



## TECHNICAL REPORT

CTR/33/04

September 2004

**Portfolio Optimisation Models and Properties of  
Return Distributions**

Diana Roman, Ken Darby-Dowman,  
Gautam Mitra

# Portfolio Optimisation Models and Properties of Return Distributions

Diana Roman, Ken Darby-Dowman\*, Gautam Mitra

*CARISMA: The Centre for the analysis of risk and Optimisation Modelling Applications*

*Department of Mathematical Sciences  
Brunel University, Uxbridge, Middlesex, UB8 3PH, UK*

---

## Abstract

Mean-risk models have been widely used in portfolio optimisation. However, such models may produce portfolios that are dominated with respect to second order stochastic dominance and therefore not optimal for rational and risk-averse investors. This paper considers the problem of constructing a portfolio which is nondominated with respect to second order stochastic dominance and whose return distribution has specified desirable properties. The problem is multi-objective and is transformed into a single objective problem by using the reference point method, in which target levels, known as aspiration points, are specified for the objective function values. A model is proposed in which the aspiration points relate to ordered return outcomes of the portfolio return. The model is extended by additionally specifying reservation points, which act pre-emptively in the optimisation. The theoretical properties of the models are studied. The performance of the models on real data drawn from the Hang Seng index is also investigated.

**Keywords:** portfolio selection; stochastic dominance; risk aversion; multi-criteria optimisation; reference point method.

---

---

\* Corresponding author: [Kenneth.Darby-Dowman@brunel.ac.uk](mailto:Kenneth.Darby-Dowman@brunel.ac.uk)

## 1. Introduction and motivation

The portfolio selection problem with one investment period is a particular case of the general problem of choosing between random variables when larger outcomes are preferred. There is, therefore, a need for models that make a choice between random variables (models for preference). Basically, there are three main types of such models: mean-risk models, expected utility maximisation and stochastic dominance models.

Mean-risk models are convenient from a computational point of view, but in many cases, depending on the risk measure that is used, they lack a rational, theoretical basis for making a choice. Thus, the validity of the results provided is often questionable. Expected utility maximisation is usually difficult to put into practice since the choice of a suitable utility function is somewhat subjective. Stochastic dominance has a sounder theoretical basis for making choice under risk but unfortunately is very demanding from a computational point of view.

In this paper, we present a model for portfolio selection, which combines the advantages of the previously mentioned approaches, in the sense that it provides a meaningful solution, corresponding to observed economic behaviour, and is practical from a computational point of view. The solution is meaningful in the sense that the selected portfolio is nondominated with respect to second order stochastic dominance (SSD). It is therefore optimal for every rational and risk averse investor. In addition, this portfolio has a return distribution close to a user-specified target (reference) distribution.

In solving the model, there are three possible cases to consider. Firstly, if the reference distribution (appealing as it may seem) is not optimal with respect to SSD, the model provides a solution, which comes close to the reference, but is optimal with respect to SSD. In this case, the achieved return distribution is “better than target”. In the second case, if the reference distribution is optimal with respect to SSD, the model provides a solution, which yields exactly the reference distribution. Finally, the reference distribution may not be attainable (in the sense that some of its outcomes are too high such that there are no feasible portfolio weights that could produce such a portfolio distribution). In this case, the model provides a SSD efficient portfolio whose return distribution comes uniformly close to the reference distribution.

The rest of the paper is organised as follows. In section two, models for choice under risk are reviewed, together with their advantages and disadvantages. Mean-risk models, expected utility maximisation and stochastic dominance models are considered. Section three presents the mathematical formulation of the proposed model. The SSD efficient solutions in the portfolio selection problem are Pareto efficient solutions of a multi-objective linear problem. In order to select a specific solution to this problem (the one that comes closest to the reference distribution), the reference point method (see Wierzbicki 1982, Steuer 1986, Makowski and Wierzbicki 2003) is used. This is an interactive method used in multi-objective programming for generating an efficient solution, in which desired levels for the objective functions (called aspiration points) are introduced. The multi-objective problem is transformed into a single objective optimisation problem by constructing a scalarizing achievement function, which is then optimised. We present a model where the aspiration points relate to the ordered return outcomes of the portfolio. This model is then extended by additionally introducing reservation points, which act pre-emptively in the optimisation. Section four presents the

computational results. A dataset drawn from the Hang Seng index with price history for 20 stocks over 131 months is used to explore and evaluate the behaviour of the models. Section five presents the conclusions.

## 2. Preliminaries: Models for choice under risk

### 2.1. The portfolio selection problem

The problem of portfolio selection with one investment period is an example of the general problem of deciding between random variables when larger outcomes are preferred. Decisions are required on the amount (proportion) of capital to be invested in each of a number of available assets such that at the end of the investment period the return is as high as possible. Consider a set of  $n$  assets, with asset  $j$  in  $\{1, \dots, n\}$  giving a return  $R_j$  at the end of the investment period.  $R_j$  is a random variable, since the future price of the asset is not known. Let  $x_j$  be the proportion of capital invested in asset  $j$  ( $x_j = w_j/w$  where  $w_j$  is the capital invested in asset  $j$  and  $w$  is the total amount of capital to be invested), and let  $x = (x_1, \dots, x_n)$  represent the portfolio resulting from this choice. This portfolio's return is the random variable:  $R_x = x_1 R_1 + \dots + x_n R_n$ , with distribution function  $F(r) = P(R_x \leq r)$  that depends on the choice  $x = (x_1, \dots, x_n)$ .

To represent a portfolio, the weights  $(x_1, \dots, x_n)$  must satisfy a set of constraints that forms a feasible set  $M$  of decision vectors. The simplest way to define a feasible set is by the requirement that the weights must sum to 1 and short selling is not allowed. For this basic version of the problem, the set of feasible decision vectors is

$$M = \{ (x_1, \dots, x_n) / \sum_{j=1}^n x_j = 1, x_j \geq 0, \forall j \in \{1, \dots, n\} \}.$$

Consider a different portfolio defined by the decision vector  $y = (y_1, \dots, y_n) \in M$ , where  $y_j$  is the proportion of capital invested in asset  $j$ . The return of this portfolio is given by the random variable  $R_y = y_1 R_1 + \dots + y_n R_n$ .

The problem of choosing between portfolio  $x = (x_1, \dots, x_n)$  and portfolio  $y = (y_1, \dots, y_n)$  becomes the problem of choosing between random variables  $R_x$  and  $R_y$ . The criteria by which one random variable is considered "better" than another random variable need to be specified and models for choosing between random variables (models for preference) are required. The purpose of such models is firstly, to define a preference relation among random variables and secondly, to identify random variables that are non-dominated with respect to that preference relation.

The next issue is to consider a practical representation for the random variables that describe asset and portfolio returns. We treat these random variables as discrete and described by realisations under  $T$  states of the world, generated using scenario generation or finite sampling of historical data. Let state  $i \in \{1, \dots, T\}$  occur with probability  $p_i$ ,

$\sum_{i=1}^T p_i = 1$ . Let  $r_{ij}$  be the return of asset  $j$  under scenario  $i$ ,  $i \in \{1, \dots, T\}$ ,  $j \in \{1, \dots, n\}$ . Thus, the random variable  $R_j$  representing the return of asset  $j$  is finitely distributed over  $\{r_{1j}, \dots, r_{Tj}\}$  with probabilities  $p_1, \dots, p_T$ . The random variable  $R_x$  representing the return of

portfolio  $x=(x_1,\dots,x_n)$  is finitely distributed over  $\{R_{x1},\dots, R_{xT}\}$ , where  $R_{xi}=x_1r_{i1}+\dots+x_n r_{in}$ ,  $\forall i\in \{1,\dots,T\}$ .

The main approaches for modelling choice among random variables: mean-risk models, expected utility maximisation and stochastic dominance models are now discussed.

## 2.2. The mean-risk model

In mean- risk models, two scalars are attached to each random variable: the expected value (mean) and the value of a risk measure. Preference is then defined using a trade-off between the mean where a larger value is desirable and risk where a smaller value is desirable.

*Definition:* In the mean-risk approach with the risk measure denoted by  $\rho$ , random variable  $R_x$  dominates (is preferred to) random variable  $R_y$  if and only if:  $E(R_x)\geq E(R_y)$  and  $\rho(R_x)\leq\rho(R_y)$  with at least one strict inequality.

In this approach, the choice  $x$  (or the random variable  $R_x$ ) is efficient (non-dominated) if and only if there is no other choice  $y$  such that  $R_y$  has higher expected value and less risk than  $R_x$ . This means that, for a given level of expected return,  $R_x$  has the lowest possible risk, and, for a given level of risk, it has the highest possible expected return. Plotting the efficient portfolios in a mean-risk space gives the efficient frontier.

In order to find efficient portfolios, we solve an optimisation problem with decision variables  $x_1,\dots,x_n$ :

Minimise  $\rho(R_x)$   
 Subject to:  $E(R_x)\geq d$  and  $(x_1,\dots,x_n)\in M$ ,

Where  $d$  represents the desired level of expected return for the portfolio.

Varying  $d$  and repeatedly solving the corresponding optimisation problems produces the efficient frontier of the mean-risk model.

## 2.3. Expected Utility Maximisation

In expected utility maximisation, a single scalar value (the expected utility) is attached to each random variable. Preference is then defined by comparing expected utilities with a larger value preferred.

A utility function is a real-valued function defined on real numbers (representing possible outcomes of the random variable), which measures the relative value of outcomes.

Expected Utility Theory provides a basis for extending a utility function defined on real numbers (outcomes) to a utility function defined on random variables. A utility

value is assigned to each random variable in terms of the utility values of its outcomes and the probabilities associated with these outcomes.

*Definition:* Given a utility function  $U$ , the expected utility of a random variable  $R_x$  is:  $E[U(R_x)] = \int_{-\infty}^{\infty} U(\alpha) dF(\alpha)$ ,

where  $F$  is the distribution function of  $R_x$ .

In the discrete case, if the random variable  $R_x$  has the outcomes  $\alpha_1, \dots, \alpha_T$  with probabilities  $p_1, \dots, p_T$ , the expected utility of  $R_x$  is:  $E[U(R_x)] = p_1 U(\alpha_1) + \dots + p_T U(\alpha_T)$ .

The use of utility functions is based on an implied assumption that they reflect the behaviour of investors. There are progressively stronger assumptions about the behaviour of investors, which lead to progressively stronger forms of utility functions.

The first assumption on investors' behaviour is that they prefer more to less and are hence rational. In order to reflect this non-satiation attitude, the utility function must be nondecreasing. This is the only non-arguable condition for utility functions on wealth since all investors are assumed to be rational.

The second aspect to consider is investors' attitudes towards risk. There are three possible attitudes: risk aversion, risk neutrality and risk seeking. The general assumption is that investors are risk averse, which means that, as wealth increases, each additional increment is less valuable than the previous one. In order to reflect this, the utility function must be nondecreasing and concave (see Whitmore and Findlay 1978).

Further assumptions on investors' behaviour (e.g. ruin aversion and decreasing absolute risk aversion) are also discussed in the literature (e.g. Whitmore and Findlay 1978). These attitudes are less representative of investors' behaviour and are not considered further in this paper.

## 2.4. Stochastic Dominance

A difficulty with expected utility maximisation is that the specification of a utility function is a subjective task. Stochastic dominance (or stochastic ordering of random variables) was introduced, in order to overcome this difficulty (see Whitmore and Findlay 1978). Its purpose is to rank choices (random variables) under assumptions about general characteristics of utility functions that follow from prevalent modes of economic behaviour. Stochastic dominance ensures that all individuals, whose utility functions are in the same class, rank choices in the same way.

There are progressively stronger assumptions about investors' behaviour that are used in the stochastic dominance approach, leading to first, second, and higher orders of stochastic dominance relations. In this paper, we consider only first and second order stochastic dominance relations:

- *First order stochastic dominance (FSD)* assumes that an investor prefers more to less.  $R_x$  is preferred to  $R_y$  with respect to FSD (Notation:  $R_x >_{FSD} R_y$ ) if and only if

$E[U(R_x)] \geq E[U(R_y)]$ , for every nondecreasing utility function  $U$ . In this case, all rational investors prefer  $R_x$  to  $R_y$ .

- *Second order stochastic dominance (SSD)* additionally assumes that an investor is risk-averse.  $R_x$  is preferred to  $R_y$  with respect to SSD (Notation:  $R_x > SSDR_y$ ) if and only if  $E[U(R_x)] \geq E[U(R_y)]$ , for every nondecreasing and concave utility function  $U$ . In this case, all rational and risk-averse investors prefer  $R_x$  to  $R_y$ .

SSD is stronger than FSD in the sense that it is able to order more pairs of random variables. We could have indifference between  $R_x$  and  $R_y$  with respect to FSD but prefer  $R_x$  or  $R_y$  with respect to SSD.  $R_x > FSDR_y \Rightarrow R_x > SSDR_y$ . Thus, an efficient solution with respect to SSD will be also an efficient solution with respect to FSD.

In practice we compare two random variables with respect to stochastic dominance relations by pointwise comparison of some performance functions constructed from their distribution functions. Let  $F_x$  and  $F_y$  be the distribution functions of  $R_x$  and  $R_y$  respectively.

Proposition:

$R_x > FSDR_y$  if and only if  $F_x(r) \leq F_y(r)$ , for every  $r$ .

$R_x > SSDR_y$  if and only if  $F_x^{(2)}(r) \leq F_y^{(2)}(r)$ , for every  $r$ ,

Where: 
$$F_x^{(2)}(r) = \int_{-\infty}^r F_x(t) dt, \forall r \in R$$

(for a proof see Whitmore and Findlay 1978).

Performance functions for higher order stochastic dominance are similarly defined (see for example Ogryczak and Ruszczyński 1999, 2001).

Consider the case where the two random variables  $R_x$  and  $R_y$  are discrete with  $T$  equally probable outcomes. In order to make the comparison, we rank the outcomes of  $R_x$  and  $R_y$  in ascending order to obtain two  $T$ -dimensional vectors:  $(\alpha_1, \dots, \alpha_T)$  and  $(\beta_1, \dots, \beta_T)$  such that  $\alpha_1 \leq \dots \leq \alpha_T$  and  $\beta_1 \leq \dots \leq \beta_T$ .

$R_x > FSDR_y$  if and only if  $\alpha_i \geq \beta_i$  for every  $i \in \{1, \dots, T\}$  with at least one strict inequality.

$R_x > SSDR_y$  if and only if  $\sum_{j=1}^i \alpha_j \geq \sum_{j=1}^i \beta_j$  for every  $i \in \{1, \dots, T\}$  with at least one strict inequality.

Consider a simple example:

$R_x$ :  $P(R_x=1) = P(R_x=4) = P(R_x=3) = P(R_x=2) = 1/4$

$R_y$ :  $P(R_y=3) = P(R_y=5) = P(R_y=0) = P(R_y=2) = 1/4$ .

Ordering the outcomes of  $R_x$  and  $R_y$  we obtain two 4-dimensional vectors: (1,2,3,4) and (0,2,3,5) respectively. Neither  $R_x$  nor  $R_y$  dominates with respect to FSD. If we cumulate the outcomes of these vectors, we obtain the vectors: (1,3,6,10) and (0,2,5,10), respectively. Since all the components of the first vector are greater than or equal to those of the second one, it follows that  $R_x$  dominates  $R_y$  with respect to SSD.

This was to be expected, since  $R_x$  has a higher outcome than  $R_y$  at the lowest level at the expense of a lower outcome at the highest level. Thus, a risk-averse investor will prefer  $R_x$  to  $R_y$ .

## 2.5. A summary comparison of models for choice under risk

### Mean-risk models

Mean-risk models are practical from a computational point of view. However, depending on the risk measure to be used, they may lack a rational basis for making choices. Consider an example to support this statement. Consider a mean-risk model with a symmetric risk measure such as the variance and two random variables  $X$  and  $Y$ .  $X$  has one certain outcome  $W_0$ .  $Y$  has 2 outcomes:  $W_0$  and  $W_1$ , each with probability  $\frac{1}{2}$ . Let  $W_1 > W_0$ . Obviously  $Y$  should be preferred since it yields at least as much as  $X$ . However, according to a mean-variance model, neither one is preferred, since, although  $X$  has a smaller expected value, it also has zero risk.

### Expected utility maximisation

As discussed earlier, with expected utility maximisation it is generally difficult to objectively specify an appropriate utility function.

### Stochastic dominance models

Stochastic dominance relations are the soundest theoretical basis for making rational choices, corresponding to observed economic behaviour. Unfortunately, they are difficult to apply in practice. Finding the entire set of efficient solutions with respect to a stochastic dominance relation is practically intractable, when the number of random variables to be compared is not finite (as in the case of portfolio selection problem).

The model that we present here selects a solution that is efficient with respect to second order stochastic dominance and is therefore a rational basis for making choice. In addition, the selected solution is not an arbitrary SSD efficient solution but one whose distribution is as close as possible to a specified distribution.

## 3. A model for portfolio selection

### 3.1. SSD efficiency and multi-objective optimisation

Consider a portfolio  $x=(x_1, \dots, x_n)$  with return  $R_x$  and the case of  $T$  scenarios occurring with probabilities  $p_1, \dots, p_T$ . The realisation of  $R_x$  under scenario  $i \in \{1, \dots, T\}$  is  $x_1r_{i1} + \dots + x_nr_{in}$ , where  $r_{ij}$  is the return of asset  $j$  under scenario  $i$ . Consider further, the case of equally probable scenarios:  $p_i = \frac{1}{T}$  for  $i \in \{1, \dots, T\}$ . Since we search for solutions  $x$  such that  $R_x$  has outcomes as large as possible under each scenario, the portfolio selection problem may be viewed as a multi-objective problem with  $T$  objective functions to maximise:

$$\text{Max } \{f(x)=(f_1(x), \dots, f_T(x)): x \in M\}, \quad (1)$$



where:  $f_i(x) = x_1 r_{i1} + \dots + x_n r_{in}$  is the portfolio return in scenario  $i \in \{1, \dots, T\}$ ,  
 $M$  is the set of feasible decision vectors.

*Definition:*  $y \in \mathbb{R}^T$  is an achievement vector if there exists a feasible solution  $x \in M$  such that  $y = f(x)$ .

Usually, in multi-objective optimisation, two achievement vectors are compared using the Pareto preference relation, defined as follows:

An achievement vector  $y = (y_1, \dots, y_T)$  *Pareto dominates* another vector  $y' = (y'_1, \dots, y'_T)$  if  $y_i \geq y'_i$  for every  $i \in \{1, \dots, T\}$ , with at least one strict inequality.

$x_0$  is a Pareto efficient (non-dominated) solution of (1) if and only if there does not exist a feasible  $x$  such that  $f(x)$  Pareto dominates  $f(x_0)$ . In other words, a Pareto efficient solution is a feasible solution such that, in order to improve upon one objective function, at least one other objective function must assume a worse value.

However, in the portfolio optimisation problem, in addition to Pareto dominance, there are other considerations. Thus, we do not look for simply Pareto efficient solutions.

Firstly, an achievement vector represents a random variable with  $T$  equally probable outcomes for which the ordering is immaterial. We do not compare vectors, but the distributions that these vectors represent. We compare distributions with respect to the first order stochastic dominance relation (FSD) by comparing the  $k$ -th worst outcomes of the achievement vectors, for  $k=1, \dots, T$ . Thus, the distribution represented by  $y$  dominates that of  $y'$  with respect to FSD if the  $k$ -th worst outcome of  $y$  is greater than the  $k$ -th worst outcome of  $y'$ ,  $\forall k \in \{1, \dots, T\}$  (see section 2.4).

Secondly, if we compare distributions from the point of view of a risk-averse decision-maker, we consider a surplus more valuable at low levels of outcomes than at high levels. Thus, transferring a small amount from an outcome to any worse outcome results in a preferred achievement vector, since it represents a less risky distribution. To reflect this, we compare two achievement vectors by comparing their cumulative outcomes (that is, the sums of their worst  $k$  outcomes, for each  $k \in \{1, \dots, T\}$ ) (Ogryczak 2000, 2002). In terms of distributions, this means comparing the two distributions represented by those vectors with respect to second order stochastic dominance. Thus, the distribution represented by  $y$  dominates that of  $y'$  with respect to SSD if the sum of the worst  $k$  outcomes of  $y$  is greater than the sum of worst  $k$  outcomes of  $y'$ ,  $\forall k \in \{1, \dots, T\}$  (see section 2.4).

*Remarks:*

- The Value-at-Risk (VaR) at confidence level  $\frac{k}{T} \in (0,1)$  of a random variable  $R_X$  is

$$VaR_{\frac{k}{T}}(R_X) = -\inf\{z / P(R_X \leq z) > \frac{k}{T}\} \text{ (see Acerbi and Tasche 2002). It easily}$$

follows that for a random variable with  $T$  equally probable outcomes, the  $k$ -th worst outcome is minus VaR at confidence level  $k/T$ . Thus, in this case, a random variable dominates another one with respect to FSD if its VaR at confidence level  $k/T$  is smaller, for every  $k$  in  $1, \dots, T$ .

- For a random variable with T equally probable outcomes, the Conditional Value-at-Risk (CVaR) at confidence level  $k/T$  is minus the mean of its worst k outcomes (this follows from the CVaR definition of Rockafellar and Uryasev 2002). Thus, in this case, a random variable dominates another one with respect to SSD if its CVaR at confidence level  $k/T$  is smaller, for every k in  $1, \dots, T$ .

In order to find the FSD efficient solutions of the portfolio selection problem, the objective functions in (1) are replaced by the k-th worst portfolio returns, for  $k=1, \dots, T$  respectively (see Ogryczak 2002).

Of greater interest are the SSD efficient solutions, which are obtained by replacing the objective functions in (1) with the cumulative outcomes: sums of the worst k portfolio returns, for  $k=1, \dots, T$ .

*Proposition:* For every  $k \in \{1, \dots, T\}$ , the sum of the worst k outcomes of a given achievement vector  $y=(y_1, \dots, y_T)$  is the optimal value of the following LP problem:

$$\begin{aligned} & \text{Max}(kt_k - \sum_{i=1}^T d_{ki}) \\ & \text{Subject to: } t_k - y_i \leq d_{ki}, \text{ for } i=1, \dots, T \\ & \quad d_{ki} \geq 0, \text{ for } i=1, \dots, T. \end{aligned}$$

The proof follows easily from the CVaR optimisation formula (Rockafellar and Uryasev 2000). An independent proof is given by Ogryczak (2002).

For each  $k \in \{1, \dots, T\}$ , the above problem has one free variable  $t_k$  and T non-negative variables  $d_{ki}$ ,  $i=1, \dots, T$ . The value of  $t_k$  is the k-th worst outcome of the achievement vector y (see Rockafellar and Uryasev 2000 or Ogryczak 2002). For each  $i \in \{1, \dots, T\}$ ,  $d_{ki}$  takes value 0 if  $y_i$  is greater than or equal to the k-th worst outcome  $t_k$ ; otherwise,  $d_{ki}=t_k-y_i$ . This means that at most  $(k-1)$  values of  $d_{ki}$  are strictly positive.

Thus, we can express the SSD efficient solutions for the portfolio selection problem as Pareto efficient solutions of the following multi-objective linear problem:

$$\text{Max}(t_1 - \sum_{i=1}^T d_{1i}, 2t_2 - \sum_{i=1}^T d_{2i}, \dots, Tt_T - \sum_{i=1}^T d_{Ti}) \quad (2)$$

*Subject to:*

$$t_k - f_i(x) \leq d_{ki}, \text{ for } i, k=1, \dots, T$$

$$d_{ki} \geq 0, \text{ for } i, k=1, \dots, T$$

$$\sum_{j=1}^n x_j = 1$$

$$x_j \geq 0, \text{ for } j=1, \dots, n$$

where  $f_i(x) = \sum_{j=1}^n r_{ij}x_j$  is the portfolio return in scenario i,  $i=1, \dots, T$ .

The problem has T free variables  $t_1, t_2, \dots, t_T$ , with  $t_k$  being the k-th worst outcome of the distribution that results,  $T^2$  non-negative variables  $(d_{ki})_{k,i=1, \dots, T}$  with the

interpretations given above and  $n$  non-negative variables giving the optimal portfolio weights.

*Remark:* An equivalent formulation is obtained by dividing the  $k$ -th objective function in (2) by  $k$ , for each  $k \in \{1, \dots, T\}$ , thus obtaining the means of the worst  $k$  outcomes as the objective functions.

### 3.2. The reference point method

The problem (2) has an infinite number of efficient solutions and additional criteria are needed in order to select a specific solution.

By specifying desirable outcomes for the portfolio return under the various scenarios, we are specifying a desirable (target) distribution. This results in desirable outcomes for the cumulative returns, and hence desirable outcomes for the objective functions in (2). We want to find the asset weights  $x_1, \dots, x_n$  that produce a SSD efficient distribution which comes closest to the target distribution.

The method used is the Reference Point Method (Wierzbicki 1982, Makowski and Wierzbicki 2003). In order to select a Pareto efficient solution of a multi-objective problem, aspiration levels (called reference points) are specified for some or all of the objective functions in the multi-objective problem. Depending on these reference points, a special scalarizing achievement function is constructed, which, when maximised, produces an efficient solution of the original multi-objective problem. Thus, the multi-objective problem is transformed into a single-objective problem. The process is interactive in that a solution is presented to the decision maker who can modify the aspiration levels in the light of the solutions already presented and obtain a further candidate solution.

Let  $z^* = (z^*_1, \dots, z^*_T)$  be the target (i.e.,  $z^*_k$  is the aspiration level for the  $k$ -th objective function). The simplest form of achievement function is:

$$\gamma_{z^*}(z) = \min_{1 \leq k \leq T} (z_k - z^*_k) + \varepsilon \sum_{k=1}^T (z_k - z^*_k) \quad (3)$$

where  $z = (z_1, \dots, z_T)$  is the current achievement vector and  $\varepsilon$  is an arbitrary small parameter.

The terms  $(z_k - z^*_k)$  in (3) are usually replaced with more complicated functions of  $z_k$  and  $z^*_k$ ,  $\gamma_k(z_k, z^*_k)$ , which must satisfy certain properties: being monotonically increasing with respect to  $z_k$  and taking value 0 if  $z_k = z^*_k$  (see for example Makowski and Wierzbicki 2003). These functions are called partial achievement functions since they measure the actual achievement of outcome  $z_k$  with respect to its corresponding aspiration level  $z^*_k$ . Various functions  $\gamma_k(z_k, z^*_k)$  provide a wide modelling environment for measuring individual achievements. Another example of such functions  $\gamma_k(z_k, z^*_k)$  are piece-wise linear functions with two segments whose gradients depend on whether the current outcome is above or below its aspiration point (see Makowski and Wierzbicki 2003).

Remarks:

- The scalarizing achievement function is essentially defined by the worst partial achievement:  $\min_{1 \leq k \leq T} \gamma_k(z_k, z_k^*)$ , which in our case means  $\min_{1 \leq k \leq T} (z_k - z_k^*)$ . The latter term is introduced in order to guarantee the solution efficiency in the case where the maximisation of the worst partial achievement has non-unique optimal solutions.
- Different values of  $\varepsilon$  can lead to different efficient solutions. A solution that maximises  $\gamma_{Z^*}$  will be a Pareto efficient solution of the multi-objective problem with a bound of  $M=1+1/\varepsilon$  on trade-off coefficients (see Makowski and Wierzbicki 2003 for more details).

In addition, one could introduce reservation levels  $z_k^r$  for the k-th objective function. In this case, the partial achievement functions will be  $\gamma_k(z_k, z_k^*, z_k^r)$ . The difference between aspiration and reservation levels is, as the name implies, that whilst it is desirable to achieve aspiration levels, reservation levels should be achieved if at all possible. The scalarizing achievement function must ensure that reservation levels preempt aspiration levels. Thus, reservation levels represent soft lower bounds of the objective functions.

In this paper, we consider piecewise linear partial achievement functions that use both aspiration and reservation levels (see Makowski and Wierzbicki 2003, Kaleta et al. 2003):

$$\gamma_k(z_k, z_k^*, z_k^r) = \begin{cases} \frac{\alpha(z_k - z_k^r)}{z_k^* - z_k^r}, & \text{for } z_k \leq z_k^r, \\ \frac{z_k - z_k^r}{z_k^* - z_k^r}, & \text{for } z_k^r < z_k < z_k^*, \\ \frac{\beta(z_k - z_k^*)}{z_k^* - z_k^r} + 1, & \text{for } z_k \geq z_k^*, \end{cases} \quad (4)$$

where  $\alpha, \beta$  are arbitrarily defined parameters such that  $0 < \beta < 1 < \alpha$ .

The parameter  $\alpha > 1$  represents the decision maker's dissatisfaction for outcomes worse than their corresponding reservation level, while the parameter  $\beta < 1$  represents satisfaction for outcomes better than their aspiration level (Kaleta et al. 2003).

The partial achievement functions (4) have attractive properties, which allow a simple interpretation of results. They take:

- negative values if the outcome is below its reservation level ( $z_k < z_k^r$ );
- value 0 for  $z_k = z_k^r$ ;
- value between 0 and 1 for outcomes between reservation and aspiration level
- value 1 for  $z_k = z_k^*$ ;
- value greater than 1 for outcomes better than the aspiration level.

Thus, they map outcome values onto a normalized scale of decision maker's satisfaction (Kaleta et al. 2003).

Moreover, under the condition  $0 < \beta < 1 < \alpha$ , the partial achievement function is strictly increasing and concave. Thus, it can be expressed in the form:

$$\gamma_k(z_k, z_k^*, z_k^r) = \min \left\{ \frac{\alpha(z_k - z_k^r)}{z_k^* - z_k^r}, \frac{z_k - z_k^r}{z_k^* - z_k^r}, \frac{\beta(z_k - z_k^*)}{z_k^* - z_k^r} + 1 \right\}, \quad (5)$$

which is useful when writing the model, since we can express the partial achievement functions using linear constraints:

$$\begin{aligned} \gamma_k(z_k, z_k^*, z_k^r) &\leq \frac{\alpha(z_k - z_k^r)}{z_k^* - z_k^r}; \\ \gamma_k(z_k, z_k^*, z_k^r) &\leq \frac{z_k - z_k^r}{z_k^* - z_k^r}; \\ \gamma_k(z_k, z_k^*, z_k^r) &\leq \frac{\beta(z_k - z_k^*)}{z_k^* - z_k^r} + 1. \end{aligned}$$

Since the achievement function is maximised, the largest possible values of  $\gamma_k(z_k, z_k^*, z_k^r)$  are required. Thus, the linear constraints above are equivalent to (5).

*Remark:* When the aspiration and the reservation levels for an objective function are chosen, they must lie between the lower and the upper bound of that objective function. So, if  $z_k^{\min}$  and  $z_k^{\max}$  are the worst and best values for the k-th objective function, respectively, we require:

$$z_k^{\min} \leq z_k^r \leq z_k^* \leq z_k^{\max} \quad (6)$$

Provided condition (6) is satisfied for every  $k \in \{1, \dots, T\}$ , the maximisation of the scalarizing achievement function produces an efficient solution of the initial multi-objective problem (2) (see Makowski and Wierzbicki 2003). Moreover, the converse is true, in the sense that for every efficient solution, there exist aspiration (and reservation) levels such that this efficient solution maximises the corresponding achievement function (see Makowski and Wierzbicki 2003).

### 3.3. Models for portfolio selection

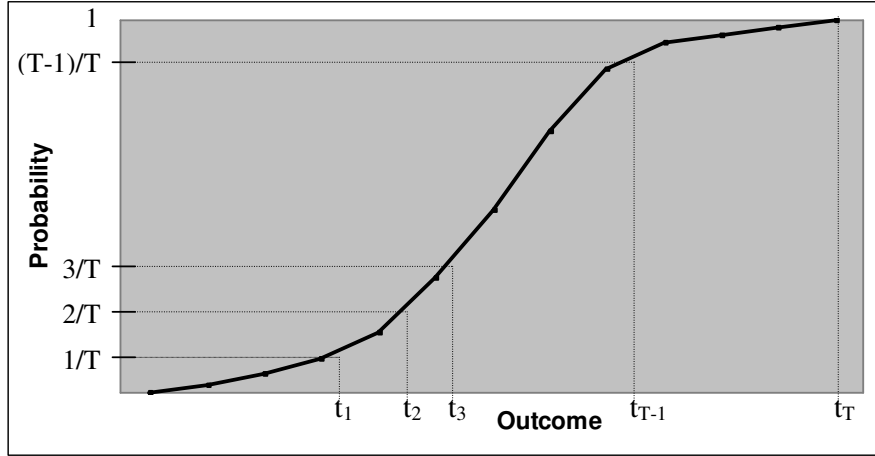
In this section, two models for portfolio selection (models M1 and M2) are presented. The motivation for model M1 is to find a SSD efficient portfolio whose return distribution comes close to a “target” aspiration distribution. Model M2 considers also a “reservation” distribution, which represents a first level target.

For model M1, let  $asp_k$  be the aspiration level for the k-th cumulative outcome (sum of the worst k outcomes),  $k=1, \dots, T$ . These levels imply aspiration levels for the ordered outcomes themselves and hence define a distribution, which we refer to as the “target”, “reference” or “aspiration” distribution.

Intuitively, the stochastically larger random variable has a smaller distribution function (see Figure 1). Thus, our aim is to obtain a distribution whose outcomes are positioned further to the right.

*Remark:* One way to choose the aspiration levels  $asp_k$  is to choose desirable values for the ordered outcomes of the distribution and then cumulate them. The fact that there is no “universal recipe” for choosing the aspiration levels should be regarded as an advantage of this method. We can analyse the distribution of the portfolio

that we obtain, and, if not satisfactory, we can modify the aspiration levels and obtain a further candidate solution portfolio.



**Figure 1: cumulative distribution function and ordered outcomes**

We will keep the same notation used in (2). In addition, let  $z_k$  represent the sum of the worst  $k$  outcomes, for  $k=1, \dots, T$ , and let  $\delta = \min_{k=1 \dots T} (z_k - asp_k)$  represent the worst partial achievement.  $z_k$  and  $\delta$  are free variables.

The following model for portfolio optimisation uses the achievement function defined by (3):

$$(M1): \text{Maximise } \delta + \varepsilon \left( \sum_{k=1}^T z_k - \sum_{k=1}^T asp_k \right).$$

$$\text{Subject to: } z_k = kt_k - \sum_{i=1}^T d_{ki}, \text{ for } k=1, \dots, T$$

$$z_k - asp_k \geq \delta, \text{ for } k=1, \dots, T.$$

$$t_k - \sum_{j=1}^n x_j r_{ij} \leq d_{ki}, \text{ for } k=1, \dots, T, i=1, \dots, T$$

$$\sum_{j=1}^n x_j = 1.$$

The nature of the decision variables is discussed below:

- $t_k$  = the  $k$ -th worst outcome of the portfolio return, for  $k=1, \dots, T$  (free variable);
- $z_k$  = the sum of the worst  $k$  outcomes of the portfolio return, for  $k=1, \dots, T$  (free variable);
- $\delta = \min_{k=1 \dots T} (z_k - asp_k)$  = the worst partial achievement (free variable);
- $(d_{ki})_{k,i=1, \dots, T}$  are non-negative variables.  $d_{ki}=0$  if the portfolio return in scenario  $i$  is greater than or equal to the  $k$ -th worst outcome  $t_k$ ; otherwise,  $d_{ki}$  is the difference between the  $k$ -th worst outcome  $t_k$  and the portfolio return in scenario  $i$ ;
- $x_j, j=1, \dots, n$  are the portfolio weights to be determined (non-negative variables)

In solving the model, there are three cases to consider, depending on the reference distribution chosen by the decision maker:

Case 1: If the reference distribution is not optimal with respect to SSD, the optimal value of the objective function will be strictly positive. In this case, the model improves on the reference distribution, in the sense that it provides a solution whose corresponding distribution is SSD efficient.

Case 2: If the reference distribution is optimal with respect to SSD, the optimal value of the objective function will be zero. In this case, the solution obtained matches exactly the reference distribution.

Case 3: If the reference distribution is not attainable, in the sense that (some of) the outcomes are too high so that there are no feasible portfolio weights that could yield such a distribution, the optimal value of the objective function will be strictly negative. In this case, the model provides a SSD efficient solution whose distribution of outcomes comes as close as possible (uniformly) to the reference levels.

The proof follows from the way that the reference point method works, depending on whether the aspiration points form an efficient, an inefficient or an unattainable achievement vector of the multi-objective problem (see Makowski and Wierzbicki 2003).

For the model M2, we consider both aspiration ( $asp_k$ ) and reservation ( $res_k$ ) levels for (some of) the cumulative outcomes  $z_k$ . The partial achievement functions are defined by (5).

In addition to the variables used in model (M1), there are another T free variables:  $\delta_k$ , for  $k=1, \dots, T$ , representing the values of the partial achievement functions.

$$(M2): \text{Maximise } \delta + \varepsilon \sum_{k=1}^T \delta_k$$

$$\text{Subject to: } \delta_k \geq \delta, \text{ for } k=1, \dots, T$$

$$\delta_k \leq \frac{\alpha(z_k - res_k)}{asp_k - res_k}, \text{ for } k=1, \dots, T$$

$$\delta_k \leq \frac{z_k - res_k}{asp_k - res_k}, \text{ for } k=1, \dots, T$$

$$\delta_k \leq \frac{\beta(z_k - asp_k)}{asp_k - res_k} + 1, \text{ for } k=1, \dots, T$$

$$z_k = kt_k - \sum_{i=1}^T d_{ki}, \text{ for } k=1, \dots, T$$

$$t_k - \sum_{j=1}^n x_j r_{ij} \leq d_{ki}, \text{ for } k=1, \dots, T, i=1, \dots, T$$

$$\sum_{j=1}^n x_j = 1$$

$$x_j \geq 0, \text{ for } j=1, \dots, n$$

$$d_{ki} \geq 0, \text{ for } k=1, \dots, T, i=1, \dots, T.$$

In solving model (M2), there are five cases to consider, depending on the aspiration and reservation distributions chosen by the decision maker:

Case 1: If the reservation distribution (and hence the aspiration distribution) is not attainable, the optimal value of the objective function is strictly negative. In this case, the model provides a SSD efficient solution whose return distribution is uniformly as close as possible to the reservation distribution.

Case 2: If the reservation distribution is SSD efficient (and the aspiration distribution is not attainable), the optimal value of the objective function is zero. In this case, the model provides a SSD efficient solution that matches the reservation distribution.

Case 3: If the reservation distribution is attainable but not SSD efficient and the aspiration distribution is not attainable, the optimal value of the objective function is between zero and one. In this case, the model provides a SSD efficient solution, whose return distribution improves on the reservation distribution and is uniformly as close as possible to the aspiration levels.

Case 4: If the aspiration distribution is SSD efficient (which implies that the reservation distribution is not), the optimal value of the objective function is equal to one. The solution obtained matches the aspiration distribution.

Case 5: If the aspiration distribution (and hence the reservation distribution) is not SSD efficient, the optimal value of the objective function is greater than one. In this case, the model provides a solution, whose return distribution improves on the aspiration distribution, in the sense that is SSD efficient.

### 3.4. Connection with mean-risk models

#### a. Mean-CVaR models

The  $k$ -th objective function  $z_k$  in (2) represents the sum (or, equivalently, the mean) of the worst  $k$  outcomes, for  $k=1, \dots, T$ . By attaching aspiration points only for the  $k$ -th (with  $k < T$ ) and the  $T$ -th objective function, we obtain a mean-CVaR model, with CVaR at confidence level  $k/T$ .

The classical mean-CVaR model, with CVaR at confidence level  $k/T$ , minimises CVaR using the optimisation formula of Rockafellar and Uryasev (2000), subject to conditions on expected return. Since CVaR at confidence level  $k/T$  is minus the mean of the worst  $k$  outcomes, the classical mean-CVaR model is equivalent to:

$$\begin{aligned} \text{Maximise } z_k &= kt_k - \sum_{i=1}^T d_{ki} \\ \text{Subject to: } z_T &= \sum_{i=1}^T \sum_{j=1}^n r_{ij} x_j \geq d \\ t_k - \sum_{j=1}^n x_j r_{ij} &\leq d_{ki}, \text{ for } i=1, \dots, T \end{aligned}$$



$$\sum_{j=1}^n x_j = 1$$

$$x_j \geq 0, \text{ for } j=1, \dots, n$$

$$d_{ki} \geq 0, \text{ for } i=1, \dots, T,$$

where  $d$  is the desired level for the sum of all outcomes (or, equivalently, where  $d/T$  is the desired level of expected return)

Basically, we have a multi-objective problem with 2 objective functions to maximise:  $z_k$  and  $z_T$ , and two different methods for producing an efficient solution.

The reference point approach introduces aspiration levels for both  $z_k$  and  $z_T$  and then maximises the achievement function:

$$\text{Max}[\min\{z_k - \text{asp}_k, z_T - \text{asp}_T\} + \varepsilon(z_k - \text{asp}_k + z_T - \text{asp}_T)],$$

whereas, in classical mean-CVaR optimisation,  $z_k$  is maximised whilst a lower limit is imposed on  $z_T$ :  $z_T \geq d$ .

Consider now the case where  $\text{asp}_T = d$ . Whether or not both approaches will produce the same solution depends on the aspiration level  $\text{asp}_k$ . Both solutions will be efficient, and, if  $\text{asp}_k$  is sufficiently low, they will be the same, with  $z_T = d$  and the best possible value for  $z_k$ .

If  $\text{asp}_k$  is too high, a different solution will be obtained with the reference point method, in which the value of  $z_k$  will be improved at the expense of decreasing the level for  $z_T$ . This solution will be efficient in a mean-CVaR model with a lower level imposed for the expected return; it has a better CVaR and a worse mean. Computationally, both models are of similar complexity.

The advantage of the reference point method is that it can deal with more than two reference levels. Although computationally more demanding, the distributions of portfolio returns can be modelled in a more refined way. In addition, this approach can produce any SSD efficient solution in the sense that there exist aspiration (and reservation) levels such that the solution maximises the corresponding achievement function.

It may be noted that if the  $k$ -th objective function in (2) is maximised, a particular efficient solution of (2) is obtained: a portfolio that minimises CVaR at confidence level  $k/T$ . In a mean-CVaR model, with CVaR at confidence level  $k/T$ , this corresponds to the lower extremity of the efficient frontier – the minimum risk portfolio.

### *b. The maximum mean portfolio*

The case of maximising the last objective function in (2) corresponds to maximising the expected return in a mean-risk model, irrespective of the risk measure used. In this case, the solution corresponds to the upper extremity of the efficient frontier in any mean-risk model.

### *c. Weighted CVaR*

It is well known that, for a multi-objective linear problem, we can obtain an efficient solution by the “weighted-sum approach” (see Steuer 1986, Ogryczak 2000).

Consider the multi-objective linear problem:

Max  $\{Cx \text{ such that } x \in Q\}$ , where  $C$  is the  $T \times n$  criterion matrix,  $Q$  is the set of feasible solutions.

Let  $\Lambda$  denote the set of weighting vectors:

$$\Lambda = \{\lambda \in R^T / \lambda_i > 0, \sum_{i=1}^T \lambda_i = 1\}.$$

An optimal solution of the single objective linear problem: Max  $\{\lambda^{*T}Cx \text{ such that } x \in Q\}$  with  $\lambda^* \in \Lambda$  is an efficient solution of the initial multi-objective problem.

The converse is also true: if  $x^* \in Q$  is an efficient solution of the initial multi-objective problem, then there exists  $\lambda^* \in \Lambda$  such that  $x^*$  is an optimal solution of the single objective problem Max  $\{\lambda^{*T}Cx \text{ such that } x \in Q\}$  (see Steuer 1986).

The objective functions in (2) can be considered as minus CVaR at confidence levels  $1/T, 2/T, \dots, k/T, \dots, T/T$ . Thus, every efficient solution of (2) can be considered as one that minimises a “weighted CVaR” measure:

$$\lambda_1 CVaR_{1/T} + \lambda_2 CVaR_{2/T} + \dots + \lambda_T CVaR_{T/T}.$$

*Remark:* A weighting approach for solving the portfolio optimisation problem via multi-objective problem (2) was suggested by Ogryczak (2000), Ogryczak and Sliwinski (2003). The advantage of using the reference point method instead of this is that the reference points provide a better way of controlling the selection of a solution, which is meaningful for the decision-maker, whereas the values of the weights to be attached are more difficult to interpret.

#### d. VaR constraints

We can impose (hard) constraints on VaR at a specified confidence level

$\alpha \in (0, 1)$ . Suppose  $\alpha$  lies between  $\frac{k-1}{T}$  and  $\frac{k}{T}$ . In this case, the value of the variable  $t_k$  in model (M1) will be minus VaR at confidence level  $\alpha$ . Thus, an additional constraint such as  $t_k \geq d$  may be added, where  $d$  is a desired lower limit for the value of the  $k$ -th worst outcome (or VaR at confidence level  $\alpha$ ). Of course, if the value of  $d$  is too high, infeasibility will result.

If a high reference point is used for the last objective function in the model (which represents the expected return of the portfolio), an approximation to the corresponding solution of a mean-VaR model will be obtained. This solution probably will not have the best expected return for the specified VaR level, but will be SSD efficient and with other outcomes close to the reference points.

#### e. Modelling the tail of distribution

It is possible to set reference levels only for the tail of the portfolio return distribution (i.e. the worst outcome, the sum of the worst two outcomes,...). The specified reference values are more likely to be achieved, and, in addition, the process is computationally simpler.

Another approach would be to use hard constraints for the tail of the distribution, while using reference values for (the other) general outcomes. For example, in model (M1) we can impose additional VaR constraints:  $t_1 \geq d_1, t_2 \geq d_2, \dots, t_k \geq d_k$  for a specific value of k. Of course, infeasibility may result in this case.

#### 4. Computational results

In this section, we present an experimental analysis of the models (M1) and (M2) presented in section 3.3. The returns of the portfolios produced by the models are evaluated in terms of their statistics (mean, median, standard deviation, skewness, kurtosis, range, minimum and maximum). Both in-sample and out-of-sample testing are performed. A data set, drawn from the Hang Seng index, was used for testing the models. Logarithmic returns of the 20 stocks that belonged to the index throughout the period 4 January 1993 - 1 December 2003 were considered. The data set consists of monthly returns and has 131 time periods, considered as equally probable scenarios ( $n=20, T=131$ ). For the out-of sample analysis, the behaviour of the portfolios obtained is examined in the eight months following the date of selection (January 2004-August 2004). The models were processed using CPLEX 9.0 optimisation solver and the value of  $\epsilon$  in the models was fixed at 0.00005.

In solving model (M1), there are three cases, depending on the aspiration distribution (see section 3.3). We examine each of these situations, by choosing appropriate aspiration distributions. The return distribution of the portfolio obtained in each case is analysed. In cases where the distribution obtained was considered unsatisfactory, we introduced reservation levels for some of the outcomes (mainly, in the left tail) and ran model (M2). We analysed the differences in the return distributions obtained with each model.

##### Case 1:

Here we examine the situation where the reference distribution is not SSD efficient. We chose the return distribution of the Hang Seng index as the reference distribution: the 131 monthly returns of the index were ordered and after that cumulated, in order to obtain the sum of the worst k outcomes, for  $k=1 \dots 131$ . The cumulative outcomes of the index were introduced as aspiration levels in model (M1).

On solving the model, we obtain a positive optimal value of the objective function (0.185), which indicates that the index is not SSD efficient, in contrast with the portfolio obtained by the model. We denote this portfolio as P1. The value of  $\delta$  (the worst partial achievement) is 0.18, which indicated that the cumulated outcomes of portfolio P1 are greater than the cumulated outcomes of the index by at least 0.18.

Table 1 shows the in-sample performance of the two portfolios considered, with P1 outperforming the index on seven of the eight measures used: higher mean, higher

median, lower variance, higher skewness (P1 has positive skewness, while the index has negative skewness), lower kurtosis, lower range, higher minimum but with lower maximum.

In Figure 2 we present the cumulative distribution functions of returns of portfolio P1 and the Hang Seng index. In principle, these are “step functions”, but in Figure 2 they are displayed as piece-wise linear functions by connecting the chosen bins of the outcomes and their corresponding probabilities. It is obvious that P1 has a cumulative distribution function that is “smaller” for low outcomes, so at low levels, its outcomes are distributed further to the right than those of the index. This was to be expected, since P1 dominates the index with respect to SSD.

	portfolio	
	P1	index
Mean	0.0073	0.0059
Median	0.0117	0.0094
St. Deviation	0.0635	0.0861
Skewness	0.0240	-0.0864
Kurtosis Excess	0.9901	2.1539
Range	0.3890	0.6128
Minimum	-0.1688	-0.3482
Maximum	0.2202	0.2645

Table 1: statistics for the return distributions of P1 and Hang Seng index (case 1)

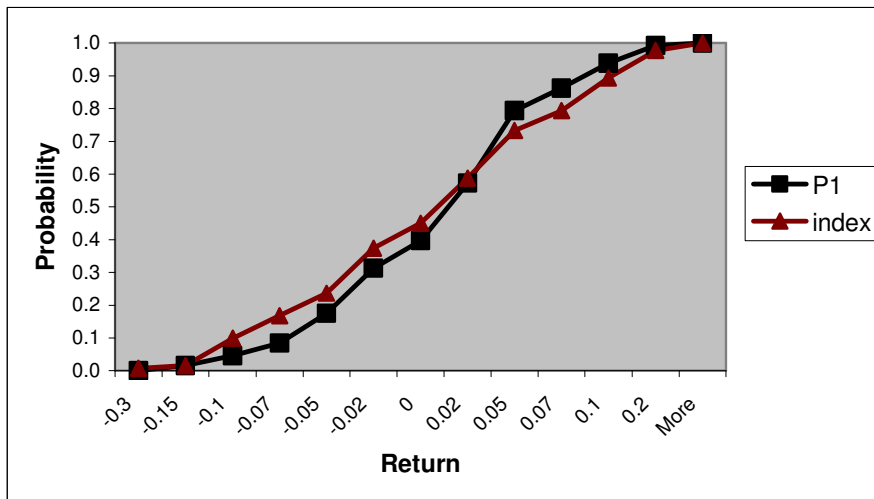
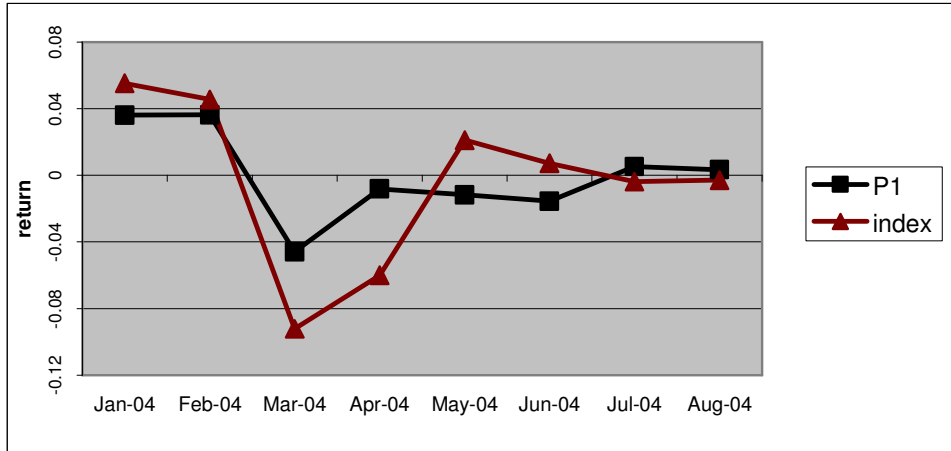


Figure 2: cumulative distribution functions for returns of P1 and Hang Seng index

For an out-of-sample evaluation, we considered the next eight time periods following the date of selection (January 2004-August 2004). The results are illustrated in Figure 3 and the ex-post parameters are presented in Table 2. The differences in the ex-post parameters for these two portfolios are very similar to those of the in-sample analysis: P1 has almost all the ex-post parameters better than those of the index.



**Figure 3: Out-of-sample evaluation of The Hang-Seng index and portfolio P1 (Jan-Aug 2004)**

	portfolio	
	P1	index
Mean	-0.00002	-0.00378
St. Deviation	0.02727	0.05018
Skewness	-0.06427	-0.79758
Kurtosis Excess	0.02027	-0.05637
Minimum	-0.04586	-0.09224
Maximum	0.03631	0.05518
Range	0.08217	0.14742

*Table 2: parameters for the out-of-sample performance of portfolios P1 and Hang Seng index (Jan-Aug 2004).*

Case 2:

Here we examine the case where the reference distribution is SSD efficient. The reference distribution is chosen to be the return distribution of the asset with the highest expected return. This portfolio, composed of just one asset, is efficient in any mean-risk model, irrespective of the risk measure – it lies at the upper extreme of the efficient frontier. Most importantly, it is also SSD efficient - it maximises the last objective function in (2), representing the sum of all outcomes. We call this portfolio Mean\_max. Running model (M1) having as aspiration points the cumulated outcomes of this distribution, we obtain, as expected, an optimal objective value of 0 and the same portfolio weights: a value of one corresponding to the asset with the highest expected return and zero for the remaining weights. This distribution is SSD efficient and therefore cannot be improved. In order to illustrate this case we performed an “inverse” procedure by specifying the outcomes for the distribution and then determining the portfolio weights that produce these outcomes.

This distribution is not particularly desirable but was chosen to demonstrate model behaviour. Though SSD efficient and also efficient in all mean-risk models, it has undesirable properties, such as a long left tail (worst outcome= -0.72, second worst outcome= -0.54, etc.). In such a situation model (M2) can play a useful role. As an example, consider reservation levels for the worst five outcomes to be: -0.3, -0.3, -0.2,

-0.2, -0.2. Cumulating these values, reservation levels for the first five cumulated outcomes are obtained. Aspiration levels were defined by considering the optimal values for the first five cumulated outcomes and the cumulated outcomes of Mean\_max for the remaining ones. On solving this instance of model (M2), portfolio P2 is obtained and an optimal value of 0.91, which indicates, as expected, that the reservation distribution is not SSD efficient and the aspiration distribution is not attainable.

As shown in Table 3, although the return distribution of Mean\_max has a higher expected value than that of P2, it has also higher volatility and higher risk of lower returns (negative skewness, high kurtosis, very low minimum).

	portfolio	
	Mean_max	P2
Mean	0.01273	0.00491
Median	0.01604	0.00568
St. Deviation	0.14332	0.06084
Skewness	-1.40261	0.12335
Kurtosis Excess	5.74807	0.81840
Range	0.99809	0.34853
Minimum	-0.72631	-0.14352
Maximum	0.27179	0.20501

Table 3: statistics for the return distributions of Mean\_max and P2 (case 2)

For an out-of-sample evaluation, the performance of the two portfolios over the eight time periods (January 2004-August 2004) is illustrated in Figure 4 and the ex-post parameters are presented in Table 4. Except for the maximum of out-of-sample returns, which is a little higher for Mean\_max, all the other ex-post parameters are better for portfolio P2. Even the mean of the out-of-sample returns is higher for P2, which shows the importance of modelling the tail of a distribution.

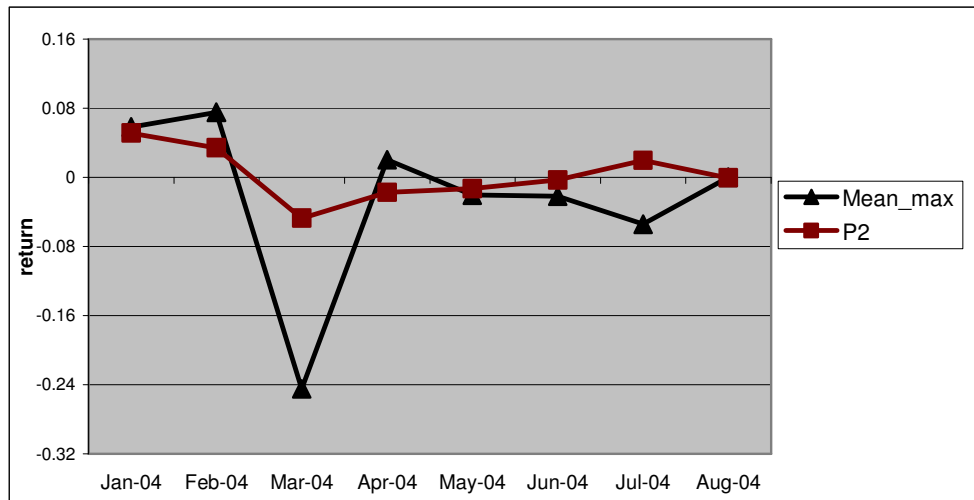


Figure 4: Out-of-sample evaluation of portfolios Mean\_max and P2.2 (Jan-Aug 2004)

	portfolio	
	Mean_max	P2
Mean	-0.02337	0.00303
St. Deviation	0.09914	0.03122
Skewness	-1.81243	0.05663
Kurtosis Excess	4.12961	-0.19501
Minimum	-0.24471	-0.04709
Maximum	0.07535	0.05131
Range	0.32006	0.09841

Table 4: parameters for the out-of-sample performance of portfolios Mean\_max and P2 (Jan-Aug 2004)

Case 3:

Here we consider the case when the reference distribution is unattainable. We chose as aspiration points the optimal values for the cumulated outcomes. As expected, in solving model (M1) with these aspiration points, a negative optimal value for the objective function is obtained (-0.55). The portfolio obtained is denoted by P3.1.

Model M2 can be used to obtain a portfolio whose return distribution has a shorter tail. By defining reservation points for the first seven cumulated outcomes and keeping the same aspiration levels as before (the best values for the cumulated outcomes) and solving model (M2), a portfolio P3.2 is obtained with an optimal value of the objective function between 0 and 1 (0.804), which indicates that the reservation distribution is not SSD efficient and the aspiration distribution is not attainable. (In this example, the reservation levels were chosen by first imposing minimum desirable outcomes for the worst seven outcomes of the portfolio distribution: -0.15, -0.13, -0.13, -0.12, -0.11, -0.11, -0.11. By cumulating these, reservation levels for the first seven cumulated outcomes were obtained).

As shown in Table 5, although the return distribution of P3.2 has a lower mean (and a lower maximum), most of its statistics are better, especially those concerning the distribution's tail (higher skewness, lower kurtosis, lower range, higher minimum).

	portfolio	
	P3.1	P3.2
Mean	0.00856	0.00324
Median	0.00828	0.00413
St. Dev.	0.06785	0.06142
Skewness	-0.15626	0.23144
Kurtosis Excess	1.02569	0.59529
Range	0.42907	0.32448
Minimum	-0.21090	-0.13603
Maximum	0.21816	0.18845

Table 5: statistics for the return distributions of P3.1 and P3.2 (case 3)

The out-of-sample returns of portfolios P3.1 and P3.2 over the eight time periods (January 2004-August 2004) are illustrated in Figure 5 and the ex-post parameters are presented in Table 6. The better performance of portfolio P3.2 can be clearly seen. All the ex-post parameters (except for variance, which is slightly higher) are better.

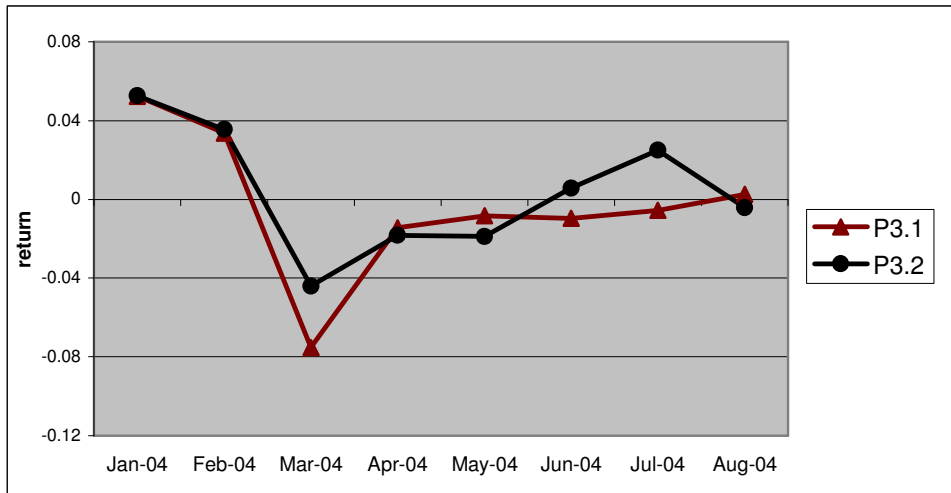


Figure 5: Out-of-sample evaluation of portfolios P3.1 and P3.2 (Jan-Aug 2004)

	portfolio	
	P3.1	P3.2
Mean	-0.00307	0.00421
St. Deviation	0.03753	0.03207
Skewness	-0.56962	0.09690
Kurtosis Excess	1.73200	-0.80341
Range	0.12756	0.09674
Minimum	-0.07525	-0.04407
Maximum	0.05231	0.05267

Table 6: parameters for the out-of-sample performance of portfolios P3.1 and P3.2 (Jan-Aug 2004)

## 5. Concluding remarks

The portfolio selection problem is presented as a particular case of decision making among random variables. The concern is to make meaningful decisions, according to observed economic behaviour, and at the same time preserve computational tractability. Meaningful decisions are usually considered those consistent with an axiomatic model of preference (stochastic dominance). Here, we considered the problem of choosing a portfolio that is efficient with respect to second order stochastic dominance (SSD). Thus, decision makers are assumed to be rational and risk-averse. The number of portfolios that are efficient with respect to second order stochastic dominance is infinite and additional criteria are needed in order to select a specific, meaningful solution. We proposed to do this by specifying a reference distribution, i.e. specifying return outcomes that would be desirable under various scenarios, and finding the portfolio whose return distribution comes closest to the reference.

The models use target levels for the sums of the worst  $k$  outcomes, for  $k \in \{1, \dots, T\}$ , where  $T$  is the number of (equally probable) scenarios considered. The first model uses only aspiration levels. The second one uses, in addition, reservation levels,



which pre-empt the aspiration levels in the optimisation. Both models produce SSD efficient portfolios irrespective of the reference distribution used.

The models were evaluated using a dataset drawn from the Hang Seng index and using different reference distributions. Both in sample and out-of sample analyses were performed. The portfolios produced by the models were evaluated by comparing statistics of their return distributions. Of primary interest in the evaluation is the status of the reference distribution with respect to SSD efficiency. In cases where the reference distribution was not SSD efficient, the portfolio produced by the first model (M1) had a return distribution with better statistics than the reference, and, in addition, better out-of-sample results. When the reference distribution was SSD efficient, model M1 produced a portfolio whose return distribution matched the reference. The robustness of the model is such that, when the reference distribution is unattainable, a portfolio is produced whose return distribution is as close as possible to the reference.

Model M1 was extended in order to gain greater modelling power in selecting desirable portfolios. In cases where the return distributions of the portfolios produced have undesirable outcomes under specific scenarios, the second model (M2) can be used to produce portfolios which improve on these outcomes. In particular, we used model M2 to model the left tail of the return distributions. The effectiveness of model M2 was demonstrated by producing portfolios that outperformed those produced by M1, both in terms of in sample statistics and out-of-sample results.

It is envisaged that, as in many multi-objective applications, the models are used interactively with the decision maker specifying new aspiration (and reservation) levels in order to obtain portfolios with progressively better return distributions.

## REFERENCES

- [1] ACERBI, C., D. TASCHE (2002): On the Coherence of Expected Shortfall, *Journal of Banking and Finance*, **26**, 1487-1503.
- [2] KALETA, M., OGRYCZAK, W., TOCZYLOWSKI, E., ZOLTOWSKA, I. (2003): On Multiple Criteria Decision Support for Suppliers on the Competitive Electric Power Market, *Annals of Operations Research*, **121**, 79-104.
- [3] MAKOWSKI, M. and A.P. WIERZBICKI (2003): Modelling Knowledge. Model Based Decision Support and Soft Computations, *Applied Decision Support with Soft Computing*, **124**, 3-60, Springer-Verlag, Berlin, NY; also available at: [www.iiasa.ac.at/~marek/pubs](http://www.iiasa.ac.at/~marek/pubs).
- [4] OGRYCZAK, W., A. RUSZCZYNSKI (1999): From Stochastic Dominance to Mean- Risk Models: Semideviations as Risk Measures, *European Journal of Operational Research*, **116**, 33-50.
- [5] OGRYCZAK, W. (2000): Multiple Criteria Linear Programming Model for Portfolio Selection, *Annals of Operations Research*, **97**, 143-162.

- [6] OGRYCZAK, W., A. RUSZCZYNSKI (2001): On Consistency of Stochastic Dominance and Mean- Semideviations Models, *Mathematical Programming*, **89**, 217-232.
- [7] OGRYCZAK, W. (2002): Multiple Criteria Optimization and Decisions under Risk, *Control and Cybernetics*, **31**, 975-1003.
- [8] OGRYCZAK, W., SLIWINSKI, T. (2003): On solving linear programs with the ordered weighted averaging objective, *European Journal of Operations Research*, **148**, 80-91.
- [9] ROCKAFELLER, R.T., S. URYASEV (2000): Optimization of Conditional Value-at-Risk, *Journal of Risk*, **2**, 21-42.
- [10] ROCKAFELLER, R.T., S. URYASEV (2002): Conditional Value-at-Risk for general loss distributions, *Journal of Banking and Finance*, **26**, 1443-1471.
- [11] STEUER, R. (1986): *Multiple Criteria Optimization - Theory, Computation and Application*, John Wiley & Sons, New York.
- [12] WHITMORE, G.A., M.C. FINDLAY (Eds.) (1978). *Stochastic Dominance: An Approach to Decision- Making Under Risk*, D.C.Heath, Lexington, MA.
- [13] WIERZBICKI, A.P. (1982): A Mathematical Basis for Satisficing Decision Making, *Mathematical Modeling*, **3**, 391-405.