obtained from historical data. From Theorem 6, problem (25) is then equivalent to:

$$\inf \zeta \tag{26}$$

$$s.t.\ \sup_{q \in \mathcal{P}} \left\{ -\sum_{i=1}^{N} q_i r_i^T x - \psi\left(q\right) \right\} \leq \zeta \tag{27}$$

$$\mu^T x \geq R \tag{28}$$

$$x \in X. \tag{29}$$

Formulation (26)-(29) describes the family of robust models that we address, where ψ denotes the penalty function inducing the convex risk measure ρ. Note that, according to the observations in Section 3.2, constraint (27) captures a relaxed notion of robustness. This constraint states that the weight of the probability q depends on a suitably defined penalty function $\psi(q)$, 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 (26)-(29) where the penalty function $\psi(q)$ is defined in terms of an arbitrary norm $||\cdot||$. This will lead to the investigation of special convex risk measures in formulation (26)-(29). Since $\psi(q)$ is defined in terms of a norm, the particular convex risk measures under investigation will be called *norm-risk measures*, while the related models will be referred to as *norm-portfolio models*. The more general formulation of *norm-portfolio* models studied in this paper is the following:

$$\begin{aligned} &\inf \quad \zeta \\ &s.t.\ \sup_{q \in \mathcal{P}} \left\{ -\sum_{i=1}^{N} q_i r_i^T x - \lambda \left\|q, p\right\| \right\} \leq \zeta \\ &\mu^T x \geq R \\ &x \in X \end{aligned} \tag{30}$$

where $\|q, p\|$ denotes an arbitrary norm, defining a distance measure between the generic probability q and the reference probability p. The non-negative scalar λ is used to gauge this distance measure, as better explained below.

In this study, the following norms are addressed: the $||\cdot||_\infty$ norm, the $||\cdot||_1$ norm, the D-norm (Bertsimas et al. 2004) and the Euclidean norm. In the first three cases, the norm-portfolio models can be formulated as linear programming problems (LP); in the last case, i.e. when the penalty function is defined in terms of the Euclidean norm, the norm-portfolio models can be reduced to second order cone programming problems (SOCP).

4.1.1 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. Let $p_i = \frac{1}{N}$ for $i = 1, ...N$. Consider the vector $(p - q) \in \mathbb{R}^N$ and its *infinity norm* (also denoted by $||\cdot||_\infty$):

$$||p - q||_\infty = \max_i |p_i - q_i| . \tag{31}$$

The first family of norm-portfolio models we study is obtained by defining the penalty function $\psi(q)$ as follows:

$$\psi(q) = \lambda||p - q||_\infty = \lambda \max_i |p_i - q_i| \tag{32}$$

where, as indicated before, the parameter λ is a non-negative scalar used to gauge the distance between the probabilities. In other words, we use λ to assign a different weight to the probability q on the basis of its 'distance' from the reference probability p. Let us replace $\psi(q)$ in (26)-(29) by (32), obtaining the following formulation:

$$\begin{aligned} & \inf\ \zeta \\ s.t. \quad & \sup_{\substack{r=\sum_{i=1}^N r_i q_i \\ \sum_{i=1}^N q_i = 1 \\ q_i \geq 0, i=1,\ldots,N}} \left\{ -r^T x - \lambda \max_i |p_i - q_i| \right\} \leq \zeta \\ & \mu^T x \geq R, \quad x \in X. \end{aligned} \tag{33}$$

If we analyse the role of parameter λ in problem (33), then it is easy to observe that $\lambda = 0$ is the case where the investor gives the same weight (i.e. zero) to all probability measures q without considering their distance from the reference probability p; on the other hand, by increasing the value of λ the investor gives a smaller weight to those probabilities that are far from the reference probability. The extreme case $\lambda \to \infty$ is the case where the investor considers only the reference probability, i.e. $q = p$. In terms of robustness, the case $\lambda = 0$ describes a model more conservative and hence more robust. By increasing the value of parameter λ, the models become less conservative and hence theoretically less robust.

Problem (33) is not a linear programming model. However, it is well-known that the norm $||\cdot||_\infty$ can be linearized. Therefore:

Theorem 7 *Under the norm $||\cdot||_\infty$, the norm-risk measure leads to a LP model.*

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

Let us now consider the L_1 *norm* (also denoted by $\|\cdot\|_1$):

$$||p-q||_1 = \sum_{i=1}^{N} |p_i - q_i|. \tag{34}$$

The second family of *norm-portfolio* models we study is obtained by defining the penalty function $\psi(q)$ as follows:

$$\psi(q) = \lambda||p-q||_1 = \lambda \sum_{i=1}^{N} |p_i - q_i| \tag{35}$$

where, as before, the parameter λ is a non-negative scalar used to gauge the distance between the probabilities. Let us replace $\psi(q)$ in (26)-(29) by (35). Since also the $\|\cdot\|_1$ can be linearized, the following theorem holds true:

Theorem 8 *Under the L_1 norm, the norm-risk measure leads to a LP problem.*

4.1.3 The D-norm case

Here we address a particular norm, called the D-norm, which has been introduced by Bertsimas et al. (2004). We show that also under this scenario the norm-portfolio model can be reduced to a LP problem. We introduce the following definition of D-norm:

Definition 9 *Given a non-negative integer m ($m \le N$), the D-norm of vector $(p-q) \in \mathbb{R}^N$ is* [3]*:*

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

In other words, the D-norm of $p-q$ is the sum of the m largest absolute values of the entries of $p-q$. Therefore, setting $m = 1$ the D-norm coincides with the L_∞ norm. On the other hand, setting $m = N$ then the D-norm coincides with the L_1 norm, i.e. the L_1 and the L_∞ norms are special cases of the D-norm.

Let us consider (26)-(29), and set

$$\psi(q) = \lambda||p-q||_m, \quad \lambda \ge 0 \tag{36}$$

within the model, for a suitable m. The analogous of Theorems 7 and 8 is:

Theorem 10 *Under the D-norm, the norm-risk measure leads to a LP model.*

Proof. See the appendix. ■

[3] In their study, Bertsimas et al. address the general case where m is a non-negative real value.

4.1.4 The Euclidean norm case

Let us now consider the case where the penalty function $\psi(q)$ is described in terms of the Euclidean norm. In such a case, the *norm-portfolio* model can be formulated as a Second Order Cone Programming problem (SOCP).

Theorem 11 *Under the Euclidean norm, the norm-risk measure leads to a SOCP model.*

Proof. See the appendix. ■

4.2 Coherent variant of the norm-portfolio models

In this section we study the variant of the norm-portfolio models where the considered risk measure is also *coherent*. As reviewed in Section 3.1, a coherent risk measure arises from a family $\mathcal{Q}$ of probability measures by computing the worst expected loss when q varies in $\mathcal{Q}$, that is to say:

$$\rho(Y) = \sup_{q \in \mathcal{Q}} E_q[-Y], \quad Y \in \Phi. \tag{37}$$

By exploiting this characterization, the coherent version of the *norm-portfolio* family (30) can therefore be defined as follows:

$$\begin{aligned} & \inf \zeta \\ s.t. \quad & \sup_{\substack{r=\sum_{i=1}^N r_i q_i \\ \sum_{i=1}^N q_i = 1 \\ q_i \geq 0,\ i=1,\dots,N \\ ||p-q|| \leq \pi,}} -r^T x \leq \zeta \\ & \mu^T x \geq R, \quad x \in X \end{aligned} \tag{38}$$

where $||\cdot||$ denotes a generic norm and π is a parametric upper bound on the distance between the generic probability q and the reference probability p. Let us now specialize (38) by means of the infinity norm (in an analogous way we can specialize the family via the L_1 norm, the D-norm and the Euclidean norm):

$$\begin{aligned} & \inf \zeta \\ s.t. \quad & \sup_{\substack{r=\sum_{i=1}^N r_i q_i \\ \sum_{i=1}^N q_i = 1 \\ q_i \geq 0,\ i=1,\dots,N \\ ||p-q||_\infty \leq \pi,}} -r^T x \leq \zeta \\ & \mu^T x \geq R, \quad x \in X. \end{aligned} \tag{39}$$

As indicated before, π is an upper bound on the distance between the probability q and the reference probability p. Consequently, π belongs to the interval $[0, 1]$. The extreme values describe the case where the distance between the probabilities is null, i.e. only $q = p$ is taken into consideration (case $\pi = 0$), and the case where the entire set of probabilities is addressed (case $\pi = 1$). In terms of robustness, the case $\pi = 0$ is therefore the less conservative, and hence the less robust.

Observe that, as for the general convex case, (39) can be reduced to a Linear Programming Problem. Further observe that also the coherent variant implements a relaxed form of robustness, by choosing $\psi(q) = 0$ for all $q \in \mathcal{Q}$ (i.e. q such that $||p - q||_\infty \leq \pi$), and $+\infty$ otherwise.

4.2.1 Some considerations about the bound π

Consider parameter π in model (39). There exists one value of π that enables to establish an interesting relation between the coherent risk measure introduced in Section 4.2, based on the norm $|| \cdot ||_\infty$, and the well-known CVaR (which is also coherent).

As before, let N denote the number of the samples. In addition, let α be the confidence level chosen in the definition of the $CVaR_\alpha$ risk measure. Assume $\alpha \geq 1/2$ (so $(1-\alpha) \leq 1/2$), as usual in practical applications, and $(1-\alpha) \geq \frac{1}{N}$. Let η and τ be, respectively, the quotient of $(1-\alpha) : \frac{1}{N}$ and its rest. Then introduce the following value:

$$p^*_\alpha = \frac{p_i}{1-\alpha} + \frac{\tau}{\eta(1-\alpha)} = \frac{1}{N(1-\alpha)} + \frac{\tau}{\eta(1-\alpha)}. \tag{40}$$

By assuming $\eta < N$, define

$$\pi^*_\alpha = \left| \frac{1}{N} - p^*_\alpha \right|. \tag{41}$$

Let us consider the case where 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. Then:

Proposition 12 *If there are N scenarios, and all scenarios have the same probability, then setting $\pi = \pi^*_\alpha$ the coherent model (39) is equivalent to the $CVaR_\alpha$ model (in terms of optimal portfolio value).*

Proof. Let us recall the definition of VaR_α:

$$VaR_\alpha(x) = \min \left\{ \gamma : Prob(-\tilde{r}^T x \geq \gamma) \leq 1 - \alpha \right\}. \tag{42}$$

Let $r_1, \dots, r_N$ be the samples related to $\tilde{r}$, and $p_i = \frac{1}{N}$ denote the probability of the sample i, $i = 1, \dots, N$. Under these assumptions, $CVaR_\alpha(x)$ can be calculated as follows (Cornuejols and Tütüncü 2007):

$$CVaR_\alpha(x) = \frac{1}{1-\alpha} \sum_{i: -r_i^T x \geq VaR_\alpha(x)} \frac{1}{N}(-r_i^T x). \tag{43}$$

Let us distinguish two cases:

Case $\tau = 0$

Let us assign the probability $\tilde{p}_i = p^*_\alpha = \frac{1}{N(1-\alpha)}$ 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 $\pi \geq \pi^*_\alpha$ within model (39), then the probability above is feasible for the coherent model; in fact the following constraints are satisfied:

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

$$\begin{aligned}
&\Leftrightarrow \sum_{i=1}^{N-\eta} \tilde{p}_i + \sum_{i=N-\eta+1}^{N} \tilde{p}_i = 1 \\
&\Leftrightarrow 0 + (\eta)\frac{1}{N(1-\alpha)} = 1 \\
&\Leftrightarrow N(1-\alpha)\frac{1}{N(1-\alpha)} = 1
\end{aligned} \tag{44}$$

3) $\max_i |p_i - \tilde{p}_i| \leq \pi^*_\alpha$

   To prove the condition above, observe that, since $(1-\alpha) \leq 1/2$, then $\pi^*_\alpha \geq \frac{1}{N}$. Let us distinguish the following two subcases:

   a) $\tilde{p}_i = 0$:

   $$\max_i |p_i - 0| \leq \pi^*_\alpha \Leftrightarrow \frac{1}{N} \leq \pi^*_\alpha$$

   that is satisfied due to the observation above;

   b) $\tilde{p}_i \neq 0$:

   $$\begin{aligned}
   &\max_i |p_i - \tilde{p}_i| \leq \pi^*_\alpha \Leftrightarrow \\
   &\Leftrightarrow \left|\frac{1}{N} - \frac{1}{N(1-\alpha)}\right| \leq \left|\frac{1}{N} - \frac{1}{N(1-\alpha)}\right|
   \end{aligned}$$

   that is clearly satisfied.

Further observe that, by setting $\pi = \pi^*_\alpha$, then $\{\tilde{p}_i\}$ is the probability that maximizes the objective function of the inner problem of model (39). 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 to the scenarios with the highest loss. But this is exactly the probability addressed in (43). Hence, we can state that an optimal portfolio for model (39), under the choice $\pi = \pi^*_\alpha$, is optimal also for the $CVaR_\alpha$ model.

Case $\tau \neq 0$

The proof is the same of the case $\tau = 0$. In this case, assign the probability $\tilde{p}_i = p^*_\alpha = \frac{1}{N(1-\alpha)} + \frac{\tau}{\eta(1-\alpha)}$ to the η scenarios with the highest loss, and $\tilde{p}_i = 0$ otherwise. Also in this case, by setting $\pi = \pi^*_\alpha$ then $\{\tilde{p}_i\}$ is the probability that maximizes the objective function of the inner problem of (39). 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 to the scenarios with the highest loss. Since the η scenarios with the highest losses are exactly the scenarios addressed by the $CVaR_\alpha$ model, an optimal portfolio for (39) is also optimal for the $CVaR_\alpha$ model. Note, however, that the optimum objective function value of (39) does not necessarily coincide with the one returned by the $CVaR_\alpha$ model.

■

The following relationships can then be established between the coherent variant of the $||\cdot||_\infty$-portfolio models and the $CVaR_\alpha$ model:

Corollary 13 *Given α, the coherent variant of the $||\cdot||_\infty$-portfolio model generalizes the $CVaR_\alpha$ model; in addition, for $\pi \geq \pi^*_\alpha$ the coherent variant is more robust than $CVaR_\alpha$, in the sense that the considered probability set includes the one (implicitely) addressed by the $CVaR_\alpha$ model.*

5 The computational analysis

The computational analysis is composed of two parts. In the first part we analysed the behavior of the norm-portfolio models and their coherent variant when the main parameters of these models (i.e. λ and π) vary. As a result of the performed *in-sample* analysis, the $\|\ \|_\infty$-models have been selected for the successive phase, together with a pool of values for the parameters λ and π.

On the other hand, the objective of the second part was to contrast the performance of the selected norm-portfolio models with alternative robust selection strategies based on convex risk measures. The CVaR optimization model and the entropic model of Ben Tal et al. (2009) have been chosen as benchmark models in such a phase. Specifically, the objective of the computational study was to compare alternative forms of relaxed robustness - the relaxed robustness characterizing the norm-portfolio models and their coherent variant, the soft robustness characterizing the entropic model, and the more classical CVaR robustness - in order to assess how "simple" penalty functions such as the ones incorporated by the norm-portfolio models compare in practice to alternative penalty functions, and in particular to more sophisticated penalty functions such as the divergence based functions incorporated by the entropic model.

In addition, the models above have been compared to (a slight variant of) Tütüncü-Koenig model (Tütüncü and Koenig 2004), in order to compare robust models based on convex risk measures to more classical robust approaches. To the best of our knowledge, this is the first attempt at comparing robust portfolio selection strategies of different kinds (i.e. classical robust models based on uncertainty sets versus relaxed robust models based on convex risk measures).

Three real market data sets have been used to compare the robust models: two data sets provided by Tütüncü (Tütüncü and Koenig 2004), and one data set provided by Byrne (Byrne and Lee 2004). An *out-of-sample* analysis has been performed for each data set. The scenario information based on past history has been used to compare the values of the portfolios generated through the robust selection strategies under investigation, and to analyse their fluctuation over time. The details of the performed investigation will be illustrated next.

5.1 Plan of the experiments

5.1.1 The in-sample analysis

The aim of this analysis is to select suitable norm-portfolio models and suitable values for the parameters of these models. Concerning the values of parameter λ, we solved the norm-portfolio models for all integer values belonging to the interval $[0, 120]$ on the three data sets (the end value 120 has been selected on empirical ground). We considered also a further value in order to describe the performance of the models at infinity (i.e. $\lambda = 10^7$). By then observing the composition of the portfolios, their risk and their return when λ varies, we selected the following values of λ for the out-of-sample experiments: $\lambda = 0$, $\lambda = 5$, $\lambda = 10$, $\lambda = 15$, $\lambda = 20$, $\lambda = 25$, $\lambda = 30$, $\lambda = 35$ and $\lambda = 10^7$.

Concerning the coherent variant, we tested the models for values of parameter $\pi \in [0, 1]$, by then selecting suitable values of π on the basis of portfolio composition, risk and return. We selected the following values: $\pi = 0$ and $\pi = 1$ as extreme values, and $\pi = 0.25$, $\pi = 0.5$ and $\pi = \pi^*_\alpha$ as intermediate values. Remember that π^*_α is the value of parameter π that enables to establish an interesting relation between the coherent variant of the norm-portfolio models and the $CVaR_\alpha$ model (see 4.2.1).

5.1.2 The out-of-sample analysis

We compared the $|| \cdot ||_\infty$-model and its coherent variant to the entropic model, the $CVaR_\alpha$ model (by setting the confidence level $\alpha = 0.9$) and Tütüncü-Koenig model. Concerning Tütüncü-Koenig model, we studied the variant of model (6) where the uncertainty involves only the covariance matrix, and Q^U is assumed to be a positive semidefinite matrix. This leads to the following formulation:

$$\begin{aligned} \min\ & x^T Q^U x \\ s.t.\ & \mu^T x \geq R \\ & x \in X. \end{aligned} \tag{45}$$

The lower bound R in constraint $\mu^T x \geq R$ has been calculated as the mean value of vector μ. In addition, since parameter R is the minimum return that the investor would be willing to receive, in order to model a realistic behaviour we always chose a non-negative value for R. This setting of R is common to all implemented models.

The upper bound matrix Q^U has been generated through a method based on quantiles like in (Tütüncü and Koenig 2004), by considering moving windows of four years and computing the covariance matrix in each such window. Starting with the generated covariance matrices we then computed the 95 percentile of each element, and defined the matrix $\tilde{Q}$ that contains all such 95 percentiles. By construction, $\tilde{Q}$ is an upper bound of Q (Q corresponds in fact to the 50 percentile). However, nothing assures that $\tilde{Q}$ is a positive semidefinite (and hence a covariance) matrix. When this property does not hold, we thus solved the following subproblem that allows us to compute a covariance matrix "near" $\tilde{Q}$, which bounds $\tilde{Q}$ from above. Given the $n \times n$ matrix $\tilde{Q}$, the formulation of the subproblem is the following:

$$\begin{aligned} \min & \sum_{i=1}^{n}\sum_{j=1}^{n} q_{ij}^U \\ s.t.\ & Q^U \succeq 0 \\ & q_{ij}^U \geq \tilde{q}_{ij} \quad i = 1, \ldots, n \quad j = 1, \ldots, n. \end{aligned} \tag{46}$$

Problem (46) is a semidefinite programming problem that has been implemented in a MatLab 7.7 environment (R2008b) and solved under the Yalmip toolbox.

Concerning the soft robust approach, we studied the model that minimizes the entropic risk measure subject to the lower limit R on the expected return. The considered penalty function is thus the ϕ-divergence at level $\gamma = 1$ [4], i.e.:

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

where p denotes the reference probability. The loss function we used to define the entropic risk measure is $-\tilde{r}^T x$ ($\tilde{r}$ denotes the random return vector), the same we used in the norm-portfolio models and in the CVaR optimization model. Based on the equivalence (17), the investigated entropic model is therefore:

$$\begin{aligned} \min_{c,x}\ & c\ \delta + (c+1)log\left[\mathbb{E}_p\left(e^{\frac{-\tilde{r}^T x}{c+1}}\right)\right] \\ s.t.\ & \mu^T x \geq R \\ & x \in X \\ & c \geq 0 \end{aligned} \tag{47}$$

where $\delta = log\left(\frac{1}{\alpha}\right)$ (Ben Tal et al. 2009). (47) is a convex programming problem whose variables are c and x. As suggested in Ben Tal et al. (2009), we solved it by performing a binary search on c, by using the MatLab system cvx 6.1 for

[4] The choice of γ reflects the degree of the investor's risk-aversion. As observed in Ben-Tal et al. (2009), γ could be interpreted as the reciprocal of the risk tolerance for a CARA utility. A low value of γ thus corresponds to a high risk aversion.

convex optimization at each step of the binary search. As a result we determined an approximated value of c that minimizes the entropic objective function, then recovering the optimum portfolio x.

We performed the so-called out-of-sample analysis to compare the alternative robust strategies. For each data set, we divided the entire data sequence into 1-year investment periods via a moving window procedure and, for each investment period, we computed the optimum portfolios via the CVaR model, Tütüncü-Koenig model and the entropic model, as well as via the $||\cdot||_\infty$-model and its coherent variant (by using the values of the parameters chosen during the in-sample analysis). We then computed the "realized" return and variance of the portfolios generated in period t with the sample data available at $t+1$. As an example, by considering the first data set, we took the 12 monthly returns from January 1979 to December 1979 as the initial historical data for generating the first pool of portfolios (according to the robust strategies under investigation). We then calculated return and variance of those portfolios by observing the historical returns of the following month, i.e. January 1980. We then moved forth a month and considered the second investment period, from February 1979 to January 1980. We observed the historical returns corresponding to February 1980 and calculated return and variance of the portfolios generated, by repeating the procedure until the end of the data set is reached.

At the end of this process we thus generated $N-T$ portfolios for each investigated robust strategy, where N denotes the total numbers of monthly samples in the considered data set, while $T=12$ is the length of each investment period (expressed in months). We then evaluated the out-of-sample performance of each model according to the following statistics: *mean-realized return*, *variance* and *Sharpe Ratio*, and *portfolio turnover*. More formally, let x_t denote the portfolio generated at period t (according to a certain robust strategy). Then its realized return at time $t+1$ is $\hat{r}_{t+1} = x_t^T r_{t+1}$, where r_{t+1} denotes the historical returns at time $t+1$. After collecting the $N-T$ realized returns $\hat{r}_t$, we evaluated 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, according to the following definitions (DeMiguel and Nogales 2009):

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

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

The out-of-sample mean is the mean of the realized returns in the considered investment periods, while the out-of-sample variance is a measure of the variation of the realized returns with respect to the out-of-sample mean. The out-of-sample Sharpe Ratio thus estimates the mean realized return per unit of risk. Finally, the portfolio turnover is a measure of the portfolio fluctuation over time. Therefore, it can indirectly estimate the magnitude of the transaction costs associated with the investigated robust strategies.

Before presenting the results of the out-of-sample analysis in detail, we specify that all tested Linear Programming models, i.e. the $||\cdot||_\infty$-model, its coherent variants and the CVaR model (12), have been implemented using Tomlab/Cplex v11.2 within MatLab 7.7 (R2008b).

5.2 The first computational test

In the first experiment we used the universe of 5 asset classes in (Tütüncü and Koenig 2004): 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 (in percentages) of corresponding market indices (Russel 1000 growth, Russell 1000 value, Russell 2000 growth, Russell 2000 value and Lehman Brothers U.S. Governement/Credit Bond) from January 1979 to July 2002, i.e. a total of $N = 283$ months.

Firstly we applied the in-sample analysis to the norm-portfolio models. Concerning the $||\cdot||_\infty$-model we observed that, when λ increases, there is a slight increase of the realized return, but also severe drops in some cases. Concerning the coherent variant, in some cases there were more fluctuations when π increases, in contrast to the theoretical behaviour of this kind of models. These observations are confirmed by the statistics reported below.

$\|\|\cdot\|\|_\infty$-**model**	**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 1: *Out-of-sample mean, variance, Sharpe Ratio and portfolio turnover.*

Table 1 reports the out-of-sample mean, variance and Sharpe Ratio and the portfolio turnover related to the $||\cdot||_\infty$-model. Observe that the out-of-sample variance degrades significantly when λ increases. On the other hand, the out-of-sample mean increases, but in a less significant way. As a consequence, the Sharpe Ratio tends to decrease when λ increases. Concerning the turnover, it tends to increase when λ increases. This trend is in accordance to the theoretical results. The best results are obtained for small values of λ (i.e. for more robust scenarios). For example, the case $\lambda = 5$ is the best one in terms of both out-of-sample Sharpe Ratio and portfolio turnover. $\lambda = 5$ is thus the choice which produced the highest mean realized return per unit of risk, and which is the least expensive in terms of magnitude of the (indirectly addressed) transaction costs.

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
$\pi^*_\alpha = 0.92$	0.8951	6.8783	0.3413	0.3649
$\pi = 1$	0.8951	6.8783	0.3413	0.3649

Table 2: *Out-of-sample mean, variance, Sharpe Ratio and portfolio turnover.*

Concerning the coherent variant, the values of the statistics changed in a less significant way with respect to the $||\cdot||_\infty$-model. In addition, in contrast to the $||\cdot||_\infty$-strategy, the empirical behaviour of the coherent variant does not seem to confirm the theoretical expectation in terms of robustness. Table 2 shows in fact that the best behaviour is achieved for $\pi = 0.25$ both in terms of out-of-sample Sharpe Ratio and portfolio turnover. That is to say, the less robust scenario produced the best out-of-sample results. Observe also that the cases $\pi = \pi^*$ and $\pi = 1$ show an identical behaviour.

We then compared the best norm-portfolio models, (i.e. the scenarios $\lambda = 5$ and $\pi = 0.25$) to the other robust selection strategies. Table 3 provides the out-of-sample results for all considered models. It also provides the average computational time (in seconds) required to compute the optimum portfolio according to each strategy. The robust strategies based on convex risk measures

Models	**Mean** ($\hat{\mu}$)	**Variance** ($\hat{\sigma}^2$)	**Sharpe Ratio** ($\hat{SR}$)	**Turnover**	**Time**
$||\cdot||_\infty$**-model** ($\lambda = 5$)	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
Tütüncü-Koenig	0.9358	7.8912	0.3331	0.3116	0.1328
entropic	1.7223	1.2034	1.5700	0.4495	91.1330

Table 3: *Out-of-sample mean, variance, Sharpe Ratio, portfolio turnover and computational time for all models chosen.*

outperformed the more classical robustness incorporated by Tütüncü-Koenig model in terms of out-of-sample Sharpe Ratio. On the other hand, Tütüncü-Koenig model produced a sligthly less portfolio turnover. The main result for the first data set is thus that, by relaxing the robustness constraints in a flexible way via convex risk measures, one can potentially gain out-of-sample performance for not too high of a price in terms of portfolio variability.

More in detail, the entropic model proved to be the best model in terms of both out-of-sample mean and variance and, as a consequence, in terms of out-of-sample Sharpe Ratio. However, it was the most expensive in terms of turnover and hence, indirectly, in terms of transaction costs. Also the computational time required by this strategy is very high compared to the other robust strategies.

The $||\cdot||_\infty$-model and its coherent variant showed a quite similar behaviour, and they outperformed the CVaR model. It is interesting to observe that these models classified as "average" models for all the addressed statistics (by excluding the computational time, where they dominated), by showing that "simple" penalty functions such as the ones incorporated by the norm-portfolio models may improve the empirical Sharpe Ratio of classical robust approaches such as Tütüncü-Koenig's one and the one of CVaR, sometimes at a sligthly larger cost (expressed in terms of turnover). However, on the first data set, they proved to be not strong enough in terms of out-of-sample performance when compared to models based on more sophisticated penalty functions, such as the entropic model, although they are significantly less costly.

5.3 The second computational test

In the second experiment we used the wider set of market indices in Tütüncü and Koenig (2004): 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. Each category is represented by 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, from July 1983 to July 2002, i.e. a total of $N = 229$ months.

Tables 4 and 5 report the out-of-sample statistics for the $||\cdot||_\infty$-model and of its coherent variant.

$\|\|\cdot\|\|_\infty$**-model**	**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: *Out-of-sample mean, variance, Sharpe Ratio and portfolio turnover.*

Also in this experiment the out-of-sample variance degrades significantly when λ increases, whereas the out-of-sample mean increases. The out-of-sample

Sharpe Ratio strictly decreases. The scenario $\lambda = 0$ thus shows the best performance in terms of mean realized return per unit of risk, at the lowest cost when expressed in terms of portfolio turnover.

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
$\pi = \pi^*_\alpha = 0.92$	0.9250	4.6469	0.4291	0.4328
$\pi = 1$	0.9250	4.6469	0.4291	0.4328

Table 5: *Out-of-sample mean, variance, Sharpe Ratio and portfolio turnover.*

As before, the out-of-sample statistics of the coherent variant do not change in a significant way when π increases. However, in contrast to the first computational test, in this case the theoretical expectation in terms of robustness is confirmed by the empirical results. In fact, when π increases the model appears to be more robust as expressed by non decreasing Sharpe Ratios. In addition, when π increases the model appears to be less expensive in terms of portfolio turnover and hence in terms of transaction costs.

Now let us compare the alternative robust selection strategies in Table 6 (the scenarios $\lambda = 0$ and $\pi = \pi^*$ are reported for the norm-portfolio models). Also in this experiment the entropic model showed the best out-of-sample Sharpe Ratio but the highest cost in terms of portfolio turnover and computational time. In addition, Tütüncü-Koenig model proved to be better than the $||\cdot||_\infty$-model, its coherent variant and CVaR in terms of both out-of-sample Sharpe Ratio and turnover. Again, the $||\cdot||_\infty$-model and its coherent variant outperformed the CVaR model, and proved to be very efficient in terms of computational time.

Models	**Mean** ($\hat{\mu}$)	**Variance** ($\hat{\sigma}^2$)	**Sharpe Ratio** ($\hat{SR}$)	**Turnover**	**Time**
$||\cdot||_\infty$**-model** ($\lambda = 0$)	0.9250	4.6469	0.4291	0.4028	0.0469
coherent variant ($\pi^*_\alpha = 0.92$)	0.9250	4.6469	0.4291	0.4328	0.0156
CVaR	0.9085	4.6196	0.4227	0.4621	0.0156
Tütüncü-Koenig	0.9317	2.5412	0.5844	0.3924	0.1875
entropic	1.9583	1.6420	1.5282	0.4827	106.9952

Table 6: *Out-of-sample mean, variance, Sharpe Ratio, portfolio turnover and computational time.*

In conclusion, in the second data set the classical robustness incorporated by Tütüncü-Koenig model showed a better performance than the relaxed one incorporated by the $||\cdot||_\infty$-model and its coherent variant. These models, however, behaved better than CVaR. Also in this experiment the entropic model showed the best out-of-sample performance but, as before, this has been achieved at the highest empirical cost, expressed in terms of portfolio turnover, and at the highest computational cost, expressed in terms of average time required to compute a robust portfolio.

5.4 The third computational test

The data of the last experiment are used in Byrne (2004). They represent the total monthly returns of nine 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) and Industrials Southern and Eastern (ISE), from December 1987 to January 2002, i.e. a total of $N = 181$ monthly returns.

The computational results related to the $||\cdot||_\infty$-model and its coherent variant are reported in Tables 7 and 8, respectively.

$\|\| \cdot \|\|_\infty$**-model**	**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 7: *Out-of-sample mean, variance, Sharpe Ratio and portfolio turnover.*

In this experiment the values of the statistics did not vary in a significant way when λ (for the first model) and π (for the latter one) increases, even if the scenarios $\lambda = 0$ and $\pi = \pi^*_\alpha = 0.92$ slightly outperformed. Both models thus showed a greater stability with respect to their parameter variations. Table 9 summarizes the comparison among the investigated robust strategies in the third data set.

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
$\pi = \pi^*_\alpha = 0.92$	17.4706	103.1427	1.7202	0.1482
$\pi = 1$	17.4708	103.1414	1.7203	0.1489

Table 8: *Out-of-sample mean, variance, Sharpe Ratio and portfolio turnover.*

Models	Mean ($\hat{\mu}$)	Variance ($\hat{\sigma}^2$)	Sharpe Ratio ($\hat{SR}$)	Turnover	Time
$\|\| \cdot \|\|_\infty$**-model** ($\lambda = 0$)	17.4711	103.1393	1.7203	0.1482	0.025
coherent variant ($\pi^*_\alpha = 0.92$)	17.4706	103.1427	1.7202	0.1482	0.0258
CVaR	17.5269	104.3457	1.7158	0.1590	0.0257
Tütüncü-Koenig	11.1195	61.0785	1.4228	0.4472	0.1915
entropic	19.8666	124.1378	1.7831	0.1191	617.1563

Table 9: *Out-of-sample mean, variance, Sharpe Ratio, portfolio turnover and computational time.*

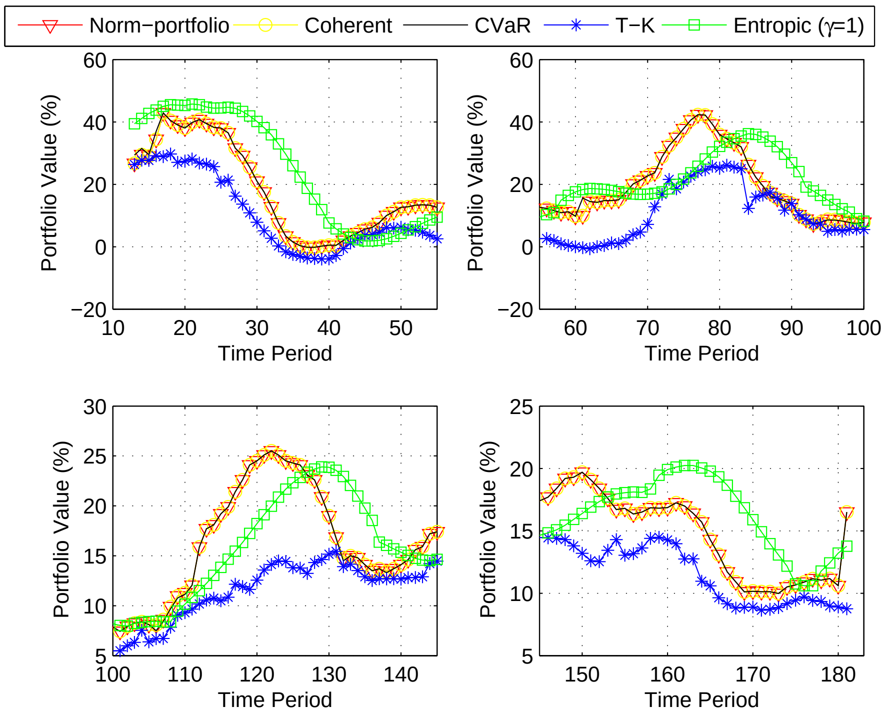
In this experiment Tütüncü-Koenig model showed the worst performance in terms of almost all criteria (with the exception of the out-of-sample variance). Again, the $||\cdot||_\infty$-model and its coherent variant outperformed CVaR for almost all statistics. In addition, also in this experiment the entropic model showed a better behaviour in terms of out-of-sample mean, Sharpe Ratio and turnover, and this was achieved at a very large computational cost. However, in this setting the norm-models and the entropic model showed a more similar behavior (at a very short computational time by considering the norm-portfolio models). In order to better understand similarities and dissimilarities of these two classes of relaxed models, we then plotted the historical trajectories of the generated optimum portfolio values in Figure 1.

According to the statistics in Table 9, the classical robust strategy incorporated by Tütüncü-Koenig model appears to be clearly inferior than the other robust strategies. Moreover, the curves of the $||\cdot||_\infty$-model, its coherent variant and CVaR are quite close; in most cases, in fact, the three trajectories overlap.

More interesting, by comparing the performance of the $||\cdot||_\infty$-model and the entropic model, we can note that, with the exception of few cases where the models generated approximately the same portfolio value (such as at times $t = 43$, $t = 104$, $t = 153$), there was not a clear superiority of one model with respect to the other one in terms of portfolio value. For example, in the period from $t = 13$ to $t = 43$ the entropic model produced the highest values of the portfolio, while from $t = 108$ to $t = 126$ there was an opposite behaviour, since the $||\cdot||_\infty$-model outperformed the entropic one in terms of portfolio value. Further observe the highest final portfolio value generated by the $||\cdot||_\infty$-strategy with respect to the one generated by the entropic strategy.

In concluding, the experimental results on the last data set showed that the relaxed robustness incorporated by the $||\cdot||_\infty$-model, its coherent variant and the entropic approach outperformed the classical robustness incorporated by Tütüncü-Koenig model not only in terms of portfolio value, but also in terms of costs. In fact, Tütüncü-Koenig model showed the worst performance at the highest cost in terms of portfolio turnover. Concerning the performance of the entropic model and of the $||\cdot||_\infty$-model, we can conclude that no clear

Figure 1: *Comparison among the chosen models*



superiority can be established for this data set; the entropic model produced in fact a sligthly better result in terms of Out-of-sample mean, Sharpe Ratio and turnover, but the analysis of the portfolio value trajectories revealed superiods where the $||\cdot||_\infty$-model outperformed the entropic one. In addition, if we focus on a long term investment strategy, it is interesting to note that the $||\cdot||_\infty$-model produced the best final portfolio value at the very lowest computational cost.

6 Conclusions

This paper provides the first attempt at comparing robust strategies of different kinds (i.e. standard robust models based on uncertainty sets versus relaxed robust models based on convex risk measures) in the framework of portfolio asset allocation. In the previous experiments, in fact, the objective was to contrast the performance of some classical portfolio selection strategies (usually the mean-variance and the minimum-variance approaches) with some specific robust selection strategies. Only in a few cases different robust methods of the same kind were compared. In addition, a new family of robust models has been proposed, which in some experiments proved to produce good results at a very low computational cost.

The main outcome of our experiments is that a relaxed robustness, such as the one incorporated by the $||\cdot||_\infty$-model, by its coherent variant and by the entropic approach, clearly outperforms the classical robustness incorporated by Tütüncü-Koenig model in terms of portfolio value. Among the relaxed robust approaches, the soft robust strategy incorporated by the entropic model always dominated in terms of out-of-sample mean portfolio value and variance, i.e. fluctuation over time. The entropic model thus classified as the most robust model in our experiments. However, this was generally achieved at the expenses of a greater portfolio variability, expressed in terms of turnover, and at a very large computational cost. The norm portfolio models, which always dominated the CVaR approach, in some cases showed an intermediate behavior, by producing good portfolio values and showing a small portfolio variability at the lowest computational cost.

Clearly, additional robust strategies have been proposed in the literature which would deserve investigation, and certainly we plan to enlarge the computational analysis to other strategies and other data sets. However, in our opinion, this computational investigation has the merit to have established some useful guidelines for further computational experiments, since it provides some preliminary indications on the relative empirical efficiency of robust selection strategies of different kind on different types of real market data.

References

[1] Artzner P, Delbaen F, Eber J, Heath D (1999) Coherent measures of risk. Math Finance 9:203-228

[2] Ben Tal A, Bertsimas D, Brown DB (2009) A soft robust model for optimization under ambiguity. Working paper

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

[4] Bienstock D (2007) Histogram Models for Robust Portfolio Optimization. J Comput Finance 11(1)

[5] Black F, Litterman R (1992) Global Portfolio Optimization. Financial Analysts J 48(5):28-43

[6] Broadie M, (1993) Computing efficient frontiers using estimated parameters. Ann Oper Res 45:21-58

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

[8] Chan LKC, Karceski J, Lakonishok J (1999) On portfolio optimization: Forecasting covariances and choosing the risk model. Rev Financial Stud 12(5):937-974

[9] Chopra VK, Ziemba WT. (1993) The effect of errors in means, variances, and covariances on optimal portfolio choices. J Portf Management 19(2):6-11

[10] Cornuejols G, Tütüncü RH (2007) Optimization Methods in Finance. Cambridge University Press

[11] DeMiguel V, Nogales FJ (2009) Portfolio Selection with Robust Estimation. Oper Res 57 (3):560-577

[12] El Ghaoui L, Oks M, Oustry F (2003) Worst Case Value-at-Risk and robust portfolio optimization: A conic programming approach. Oper Res 51(4):543-556

[13] Fabozzi FJ, Kolm PN, Pachamanova DA, Focardi SM (2007) Optimization and Management. John Wiley and Sons, Inc

[14] Föllmer H, Schied A (2002) Convex measures of risk and trading constraints. Finance and stoch 6:429-447

[15] Föllmer H, Schied A (2004) Stochastic Finance: An Introduction in Discrete Time. Walter de Gruyter, Berlin.

[16] Goldfarb D, Iyengar G (2003) Robust Portfolio Selection. Math Oper Res 28(1):1-38

[17] Halldórsson BV, Tütüncü RH (2003) An Interior-Point Method for a Class of Saddle-Point Problems. J Optim Theory and App 116(3):559-590

[18] Jagannathan R, Ma T (2003) Risk reduction in large portfolios: Why imposing the wrong constraints help?. J Finance 58(4):1651-1684

[19] Markowitz H (1952) Portfolio Selection. J Finance 7:77-91

[20] Michaud RO (1989) The Markowitz optimization enigma: is optimized optimal?. Financial Analysts J 45(1):31-42

[21] R. Recchia and M.G. Scutellà, *Robust portfolio asset allocation and risk measure*, Submitted to 4OR, 2009.

[22] Tütüncü RH, Koenig M (2004) Robust Asset Allocation. Ann Oper Res 132:157-187

[23] Zhu SS, Fukushima M (2006) Worst-Case Conditional Value at Risk with Application to Robust Portfolio Management. To appear in Oper Res

Appendix

Proof. (*Theorem 10*) Let us consider the inner problem

$$\begin{aligned} &\sup -r^T x - \lambda ||p - q||_m \\ &s.t.\ r = \sum_{i=1}^{N} r_i q_i \\ &\sum_{i=1}^{N} q_i = 1 \\ &q_i \geq 0, i = 1, \ldots, N. \end{aligned} \tag{48}$$

and introduce an auxiliary variable z in order to bound the D-norm from above. We can then state the following equivalent formulation:

$$\begin{aligned} &\sup -\sum_{i=1}^{N} q_i r_i^T x - \lambda z \\ &s.t.\quad \|p - q\|_m \leq z \\ &\sum_{i=1}^{N} q_i = 1 \\ &q_i \geq 0 \quad i = 1, \ldots N. \end{aligned} \tag{49}$$

We can replace the *sup* operator by the *max* operator since the supremum of a linear function on a closed and bounded set is always attained. Then, by exploiting the definition of D-norm, we get the following formulation:

$$
\begin{aligned}
& \max -\sum_{i=1}^{N} q_i r_i^T x - \lambda z \\
s.t. \quad & \max_{\substack{S\subseteq\{1,\ldots,N\}\\ |S|\leq m}} \left\{ \sum_{i\in S} |p_i - q_i| \right\} \leq z \\
& \sum_{i=1}^{N} q_i = 1 \\
& q_i \geq 0 \quad i = 1,\ldots,N.
\end{aligned}
\tag{50}
$$

By introducing additional auxiliary variables φ_i in order to model the absolute values $|p_i - q_i|$, then we have:

$$
\begin{aligned}
& \max -\sum_{i=1}^{N} q_i r_i^T x - \lambda z \\
s.t. \quad & \max_{\substack{S\subseteq\{1,\ldots,N\}\\ |S|\leq m}} \sum_{i\in S} \varphi_i \leq z \\
& p_i - q_i \leq \varphi_i, \qquad i = 1,\ldots,N \\
& -(p_i - q_i) \leq \varphi_i, \qquad i = 1,\ldots,N \\
& \sum_{i=1}^{N} q_i = 1 \\
& q_i \geq 0 \quad i = 1,\ldots,N.
\end{aligned}
\tag{51}
$$

The constraint

$$
\max_{\substack{S\subseteq\{1,\ldots,N\}\\ |S|\leq m}} \sum_{i\in S} \varphi_i \leq z \tag{52}
$$

includes the second inner problem of the given formulation in its left-hand side. By interpreting $\{\varphi_i\}$ as constant values, $i = 1,\ldots,N$, then the second inner problem can be formulated as the following knapsack problem:

$$
\begin{aligned}
& \max \sum_{i=1}^{N} \varphi_i y_i \\
s.t. \quad & \sum_{i=1}^{N} y_i \leq m \\
& y_i \in \{0,1\}, \quad i = 1,\ldots,N
\end{aligned}
\tag{53}
$$

Let us consider the linear relaxation of problem (53):

$$
\begin{aligned}
&\max \sum_{i=1}^{N} \varphi_i y_i \\
&s.t. \ \sum_{i=1}^{N} y_i \leq m && \rightarrow (\pi) \\
&\quad y_i \leq 1 \quad i=1,\ldots,N && \rightarrow (\pi_i) \\
&\quad y_i \geq 0 \quad i=1,\ldots,N.
\end{aligned}
\tag{54}
$$

In such a special knapsack problem, the linear relaxation provides the optimum objective function value of (53). Since the feasible set of the linear relaxation is bounded and not empty, from strong duality we can replace the linear relaxation by its dual:

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

Therefore, the inner problem can be equivalently rewritten as :

$$
\begin{aligned}
&\max - \sum_{i=1}^{N} q_i r_i^T x - \lambda z \\
&s.t. \ m\pi + \sum_{i=1}^{N} \pi_i - z \leq 0 && \rightarrow (\delta) \\
&\quad -\pi - \pi_i + \varphi_i \leq 0, \quad i=1,\ldots,N && \rightarrow (v_i) \\
&\quad \pi \geq 0, \ \pi_i \geq 0, \quad i=1,\ldots,N \\
&\quad -q_i - \varphi_i \leq -p_i, \quad i=1,\ldots,N && \rightarrow (w_i^+) \\
&\quad q_i - \varphi_i \leq p_i, \quad i=1,\ldots,N && \rightarrow (w_i^-) \\
&\quad \sum_{i=1}^{N} q_i = 1 && \rightarrow (u) \\
&\quad q_i \geq 0, \ i=1,\ldots,N
\end{aligned}
\tag{55}
$$

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

By replacing the inner problem (55), which is not empty and whose objective function is bounded from above, by its dual, we get a linear programming formulation. ■

Proof. (*Theorem 11*) Let us consider the inner problem where, as standard, variables x_i are treated as constant values:

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

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

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

Let us replace the *sup* operator by the *max* operator, and set $y = p - q$. In this way, we can substitute $q = p - y$ within the formulation. Then, since $p_i = \frac{1}{N}$ for each i, 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 is equivalent to state $(z, y) \in C_q$ where, according to the notation in (Cornuejols and Tütüncü 2007), C_q denotes a second order cone.

By introducing additional auxiliary variables s_i we get:

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

Since the problem is feasible ($y = 0, z = 0, s_i = \frac{1}{N}$ $\forall i$ is in fact a feasible solution), from the Conic duality theorem (Ben-Tal and Nemironski 2001) the optimum objective function value is equal to the one of its dual problem:

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

where C_q^* and C_l^* denote the dual cones of C_q and C_l, respectively (Ben-Tal and Nemironski 2001).

Since the dual cone of a second order cone is a second order cone, then the first two groups of constraints (i.e. the ones related to the dual cone C_q^*) can be equivalently rewritten as:

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

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

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

which is a SOCP model. ∎