Chapter 15
Risk-Averse Portfolio Optimization via Stochastic Dominance Constraints

Darinka Dentcheva and Andrzej Ruszczyński

Abstract We consider the problem of constructing a portfolio of finitely many assets whose return rates are described by a discrete joint distribution. We present a new approach to portfolio selection based on stochastic dominance. The portfolio return rate in the new model is required to stochastically dominate a random benchmark. We formulate optimality conditions and duality relations for these models and construct equivalent optimization models with utility functions. Two different formulations of the stochastic dominance constraint: primal and inverse, lead to two dual problems that involve von Neuman–Morgenstern utility functions for the primal formulation and rank dependent (or dual) utility functions for the inverse formulation. The utility functions play the roles of Lagrange multipliers associated with the dominance constraints. In this way our model provides a link between the expected utility theory and the rank dependent utility theory. We also compare our approach to models using value at risk and conditional value at risk constraints. A numerical example illustrates the new approach.

Keywords Portfolio optimization • Stochastic dominance • Stochastic order • Risk • Expected utility • Duality • Rank dependent utility • Yaari's dual utility • Value at risk • Conditional value at risk

15.1 Introduction

The problem of optimal portfolio selection is subject of major theoretical and computational studies in finance. A fundamental issue while dealing with uncertain outcomes is a theoretically sound approach to their comparison.

The theory of stochastic orders plays a fundamental role in economics (see Mosler and Scarsini 1991; Whitmore and Findlay 1978). These are relations that induce partial order in the space of real random variables in the following way. A random variable R dominates the random variable Y if $\mathsf{E}[u(R)] \geq \mathsf{E}[u(Y)]$ for all functions $u(\cdot)$ from certain set of functions, called the generator of the order. The concept of *stochastic dominance* is very popular and widely used in economics and finance because of its relation to models of risk-averse preferences (Fishburn 1964). It originated from the theory of majorization (Hardy et al. 1934) for the discrete case, and was later extended to general distributions (Quirk and Saposnik 1962; Hadar and Russell 1969; Rotschild and Stiglitz 1969). Stochastic dominance of second order is defined by the set of nondecreasing concave functions: a random variable R dominates another random variable Y in the second order if $\mathsf{E}[u(R)] \geq \mathsf{E}[u(Y)]$ for all nondecreasing concave functions $u(\cdot)$ for which these expected values are finite. Thus, no risk-averse decision maker will prefer a portfolio with return rate Y over a portfolio with return rate R.

A popular approach is the utility optimization approach. Von Neumann and Morgenstern (1944) developed the *expected utility theory*: for every rational decision maker there exists a utility function $u(\cdot)$ such that the decision maker prefers outcome R over outcome Y if and only if $\mathsf{E}[u(R)] > \mathsf{E}[u(Y)]$. This approach can be implemented also very efficiently; however, it is almost impossible to elicit the utility function of a decision maker explicitly. More difficulties arise when a group of decision makers with different utility functions have to reach a consensus. Recently, the *dual utility theory* (or *rank dependent expected utility theory*) has attracted much attention in economics. This approach was first presented by Quiggin (1982) and later rediscovered in a special case by Yaari (1987). From a different system of axioms than those of von Neumann and Morgenstern, one derives that every decision maker has a certain *rank dependent utility function $w : [0, 1] \to \mathsf{R}$*. Then a nonnegative outcome R is preferred over a nonnegative outcome Y, if and only if

$$- \int_0^1 w(p)\, dF_{(-1)}(R; p) \geq - \int_0^1 w(p)\, dF_{(-1)}(Y; p),$$

(15.1)

D. Dentcheva
Stevens Institute of Technology, Hoboken, NJ, USA

A. Ruszczyński (✉)
Rutgers University, New Brunswick, NJ, USA
e-mail: rusz@business.rutgers.edu

where, $F_{(-1)}(R; \cdot)$ is the inverse distribution function of R. For a comprehensive treatment of the rank dependent utility theory, we refer to Quiggin (1993), and for its application in actuarial mathematics, see Wang et al. 1997; Wang and Yong 1998.

Another classical approach, pioneered by Markowitz (1952, 1959, 1987), is the mean-risk approach, which compares the portfolios with respect to two characteristics. One is the expected return rate (*the mean*) and another one is *the risk*, which is given by some scalar measure of the uncertainty of the portfolio return rate. The mean-risk approach recommends the selection of Pareto-efficient portfolios with respect to these two criteria. In a mean-risk portfolio model we combine these criteria by specifying some parameter as a tradeoff between them. As a parametric optimization problem the mean-risk model can be solved numerically very efficiently, which makes this approach very attractive (Konno and Yamazaki 1991; Ruszczyński and Vanderbei 2003).

In this paper we formulate a model for risk-averse portfolio optimization and demonstrate its relation to the expected utility approach and to rank dependent utility approach. We optimize the portfolio performance under an additional constraint that the portfolio return rate stochastically dominates a benchmark return rate, for example, the return rate of an index. The model is based on the publications of Dentcheva and Ruszczyński (2003a, b, c; 2004a, b) where a new model of risk-averse optimization has been introduced. This approach has a fundamental advantage over mean-risk models and utility function models. All data for our model are readily available. In mean-risk models the choice of the risk measure has an arbitrary character, and it is difficult to argue for one measure against another. Similarly, optimization of expected utility requires the form of the utility function to be specified. Our analysis, departing from the benchmark outcome, generates implied utility function of the decision maker. It is implicitly defined by the benchmark used, and by the problem under consideration. We provide two problem formulations in which the stochastic dominance has a primal or inverse form: a Lorenz curve. The primal form has a dual problem in terms of expected utility functions, and the inverse form has a dual problem in terms of rank dependent utility functions. In this way our model provides a link between this two competing economic approaches. Duality relations with coherent measures of risk are explored in Dentcheva and Ruszczyński (2008).

15.2 The Portfolio Problem

Let $R_1, R_2, \ldots, R_n$ be random return rates of assets $1, 2, \ldots, n$. We assume that $\mathsf{E}[|R_j|] < \infty$ for all $j = 1, \ldots, n$.

Our aim is to invest our capital in these assets in order to obtain some desirable characteristics of the total return rate on the investment. Denoting by $x_1, x_2, \ldots, x_n$ the fractions of the initial capital invested in assets $1, 2, \ldots, n$ we can easily derive the formula for the total return rate:

$$R(x) = R_1 x_1 + R_2 x_2 + \ldots + R_n x_n. \tag{15.2}$$

Clearly, the set of possible asset allocations can be defined as follows:

$$X = \{x \in \mathsf{R}^n : x_1 + x_2 + \ldots + x_n = 1, \ x_j \geq 0, \ j = 1, 2, \ldots, n\}.$$

In some applications one may introduce the possibility of *short positions* (i.e., allow some x_j's to become negative). Other restrictions may limit the exposure to particular assets or their groups, by imposing upper bounds on the x_j's or on their partial sums. One can also limit the absolute differences between the x_j's and some reference investments $\bar{x}_j$, which may represent the existing portfolio, and so on. Our analysis does not depend on the detailed way this set is defined; we only use the fact that it is a convex polyhedron. All modifications discussed above define some convex polyhedral feasible sets and are, therefore, covered by our approach.

The main difficulty in formulating a meaningful portfolio optimization problem is the definition of the preference structure among feasible portfolios. If we use only the mean return rate $\mathsf{E}[R(x)]$, then the resulting optimization problem has a trivial and meaningless solution: invest everything in assets that have the maximum expected return rate. For these reasons, the practice of portfolio optimization usually resorts to two approaches.

In the first approach we associate with portfolio x some dispersion measure $\rho(R(x))$ representing the variability of the return rate $R(x)$. In the classical Markowitz model the function $\rho(R(x))$ is the variance of the return rate,

$$\rho(R(x)) = \mathsf{V}[R(x)],$$

but many other measures are possible here as well.

The mean-risk portfolio optimization problem is formulated as follows:

$$\max_{x \in X} \mathsf{E}[R(x)] - \lambda \rho(R(x)). \tag{15.3}$$

Here, λ is a nonnegative parameter representing our desirable exchange rate of mean for risk. If $\lambda = 0$, the risk has no value and the problem reduces to the problem of maximizing the mean. If $\lambda > 0$ we look for a compromise between the mean and the risk. Alternatively, one can minimize the risk function $\rho(x)$, while fixing the expected return rate $\mathsf{E}[R(x)]$ at some value m, and consider a family of problems

parametrized by m. The reader is referred to the book by Elton et al. (2006) for the modern perspective on mean-risk analysis in portfolio theory.

The general question of constructing mean-risk models that are in harmony with the stochastic dominance relations has been the subject of the analysis of the recent papers by Ogryczak and Ruszczyński (1999; 2001; 2002). We have identified there several primal risk measures, most notably central semideviations, and dual risk measures, based on the Lorenz curve, which are consistent with the stochastic dominance relations.

The second approach is to select a certain *utility function* $u : \mathsf{R} \to \mathsf{R}$ and to formulate the following optimization problem

$$\max_{x \in X} \mathsf{E}[u(R(x))]. \qquad (15.4)$$

It is usually required that the function $u(\cdot)$ is concave and nondecreasing, thus representing preferences of a risk-averse decision maker (Fishburn 1964; 1970).

Recently, a *dual (rank dependent) utility* model attracts much attention. It is based on distorting the cumulative probability distribution of the random variable $R(x)$ rather than applying a nonlinear function $u(\cdot)$ to the realizations of $R(x)$. The corresponding problem has the following form

$$\max_{x \in X} \int_0^1 F_{(-1)}(R(x), p) dw(p). \qquad (15.5)$$

Here $F_{(-1)}(R(x), p)$ is the p-quantile of the random variable $R(x)$, and $w(\cdot)$ is the *rank dependent utility function*, which distorts the probability distribution. We discuss this in Sect. 15.3.2.

The challenge in both utility approaches is to select the appropriate utility function or rank dependent utility function that represent our preferences and whose application leads to nontrivial and meaningful solutions of Equation (15.4) or (15.5).

In this paper we propose an alternative approach: introducing a comparison to a benchmark return rate into our optimization problem. The comparison is based on the stochastic dominance relation. More specifically, we consider only portfolios whose return rates stochastically dominates a certain benchmark return rate.

15.3 Stochastic Dominance

15.3.1 Direct Forms

In the stochastic dominance approach, random return rates are compared by a point-wise comparison of some performance functions constructed from their distribution functions. For a real random variable V, its first performance function is defined as the right-continuous cumulative distribution function of V:

$$F_1(V; \eta) = \mathsf{P}\{V \leq \eta\} \quad for \ \eta \in \mathsf{R}.$$

A random return V is said (Lehmann 1955; Quirk and Saposnik 1962) to *stochastically dominate* another random return S in the first order, denoted $V \succeq_{(1)} S$, if

$$F_1(V; \eta) \leq F_1(S; \eta) \quad for \ all \ \eta \in \mathsf{R}.$$

We can say that V is "stochastically larger" than S, because it takes values lower than η with smaller (or equal) probabilities than S, no matter what the target η is.

The second performance function F_2 is given by areas below the distribution function F,

$$F_2(V; \eta) = \int_{-\infty}^{\eta} F_1(V; \xi) \, d\xi \quad for \ \eta \in \mathsf{R},$$

and defines the weak relation of the *second order stochastic dominance* (SSD). That is, random return V stochastically dominates S in the second order, denoted $V \succeq_{(2)} S$, if

$$F_2(V; \eta) \leq F_2(S; \eta) \quad for \ all \ \eta \in \mathsf{R}.$$

(see Hadar and Russell 1969; Rotschild and Stiglitz 1969).

We can express the function $F_2(V; \cdot)$ as the expected shortfall (see, for example, Levy 2006; Ogryczak and Ruszczyński 1999): for each target value η we have

$$F_2(V; \eta) = \mathsf{E}[(\eta - V)_+], \qquad (15.6)$$

where $(\eta - V)_+ = \max(\eta - V, 0)$. The function $F_{(2)}(V; \cdot)$ is continuous, convex, nonnegative and nondecreasing. It is well defined for all random variables V with finite expected value. Due to this representation, the second order stochastic dominance relation $V \succeq_{(2)} S$ can be equivalently characterized by the system of inequalities on the expected shortfall below any target η:

$$\mathsf{E}[(\eta - V)_+] \leq \mathsf{E}[(\eta - S)_+] \quad for \ all \ \eta \in \mathsf{R}. \qquad (15.7)$$

Also, we obtain an equivalent characterization in terms of the expected utility theory of von Neumann and Morgenstern (see, for example, Hanoch and Levy 1969; Levy 2006; Müller and Stoyan 2002):

• For any two random variables V, S the relation $V \succeq_{(1)} S$ holds true if and only if for all nondecreasing functions $u(\cdot)$ defined on R we have

$$\mathsf{E}[u(V)] \geq \mathsf{E}[u(S)]. \qquad (15.8)$$

- For any two random variables V, S with finite expectations, the relation $V \succeq_{(2)} S$ holds true if and only if Equation (15.8) is satisfied for all nondecreasing concave functions $u(\cdot)$.

In the context of portfolio optimization, we consider stochastic dominance relations between random return rates defined by Equation (15.2). Thus, we say that portfolio x *dominates* portfolio y *in the first order*, if

$$F_1(R(x); \eta) \leq F_1(R(y); \eta) \quad for\ all\ \eta \in \mathsf{R}.$$

This is illustrated in Fig. 15.1.

Similarly, we say that x *dominates y in the second order* $(R(x) \succeq_{(2)} R(y))$, if

$$F_2(R(x); \eta) \leq F_2(R(y); \eta) \quad for\ all\ \eta \in \mathsf{R}.$$

The second order relation is illustrated in Fig. 15.2.

Recall that the individual return rates R_j have finite expected values and thus the function $F_2(R(x); \cdot)$ is well defined.
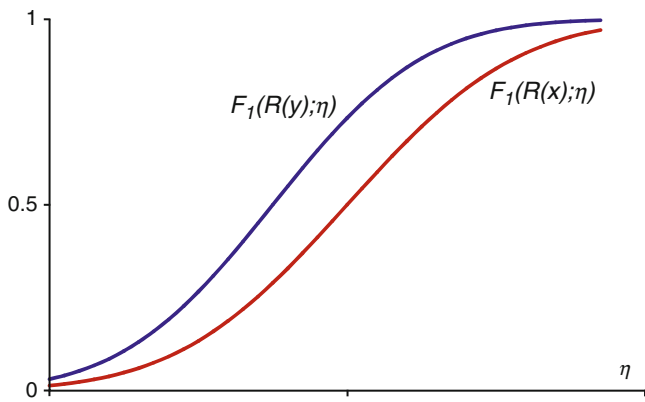

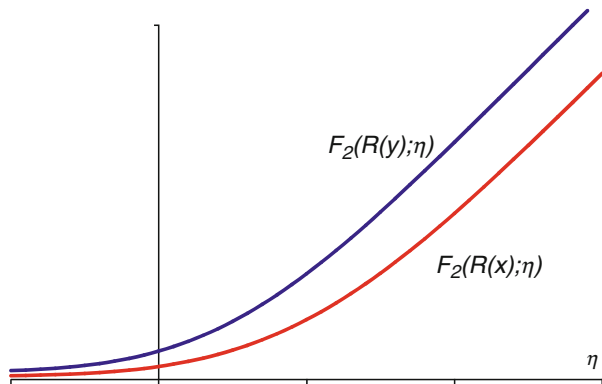**Fig. 15.1** First order stochastic dominance $R(x) \succeq_{(1)} R(y)$

Fig. 15.2 Second order dominance $R(x) \succeq_{(2)} R(y)$

15.3.2 Inverse Forms

Let us consider the inverse model of stochastic dominance, frequently referred to as *Lorenz dominance*. For a real random variable V (for example, a random return rate) we define the left-continuous inverse of the cumulative distribution function $F_1(V; \cdot)$ as follows:

$$F_{(-1)}(V; p) = \inf \{\eta : F_1(V; \eta) \geq p\} \quad for \quad 0 < p < 1.$$

Given $p \in (0, 1)$, the number $q = q(V; p)$ is called a *p-quantile* of the random variable V if

$$\mathsf{P}\{V < q\} \leq p \leq \mathsf{P}\{V \leq q\}.$$

For $p \in (0, 1)$ the set of p-quantiles is a closed interval and $F_{(-1)}(V; p)$ represents its left end. Directly from the definition of the first order dominance we see that

$$V \succeq_{(1)} S \Leftrightarrow F_{(-1)}(V; p) \geq F_{(-1)}(S; p)\ for\ all\ 0 < p < 1. \quad (15.9)$$

The first order dominance constraint can be interpreted as a continuum of probabilistic (chance) constraints, studied in stochastic optimization (see, Dentcheva 2005; Prékopa 2003).

Our analysis uses the *absolute Lorenz function* $F_{(-2)}(V; \cdot) : [0, 1] \rightarrow \mathsf{R}$, introduced in (Lorenz 1905). It is defined as the cumulative quantile:

$$F_{(-2)}(V; p) = \int_0^p F_{(-1)}(V; t) dt \quad for \quad 0 < p \leq 1,$$
$$(15.10)$$
$$F_{(-2)}(V; 0) = 0.$$

Similarly to $F_2(V; \cdot)$, the function $F_{(-2)}(V; \cdot)$ is well defined for any random variable V, which has a finite expected value. We notice that

$$F_{(-2)}(V; 1) = \int_0^1 F_{(-1)}(V; t) dt = \mathsf{E}[V].$$

By construction, the Lorenz function is convex. Lorenz functions are commonly used for inequality ordering of positive random variables, relative to their (positive) expectations (see Gastwirth 1971; Muliere and Scarsini 1989). Such a Lorenz function, $p \mapsto F_{(-2)}(V; p)/\mathsf{E}[V]$, is convex and nondecreasing. The absolute Lorenz function, however, is not monotone, when negative outcomes occur.

It is well known (see, for example, Ogryczak and Ruszczyński 2002) that we may fully characterize the second order dominance relation by using the function $F_{(-2)}(V; \cdot)$:

$$V \succeq_{(2)} S \Leftrightarrow F_{(-2)}(V; p) \geq F_{(-2)}(S; p) \ for \ all \ 0 \leq p \leq 1. \tag{15.11}$$

This characterization of stochastic dominance by Lorenz functions is widely used in economics and statistics.

We now provide an equivalent characterization by rank dependent utility functions. It is analogous to the characterization by expected utility functions.

Dentcheva and Ruszczyński (2006b) provide the following characterization.

- For any two random variables V, S the relation $V \succeq_{(1)} S$ holds true if and only if for all nondecreasing functions $w(\cdot)$ defined on [0,1] we have

$$\int_0^1 F_{(-1)}(V; p) dw(p) \geq \int_0^1 F_{(-1)}(S; p) dw(p). \tag{15.12}$$

- For any two random variables V, S with finite expectations, the relation $V \succeq_{(2)} S$ holds true if and only if Equation (15.12) is satisfied for all nondecreasing concave functions $w(\cdot)$.

The functions $w(\cdot)$ appearing in this characterization are rank dependent (dual) utility functions.

In the context of portfolio optimization, we consider stochastic dominance relations between random return rates defined by Equation (15.2). Thus, we say that portfolio x *dominates* portfolio y *in the first order*, if

$$F_{(-1)}(R(x); p) \geq F_{(-1)}(R(y); p) \quad for \ all \ p \in (0, 1).$$
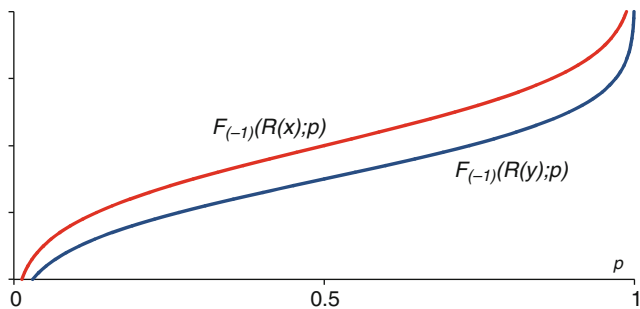
This is illustrated in Fig. 15.3.

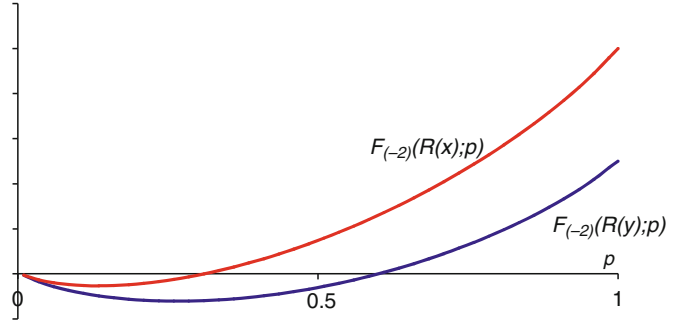**Fig. 15.3** First order stochastic dominance $R(x) \succeq_{(1)} R(y)$ in the inverse form

Fig. 15.4 Second order dominance $R(x) \succeq_{(2)} R(y)$ in the inverse form

Similarly, we say that *x dominates y in the second order* $(R(x) \succeq_{(2)} R(y))$, if

$$F_{(-2)}(R(x); p) \geq F_{(-2)}(R(y); p) \quad for \ all \ p \in [0, 1]. \tag{15.13}$$

Recall that the individual return rates R_j have finite expected values and thus the function $F_{(-2)}(R(x); \cdot)$ is well defined. The second order relation is illustrated in Fig. 15.4.

15.3.3 Relations to Value at Risk and Conditional Value at Risk

There are fundamental relations between the concepts of Value at Risk (VaR) and Conditional Value at Risk (CVaR) and the stochastic dominance constraints. The VaR constraint in the portfolio context is formulated as follows. We define the loss rate $L(x) = -R(x)$. We specify the maximum fraction ω_p of the initial capital allowed for risk exposure at risk level $p \in (0, 1)$, and we require that

$$P[L(x) \leq \omega_p] \geq 1 - p.$$

Denoting by $VaR_p(L(x))$ the left $(1 - p)$-quantile of the random variable $L(x)$, we can equivalently formulate the VaR constraint as

$$VaR_p(L(x)) \leq \omega_p.$$

The first order stochastic dominance relation between two portfolios is equivalent to the continuum of VaR constraints. Portfolio x dominates portfolio y in the first order, if

$$VaR_p(L(x)) \leq VaR_p(L(y)) \ for \ all \ p \in (0, 1).$$

The CVaR at level p, roughly speaking, has the following form

$$CVaR_p(L(x)) = E[L(x)|L(x) \geq VaR_p(L(x))].$$

This formula is precise if $\mathrm{VaR}_p(L(x))$ is not an atom of the distribution of $L(x)$. More precisely we express it as follows:

$$\mathrm{CVaR}_p(L(x)) = \frac{1}{p}\int_0^p \mathrm{VaR}_t(L(x))dt.$$

We note that

$$\mathrm{CVaR}_p(L(x)) = -\frac{1}{p}F_{(-2)}(R(x), p). \quad (15.14)$$

Another description uses extremal properties of quantiles and equivalently represents CVaR as follows (Rockafellar and Uryasev 2000):

$$\mathrm{CVaR}_p(L(x)) = \inf_\eta\left\{\frac{1}{p}\mathsf{E}[(\eta - R(x))_+] - \eta\right\}. \quad (15.15)$$

A CVaR constraint on the portfolio x can be formulated as follows:

$$\mathrm{CVaR}_p(L(x)) \le \omega_p. \quad (15.16)$$

Using Equations (15.14) and (15.13) we conclude that the second order stochastic dominance relation for two portfolios x and y is equivalent to the continuum of CVaR constraints:

$$R(x) \succeq_{(2)} R(y) \Leftrightarrow \mathrm{CVaR}_p(L(x))$$
$$\le \mathrm{CVaR}_p(L(y)) \text{ for all } p \in (0,1]. \quad (15.17)$$

Assume that we compare the performance of a portfolio x with a random benchmark Y (for example, an index return rate or another portfolio return rate) requiring $R(x) \succeq_{(2)} Y$. Then the fraction ω_p of the initial capital allowed for risk exposure at level p is given by the benchmark Y:

$$\omega_p = \mathrm{CVaR}_p(-Y), \quad p \in (0,1].$$

Assume that Y has a discrete distribution with realizations y_i, $i = 1,\dots,m$. Then relation Equation (15.7) is equivalent to

$$\mathsf{E}[(y_i - R(x))_+] \le \mathsf{E}[(y_i - Y)_+], \quad i = 1,\dots,m. \quad (15.18)$$

This result *does not* imply that the continuum of CVaR constraints Equation (15.17) can be replaced by finitely many constraints of form

$$\mathrm{CVaR}_{p_i}(R(x)) \ge \mathrm{CVaR}_{p_i}(Y), \quad i = 1,\dots,m,$$

with some fixed probabilities p_i, $i = 1,\dots,m$. The reason is that we do not know at which probability levels the CVaR constraints have to be imposed.

15.4 The Dominance-Constrained Portfolio Problem

15.4.1 Direct Formulation

The starting point for our model is the assumption that a benchmark random return rate Y having a finite expected value is available. It may have the form of $Y = R(\bar{z})$, for some benchmark portfolio $\bar{z}$. It may be an index or our current portfolio. Our intention is to have the return rate of the new portfolio, $R(x)$, preferable over Y. Therefore, we introduce the following extension of the optimization problem Equation (15.3):

$$\max\ \mathsf{E}[R(x)] - \lambda\rho(R(x)) \quad (15.19)$$

subject to

$$R(x) \succeq_{(2)} Y, \quad (15.20)$$

$$x \in X. \quad (15.21)$$

Similarly to Equation (15.3), we optimize a mean-risk objective function, but we introduce a constraint that the portfolio return dominates a benchmark. Even when $\lambda = 0$ and we maximize just the expected value of the return rate, our model will still lead to nontrivial solutions, due to the presence of the dominance constraint Equation (15.20).

To increase flexibility of model Equations (15.19)–(15.21), we may also allow a uniform shift of $R(x)$ by a constant c, as in the following model:

$$\max\ \mathsf{E}[R(x)] - \lambda\rho(R(x)) - \delta c$$

subject to

$$R(x) + c \succeq_{(2)} Y,$$

$$x \in X.$$

Here $\delta > 0$ can be interpreted a cost of the shift c. Observe that the shift c may also become negative, in which case we are rewarded for uniformity of dominating Y. The shift c may be interpreted as an additional cash added to the return, and δ is the interest to be paid when the loan is paid back.

To simplify the derivations, from now on we focus on the simplest formulation of the dominance-constrained problem:

$$\max\ \mathsf{E}[R(x)] \quad (15.22)$$

subject to

$$R(x) \succeq_{(2)} Y, \quad (15.23)$$

$$x \in X. \quad (15.24)$$

We can observe the first advantage of our problem formulation: all data in it are readily available. Moreover, the set defined by Equation (15.23) is convex (Dentcheva and Ruszczyński 2003c; 2004a, c).

Let us assume now that Y has a discrete distribution with realizations y_i attained with probabilities π_i, $i = 1, \ldots, m$. We also assume that the return rates have a discrete joint distribution with realizations r_{jt}, $t = 1, \ldots, T$, $j = 1, \ldots, n$, attained with probabilities p_t, $t = 1, 2, \ldots, T$. Then the formulation of the stochastic dominance relation Equation (15.23) resp. Equation (15.18) simplifies even further. Introducing variables s_{it} representing the shortfall of $R(x)$ below y_i in realization t, $i = 1, \ldots, m$, $t = 1, \ldots, T$, we can formulate problem Equations (15.22)–(15.24) as follows:

$$\max \sum_{i=1}^{T} p_t \sum_{j=1}^{n} x_j r_{jt} \tag{15.25}$$

subject to

$$\sum_{j=1}^{n} x_j r_{jt} + s_{it} \geq y_i, \quad i = 1, \ldots, m, \quad t = 1, \ldots, T, \tag{15.26}$$

$$\sum_{t=1}^{T} p_t S_{it} \leq \sum_{l=1}^{m} \pi_k (y_i - y_k)_+, \quad i = 1, \ldots, m, \tag{15.27}$$

$$s_{it} \geq 0, \quad i = 1, \ldots, m, \quad t = 1, \ldots, T. \tag{15.28}$$

$$x \in X. \tag{15.29}$$

Indeed, or every feasible point x of (15.22)–(15.24), setting

$$s_{it} = \max\left(0, y_i - \sum_{j=1}^{n} x_j r_{jt}\right), \quad i = 1, \ldots, m,$$
$$t = 1, \ldots, T,$$

we obtain a feasible pair (x, s) for Equations (15.26)–(15.29). Conversely, for any feasible pair (x, s) for Equations (15.26)–(15.29), inequalities Equations (15.26) and (15.28) imply that

$$s_{it} \geq \max(0, y_i - \sum_{j=1}^{n} x_j r_{jt}), \quad i = 1, \ldots, m, \quad t = 1, \ldots, T.$$

Taking the expected value of both sides and using Equation (15.27) we obtain

$$F_2(R(x); y_i) \leq F_2(Y; y_i), \quad i = 1, \ldots, m.$$

Therefore, problem Equations (15.22)–(15.24) is equivalent to problem Equations (15.25)–(15.29).

If the set X is a convex polyhedron, problem Equations (15.25)–(15.29) becomes a large scale linear programming problem. It may be solved by general-purpose linear programming solvers. However, the size of the problem increases dramatically with the number of assets n, their return realizations T, and benchmark realizations m, which makes it impractical for even moderate dimensions (in thousands). For the purpose of solving these problems, we developed a specialized decomposition method presented in Dentcheva and Ruszczyński (2006a).

15.4.2 Inverse Formulation

Assume that the return rates have a joint discrete distribution realizations r_{jt}, $t = 1, \ldots, T$ and $j = 1, \ldots, n$, attained with probabilities p_t, $t = 1, 2, \ldots, T$. Moreover, we assume that *all probabilities p_t are equal*, that is, $p_t = 1/T$, $t = 1, \ldots, T$. This is the case of empirical distributions. Correspondingly, we assume that Y has a discrete distribution with $m = T$ equally probable realizations y_t, $t = 1, \ldots, T$.

We use the symbol $R_{[t]}(x)$ to denote the ordered realizations of $R(x)$; that is,

$$R_{[1]}(x) \leq R_{[2]}(x) \leq \ldots \leq R_{[T]}(x).$$

Since $R(x)$ has a discrete distribution, the functions $F_2(R(x); \cdot)$ and $F_{(-2)}(R(x); \cdot)$ are piecewise linear. Owing to the fact that all probabilities p_t are equal, the break points of $F_{(-2)}(R(x); \cdot)$ occur at t/T, for $t = 0, 1, \ldots, m$. The same applies to $F_{(-2)}(Y; \cdot)$. It follows from Equation (15.13) that the stochastic dominance constraint Equation (15.23) can be equivalently expressed as

$$F_{(-2)}\left(R(x); \frac{t}{T}\right) \geq F_{(-2)}\left(Y; \frac{t}{T}\right), \quad t = 1, \ldots, T.$$

Note that $F_{(-2)}(R(x); 0) = F_{(-2)}(Y; 0) = 0$. We have

$$F_{(-2)}\left(R(x); \frac{t}{T}\right) = \frac{1}{T} \sum_{k=1}^{t} R_{[k]}(x), \quad t = 1, \ldots, T.$$

Therefore problem Equations (15.22)–(15.24) can be written with an equivalent inverse form of the dominance constraint:

$$\max \mathsf{E}[R(x)] \text{ subject to} \tag{15.30}$$

$$\sum_{k=1}^{t} R_{[k]}(x) \geq \sum_{k=1}^{t} y_{[k]}, \quad t = 1, \ldots, T, \tag{15.31}$$

$$x \in X. \tag{15.32}$$

It was shown in (Ogryczak and Ruszczyński 2002) that the function $x \mapsto \sum_{k=1}^{t} R_{[k]}(x)$ is concave and positively homogeneous. It is also polyhedral. Therefore, Equation (15.31) are convex polyhedral constraints. If the set X is a convex polyhedron, problem Equations (15.30)–(15.32) has an equivalent linear programming formulation.

All these transformations are possible due to the crucial assumption that the probabilities of all elementary events are equal. If they are not equal, the break points of the function $F_{(-2)}(R(x); \cdot)$ depend on x, and therefore inequality Equation (15.13) cannot be reduced to finitely many convex inequalities. This is in contrast to the primal formulation, where the discreteness of Y alone was sufficient to reduce the stochastic dominance constraint to finitely many convex inequalities.

We have to observe that the quantile formulation Equation (15.31) of stochastic dominance constraints is more involved than the primal formulation, and requires more sophisticated computational methods. Using Equation (15.31) directly would require employing nonsmooth optimization methods to solve problem Equations (15.30)–(15.32). Equivalent formulation with linear constraints has very many constraints, because of the large number of pieces of the function $x \mapsto \sum_{k=1}^{t} R_{[k]}(x)$. Still, Dentcheva and Ruszczyński (2010) developed a highly efficient cutting plane method, which significantly outperforms direct approaches.

15.5 Optimality and Duality

15.5.1 Primal Form

From now on we assume that the probability distributions of the return rates are discrete with finitely many realizations realizations r_{jt}, $t = 1, \ldots, T$, $j = 1, \ldots, n$, attained with probabilities p_t, $t = 1, 2, \ldots, T$. We also assume that there are finitely many ordered realizations of the benchmark outcome $Y : y_1 < y_2 < \cdots < y_m$. The probabilities of these realizations are denoted by π_i, $i = 1, \ldots, m$. We also assume that the set X is compact.

We define the set U of functions $u : \mathsf{R} \to \mathsf{R}$ satisfying the following conditions:

- $u(\cdot)$ is concave and nondecreasing
- $u(\cdot)$ is piecewise linear with break points y_i, $i = 1, \ldots, m$
- $u(t) = 0$ for all $t \geq y_m$

It is evident that U is a convex cone.

Let us define the function $L : \mathsf{R}^n \times \mathsf{U} \to \mathsf{R}$ as follows

$$L(x, u) = \mathsf{E}[R(x) + u(R(x)) - u(Y)]. \qquad (15.33)$$

It will play for problem Equations (15.22)–(15.24) a similar role to that of a Lagrangian. It is well defined, because for every $u \in \mathsf{U}$ and every $x \in \mathsf{R}^n$ the expected value $\mathsf{E}[u(R(x))]$ exists and is finite.

The following theorem has been proved in a more general version in (Dentcheva and Ruszczyński 2003c).

Theorem 15.1. *If $\hat{x}$ is an optimal solution of Equations (15.22)–(15.24) then there exists a function $\hat{u} \in \mathsf{U}$ such that*

$$L(\hat{x}, \hat{u}) = \max_{x \in X} L(x, \hat{u}) \qquad (15.34)$$

$$\mathsf{E}[\hat{u}(R(\hat{x}))] = \mathsf{E}[\hat{u}(Y)]. \qquad (15.35)$$

Conversely, if for some function $\hat{u} \in \mathsf{U}$ an optimal solution $\hat{x}$ of Equation (15.34) satisfies Equations (15.23) and (15.35), then $\hat{x}$ is an optimal solution of Equations (15.22)–(15.24).

We can also develop duality relations for our problem. With the function Equation (15.33) we can associate the dual function

$$D(u) = \max_{x \in X} L(x, u).$$

We are allowed to write the maximization operation here, because the set X is compact and $L(\cdot, u)$ is continuous.

The dual problem has the following form

$$\min_{u \in \mathsf{E}} D(u). \qquad (15.36)$$

The set U is a closed convex cone and $D(\cdot)$ is a convex function, so Equation (15.36) is a convex optimization problem.

Theorem 15.2. *Assume that Equations (15.22)–(15.24) has an optimal solution. Then problem Equation (15.36) has an optimal solution and the optimal values of both problems coincide. Furthermore, the set of optimal solutions of Equation (15.36) is the set of functions $\hat{u} \in \mathsf{U}$ satisfying Equations (15.34)–(15.35) for an optimal solution $\hat{x}$ of Equations (15.22)–(15.24).*

Note that all constraints of our problem are linear or convex polyhedral, and therefore we do not need any constraint qualification conditions here.

The "Lagrange multiplier" u is directly related to the expected utility theory of von Neumann and Morgenstern. We have established earlier that the second order stochastic dominance relation is equivalent to Equation (15.8) for *all* utility functions in U. Our result shows that one of them, $\hat{u}(\cdot)$, assumes the role of a Lagrange multiplier associated with Equation (15.23). A point $\hat{x}$ is a solution to Equations (15.22)–(15.24) if there exists a utility function

$\hat{u}(\cdot)$ such that $\hat{x}$ maximizes over X the objective function $\mathsf{E}[R(x)]$ augmented with this dual utility. We see that the optimization problem in Equation (15.34) is equivalent to

$$\max_{x \in X} \mathsf{E}[v(R(x))], \qquad (15.37)$$

where $v(\eta) = \eta + u(\eta)$. At the optimal solution the function $\hat{v}(\eta) = \eta + \hat{u}(\eta)$ is the implied utility function. It attaches higher penalty to smaller realizations of $R(x)$ (bigger realizations of $L(x)$). By maximizing $L(R(x), u)$ we look for x such that the left tail of the distribution of $R(x)$ is thin.

It is important to stress that the optimal function $\hat{u}(\cdot)$ is piecewise linear, with break points at the realizations $y_1, \dots, y_m$ of the benchmark Y. Therefore, the dual problem has also an equivalent linear programming formulation.

15.5.2 Inverse Form

In addition to the assumption that all involved distributions are discrete, we also assume that all probabilities p_t are equal, and that $m = T$.

We introduce the set W of concave nondecreasing functions $w : [0, 1] \to \mathsf{R}$. It is evident that W is a convex cone.

Recall the identity

$$\mathsf{E}[R(x)] = \int_0^1 F_{(-1)}(R(x); p)dp.$$

Let us define the function $\Phi : X \times \mathsf{W} \to \mathsf{R}$, as follows

$$\Phi(x, w) = \int_0^1 F_{(-1)}(R(x); p)dp + \int_0^1 F_{(-1)}(R(x); p)\,dw(p)$$

$$- \int_0^1 F_{(-1)}(Y; p)\,dw(p). \qquad (15.38)$$

It plays a role similar to that of a Lagrangian of Equations (15.30)–(15.32).

Theorem 15.3. *If* $\hat{x}$ *is an optimal solution of Equations (15.30)–(15.32) then there exists a function* $\hat{w} \in \mathsf{W}$ *such that*

$$\Phi(\hat{x}, \hat{w}) = \max_{x \in X} \Phi(x, \hat{w}) \qquad (15.39)$$

$$\int_0^1 F_{(-1)}(R(\hat{x}); p)d\,\hat{w}(p) = \int_0^1 F_{(-1)}(Y; p)d\,\hat{w}(p).$$

$$(15.40)$$

Conversely, if for some function $\hat{w} \in \mathsf{W}$ *we find an optimal solution* $\hat{x}$ *of Equation (15.39) that satisfies Equations (15.31) and (15.40), then* $\hat{x}$ *is an optimal solution of Equations (15.30)–(15.32).*

We can also develop a duality theory based on Lagrangian Equation (15.38). For every function $w \in \mathsf{W}$ the problem

$$\max_{x \in X} \Phi(x, w) \qquad (15.41)$$

is a Lagrangian relaxation of problem Equations (15.30)–(15.32). Its optimal value, $\Psi(w)$, is always greater than or equal to the optimal value of Equations (15.30)–(15.32).

We define the dual problem as

$$\min_{w \in \mathsf{W}} \Psi(w). \qquad (15.42)$$

The set W is a closed convex cone and $\Psi(\cdot)$ is a convex function, so problem Equation (15.42) is a convex optimization problem. Duality relations in convex programming yield the following result.

Theorem 15.4. *Assume that problem Equations (15.30)–(15.32) has an optimal solution. Then problem Equation (15.42) has an optimal solution and the optimal values of both problems coincide. Furthermore, the set of optimal solutions of Equation (15.42) is the set of functions* $\hat{w} \in \mathsf{W}$ *satisfying Equations (15.39)–(15.40) for an optimal solution* $\hat{x}$ *of Equations (15.30)–(15.32).*

The "Lagrange multiplier" w in this case is related to rank dependent expected utility theory. We have established earlier that the second order stochastic dominance relation is equivalent to Equation (15.12) for *all* dual utility functions in W. Our result shows that one of them, $\hat{w}(\cdot)$, assumes the role of a Lagrange multiplier associated with Equation (15.31). A point $\hat{x}$ is a solution to Equations (15.30)–(15.32) if there exists a dual utility function $\hat{w}(\cdot)$ such that $\hat{x}$ maximizes over X the objective function $\mathsf{E}[R(x)]$ augmented with this dual utility. We can transform the Lagrangian Equation (15.38) in the following way:

$$\Phi(X, w) = \int_0^1 F_{(-1)}(R(x); p)dp + \int_0^1 F_{(-1)}(R(x); p)\,dw(p)$$

$$- \int_0^1 F_{(-1)}(Y; p)dw(p) = \int_0^1 F_{(-1)}(R(x); p)d\,v(p)$$

$$- \int_0^1 F_{(-1)}(Y; p)d\,w(p),$$

where $v(p) = p + w(p)$. At the optimal solution the function $\hat{v}(p) = p + \hat{w}(p)$ is the quantile utility function implied by the benchmark Y. Since $\int_0^1 F_{(-1)}(Y;p)dw(p)$ is fixed, the problem at the right hand side of Equation (15.39) becomes a problem of maximizing the implied rank dependent expected utility in X. It attaches higher weights to quantiles corresponding to smaller probabilities p. By maximizing $\Phi(R(x), w)$ we look for x such that the left tail of the distribution of $R(x)$ is thin.

Similarly to von Neumann–Morgenstern utility function, it is very difficult to elicit the dual utility function in advance. Our model derives it from a random benchmark.

The optimal function $\hat{w}(\cdot)$ is piecewise linear, with break points at $\frac{t}{T}$, $t = 1, \ldots, T$. Therefore, the dual problem has also an equivalent linear programming formulation. This property, however, is conditioned on the assumption of equal probabilities.

15.6 Numerical Illustration

We have tested our approach on a basket of 719 real-world assets, using 616 possible realizations of their joint return rates (Ruszczyński and Vanderbei 2003). Historical data on weekly return rates in the 12 years from spring 1990 to spring 2002 were used as equally likely realizations.

Implied utility functions corresponding to dominance constraints for four benchmark portfolios.

We have used four benchmark return rates Y. Each of them was constructed as a return rate of a certain index composed of our assets. As we actually know the past return rates, for the purpose of comparison we have selected equally weighted indexes composed of the N assets having the highest average return rates in this period. Benchmark 1 corresponds to $N = 26$, Benchmark 2 corresponds to $N = 54$, Benchmark 3 corresponds to $N = 82$, and Benchmark 4 corresponds to $N = 200$. Our problem was to maximize the expected return rate, under the condition that the return rate of the benchmark portfolio is dominated. Since the benchmark point was a return rate of a portfolio composed from the same basket, we have $m = T = 616$ in this case.

We have solved the problem by our method of minimizing the dual problem that was presented in Dentcheva and Ruszczyński (2006a).

The implied utility functions from Equation (15.37) obtained by solving the optimization problem Equation (15.34) in the optimality conditions are illustrated in Fig. 15.5. We see that for Benchmark Portfolio 1, which contains only a small number of fast-growing assets, the utility function is linear on almost the entire range of return rates. Only very negative return rates are penalized.

A different situation occurs when the benchmark portfolio contains more assets and is therefore more diversified and less risky. In order to dominate such a benchmark, we have to use a utility function which introduces penalty for a broader range of return rates and is steeper. For the broadly based index in Benchmark Portfolio 4, the optimal utility function is smoother and is nonlinear even for positive return rates. It
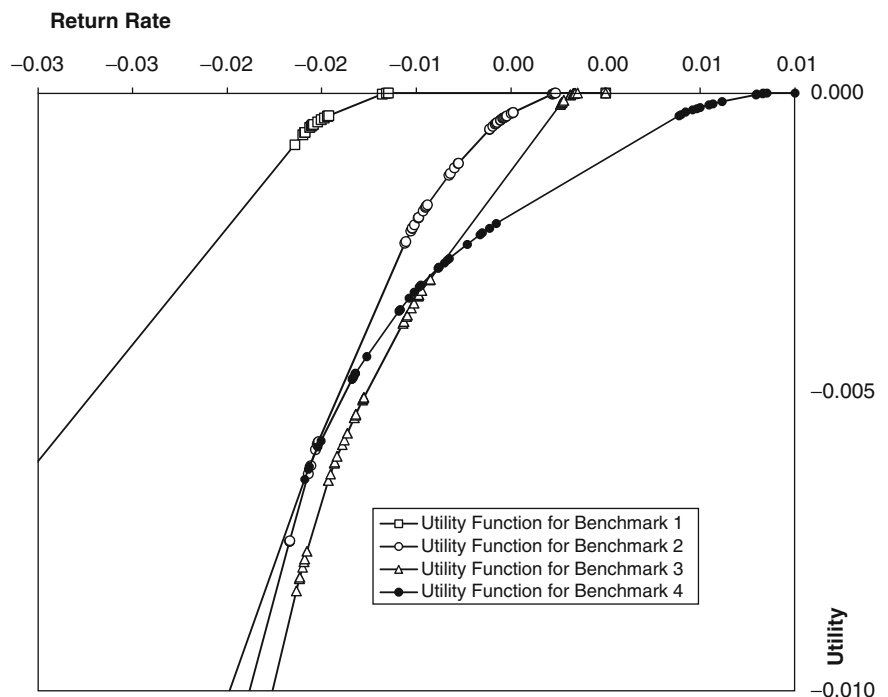

Fig. 15.5 Implied utility functions

is worth mentioning that all these utility functions, although nondecreasing and concave, have rather complicated shapes. It would be a very hard task to determine in advance the utility function that should be used to obtain a solution dominating our benchmark portfolio.

Obviously, the shape of the utility function is determined by the benchmark within the context of the optimization problem considered. If we change the optimization problem, the utility function will change.

Finally, we may remark that our model Equations (15.22)–(15.24) can be used for testing the statistical hypothesis that the return rate Y of the benchmark portfolio is nondominated.

15.7 Conclusions

We presented a new approach to portfolio selection based on stochastic dominance. The portfolio return rate in the new model is required to stochastically dominate a random benchmark, for example, the return rate of an index. We formulated optimality conditions and duality relations for these models and constructed equivalent optimization models with utility functions. Two different formulations of the stochastic dominance constraint: primal and inverse, lead to two dual problems that involve von Neuman–Morgenstern utility functions for the primal formulation and rank dependent (or dual) utility functions for the inverse formulation. The utility functions play the roles of Lagrange multipliers associated with the dominance constraints. In this way our model provides a link between the expected utility theory and the rank dependent utility theory. A numerical example illustrates the new approach and demonstrates the efficacy of the method.

Future challenges are extensions of the approach to multivariate and multistage outcomes and benchmarks.

References

Dentcheva, D. 2005. Optimization models with probabilistic constraints. in *Probabilistic and randomized methods for design under uncertainty*, G. Calafiore and F. Dabbene (Eds.). Springer, London, pp. 49–97.

Dentcheva, D. and A. Ruszczyński. 2003a. "Optimization under linear stochastic dominance." *Comptes Rendus de l'Academie Bulgare des Sciences* 56(6), 6–11.

Dentcheva, D. and A. Ruszczyński. 2003b. "Optimization under nonlinear stochastic dominance." *Comptes Rendus de l'Academie Bulgare des Sciences* 56(7), 19–25.

Dentcheva, D. and A. Ruszczyński. 2003c. "Optimization with stochastic dominance constraints." *SIAM Journal on Optimization* 14, 548–566.

Dentcheva, D. and A. Ruszczyński. 2004a. "Optimality and duality theory for stochastic optimization problems with nonlinear dominance constraints." *Mathematical Programming* 99, 329–350.

Dentcheva, D. and A. Ruszczyński. 2004b. "Convexification of stochastic ordering constraints." *Comptes Rendus de l'Academie Bulgare des Sciences* 57(3), 5–10.

Dentcheva, D. and A. Ruszczyński. 2004c. "Semi-infinite probabilistic optimization: first order stochastic dominance constraints." *Optimization* 53, 583–601.

Dentcheva, D. and A. Ruszczyński. 2006a. "Portfolio optimization with stochastic dominance constraints." *Journal of Banking and Finance* 30/2, 433–451.

Dentcheva, D. and A. Ruszczyński. 2006b. "Inverse stochastic dominance constraints and rank dependent expected utility theory." *Mathematical Programming* 108, 297–311.

Dentcheva, D. and A. Ruszczyński. 2008. "Duality between coherent risk measures and stochastic dominance constraints in risk-averse optimization." *Pacific Journal of Optimization* 4, 433–446.

Dentcheva, D. and A. Ruszczyński. 2010. Inverse cutting plane methods for optimization problems with second order stochastic dominance constraints, *Optimization* in press.

Elton, E. J., M. J. Gruber, S. J. Brown, and W. N. Goetzmann. 2006. Modern portfolio theory and investment analysis, Wiley, New York.

Fishburn, P. C.1964. Decision and value theory, Wiley, New York.

Fishburn, P.C. 1970. Utility theory for decision making, Wiley, New York.

Gastwirth, J. L. 1971. "A general definition of the Lorenz curve." *Econometrica* 39, 1037–1039.

Hadar, J. and W. Russell. 1969. "Rules for ordering uncertain prospects." *The American Economic Review* 59, 25–34.

Hanoch, G. and H. Levy, 1969. "The efficiency analysis of choices involving risk." *Review of Economic Studies* 36, 335–346.

Hardy, G. H., J. E. Littlewood and G. Polya. 1934. Inequalities, Cambridge University Press, Cambridge, MA.

Konno, H. and H. Yamazaki. 1991. "Mean-absolute deviation portfolio optimization model and its application to Tokyo stock market." *Management Science* 37, 519–531.

Lehmann, E. 1955. "Ordered families of distributions." *Annals of Mathematical Statistics* 26, 399–419.

Levy, H. 2006. Stochastic dominance: investment decision making under uncertainty, 2nd ed., Springer, New York.

Lorenz, M. O. 1905. "Methods of measuring concentration of wealth." *Journal of the American Statistical Association* 9, 209–219.

Markowitz, H. M. 1952. "Portfolio selection." *Journal of Finance* 7, 77–91.

Markowitz, H. M. 1959. Portfolio selection, Wiley, New York.

Markowitz, H. M. 1987. Mean-variance analysis in portfolio choice and capital markets, Blackwell, Oxford.

Mosler, K. and M. Scarsini (Eds.). 1991. Stochastic orders and decision under risk, Institute of Mathematical Statistics, Hayward, California.

Muliere, P. and M. Scarsini, 1989. "A note on stochastic dominance and inequality measures." *Journal of Economic Theory* 49, 314–323.

Müller, A. and D. Stoyan. 2002. Comparison methods for stochastic models and risks, Wiley, Chichester.

Ogryczak, W. and A. Ruszczyński. 1999. "From stochastic dominance to mean-risk models: semideviations as risk measures." *European Journal of Operational Research* 116, 33–50.

Ogryczak, W. and A. Ruszczyński. 2001. "On consistency of stochastic dominance and mean-semideviation models." *Mathematical Programming* 89, 217–232.

Ogryczak, W. and A. Ruszczyński. 2002. "Dual stochastic dominance and related mean-risk models." *SIAM Journal on Optimization* 13, 60–78.

Prékopa, A. 2003. Probabilistic programming. in *Stochastic Programming*, Ruszczyński, A. and A. Shapiro. (Eds.). Elsevier, Amsterdam, pp. 257–352.

Quiggin, J. 1982. "A theory of anticipated utility." *Journal of Economic Behavior and Organization* 3, 225–243.

Quiggin, J. 1993. Generalized expected utility theory – the rank-dependent expected utility model, Kluwer, Dordrecht.

Quirk, J.P and R. Saposnik. 1962. "Admissibility and measurable utility functions." *Review of Economic Studies* 29, 140–146.

Rockafellar, R. T. and S. Uryasev. 2000. "Optimization of conditional value-at-risk." *Journal of Risk* 2, 21–41.

Rothschild, M. and J. E. Stiglitz. 1969. "Increasing risk: I. A definition." *Journal of Economic Theory* 2, 225–243.

Ruszczyński, A. and R.J. Vanderbei. 2003. "Frontiers of stochastically nondominated portfolios." *Econometrica* 71, 1287–1297.

von Neumann, J. and O. Morgenstern. 1944. Theory of games and economic behavior, Princeton University Press, Princeton.

Wang, S. S., V. R. Yong, and H.H. Panjer. 1997. "Axiomatic characterization of insurance prices." *Insurance Mathematics and Economics* 21, 173–183.

Wang, S. S. and V. R. Yong. 1998. "Ordering risks: expected utility versus Yaari's dual theory of risk." *Insurance Mathematics and Economics* 22, 145–161.

Whitmore, G. A. and M. C. Findlay. (Eds.). 1978. Stochastic dominance: an approach to decision-making under risk, D.C.Heath, Lexington, MA.

Yaari, M. E. 1987. "The dual theory of choice under risk." *Econometrica* 55, 95–115.