



Institute for Empirical Research in Economics  
University of Zurich

Working Paper Series  
ISSN 1424-0459

**Published in: *Journal of Banking and Finance* 29 (4), pp. 895-926**

---

Working Paper No. 121

**Reward-Risk Portfolio Selection and Stochastic Dominance**

Enrico De Giorgi

August 2002

---

# Reward-Risk Portfolio Selection and Stochastic Dominance\*

Enrico De Giorgi<sup>†</sup>

13th August 2002

## Abstract

The portfolio selection problem is traditionally modelled by two different approaches. The first one is based on an axiomatic model of risk-averse preferences, where decision makers are assumed to possess an expected utility function and the portfolio choice consists in maximizing the expected utility over the set of feasible portfolios. The second approach, first proposed by Markowitz (1952), is very intuitive and reduces the portfolio choice to a set of two criteria, reward and risk, with possible tradeoff analysis. Usually the reward-risk model is not consistent with the first approach, even when the decision is independent from the specific form of the risk-averse expected utility function, i.e. when one investment dominates another one by second order stochastic dominance. In this paper we generalize the reward-risk model for portfolio selection. We define reward measures and risk measures by giving a set of properties these measures should satisfy. One of these properties will be the consistency with second order stochastic dominance, to obtain a link with the expected utility portfolio selection. We characterize reward and risk measures and we discuss the implication for portfolio selection.

**Keywords:** stochastic dominance, coherent risk measure, decision under risk, mean-risk models, portfolio optimization.

**JEL Classification:** G11.

## 1 Introduction

Markowitz (1952) introduced an intuitive model of return and risk for portfolio selection. This model is useful to guide one's intuition, and because of its simplicity it is also commonly used in practical finance decisions. Markowitz (1952) proposed to model return and risk in term of mean and variance, but he also suggested other measures of risk as for example the semivariance (Markowitz 1959).

The advantage of using the variance for describing the risk component of a portfolio, is principally due to the simplicity of the computation, but from the point of view of risk measurement the variance is not a satisfactory measure. First, the variance is a symmetric measure and "penalizes" gains and losses in the same way. Second, the variance is inappropriate to describe the risk of

---

\*I would like to thank Steiner Ekern for suggesting this problem. I am grateful to Thorsten Hens, Werner Hürlimann and Carlo Acerbi for valuable suggestions. Research of the author was financially supported by the Swiss National Science Foundation through NCCR FINRISK and by RiskLab, ETH Zürich, Switzerland.

<sup>†</sup>Contact address: degiorgi@math.ethz.ch. Institute for Empirical Research in Economics, University of Zurich, Blümlisalpstrasse 10, CH-8006 Zürich and RiskLab, D-MATH, Swiss Federal Institute of Technology, ETH-Zentrum, CH-8092 Zürich.

low probability events, as for example the default risk. Third, the mean-variance approach is not consistent with second order stochastic dominance and thus with the expected utility approach for portfolio selection. This is well illustrated by the  $(\mu, \sigma)$ -Paradox.

As already suggested by Markowitz (1959), Ogryczak and Ruszczyński (1997) also proposed semi-variance models, where the reward-risk approach is maintained, but the choice of semivariance instead of variance makes the model consistent with second order stochastic dominance. They also extend the consistency concept to higher order stochastic dominance by defining a more general central semideviation measure. Other risk measures have been proposed for portfolio selection, as for example Value-at-Risk (Jorion 1997, Duffie and Pan 1997) or Expected-Shortfall (Acerbi and Tasche 2001). The latter one is consistent with second order stochastic dominance. Value-at-Risk is widely used in practice: it is only consistent with respect to first order stochastic dominance (see Hürlimann 2002). Moreover, it has been shown by Artzner, Delbaen, Eber, and Heath (1997), that Value-at-Risk fails in controlling the risk of large losses with small probability, since it only considers the probability of certain losses to occur, but not the magnitude of these losses. Moreover, Value-at-Risk usually does not satisfy the sub-additivity property, which ensures - if satisfied - a reasonable behaviour of the risk measure when adding two positions. Artzner, Delbaen, Eber, and Heath (1999), being concerned with banking regulations, have proposed an axiomatic approach to the definition of a risk measure. They presented a set of four properties for measures of risk and they called measures satisfying these properties, *coherent risk measures*. Moreover, they show that a coherent risk measure can be still characterized by a non-empty set of scenarios (called generalized scenario), such that “any coherent risk measure arises as the supremum of the expected negative of final net worth for some collection of probability measures on the states of the world” (Artzner, Delbaen, Eber, and Heath 1999, Section 4.1). Unfortunately, coherent risk measures are usually not consistent with second order stochastic dominance. Pflug (1998) considered various classes of risk measures and gave the general properties for these classes, generalizing the approach of Ogryczak and Ruszczyński (1997) he introduced expectation-dispersion risk measures and showed that under some conditions they are consistent with stochastic dominance but not coherent.

We want to proceed analogously as Artzner, Delbaen, Eber, and Heath (1999), i.e. for *portfolio selection* we give and justify a set of properties for measures of reward and measures of risk. We give a characterization of these measures and we present some examples. Future researches will be devoted to the portfolio selection problem and to the question posed by Jaschke and Küchler (2000) about a possible extension of the Markowitz theory to the reward-risk framework.

As already pointed out, we would like to define a reward-risk framework for portfolio selection which is consistent (in some sense) with utility expectation. This is attained by imposing some specific axioms defining reward and risk measures. We give a brief overview of our axioms. A *reward measure* should be linear on the space of random variables. If this is not the case, it would be possible to increase (or decrease) the reward of a portfolio by just splitting it in two or more positions. Moreover, a reward measure should satisfy a risk-free condition, which states that the reward of a risk-free asset must be identical with the certain payoff (or change in value of the portfolio) provided by the risk-free asset. Finally, to ensure the link with the utility expectation approach, we impose a consistency of the reward measure with second order stochastic dominance. This last axiom implies that whenever a risk is preferred to another one by all rational, risk-averse expected utility maximizers, then this risk should provide an higher reward than the dominated risk.

Analogously, a *risk measure* is defined by the following axioms. First, the convexity, which ensures the diversification effect. In fact, if the risk measure were not convex, then it would be possible to reduce the risk of some portfolio by splitting the portfolio in two or more single positions. Second, a risk measure must be invariant under addition of a risk-free position. In a reward-risk

framework, the contribution of risk-free asset to the portfolio should be captured by the reward measure, since a risk-free asset gives a certain reward without adding risk. Third, the principle “no investment, no risk” leads to the very natural axiom that a zero position have zero risk. Finally, following the same argument introduced above for reward measure, we impose the consistency of the risk measure with second order stochastic dominance.

The paper is organized as follows: in Section 2 we present the portfolio selection problem and some standard results from decision theory. In Section 3, the definitions of reward and risk measures are given, with some example. In Section 4, we characterize reward measures and we present a possible simple characterization of risk measures. Section 5 concludes.

## 2 Portfolio selection and decision theory

We consider a two period economy. Let  $\Omega$  denote the state of nature in the final period  $t = T$ . We suppose that  $\Omega$  is finite and we write  $\Omega = \{1, \dots, S\}$ .

Let  $\mathcal{F}$  be a  $\sigma$ -algebra on  $\Omega$ . We take  $\mathcal{F} = 2^\Omega$ .  $(\Omega, \mathcal{F})$  define our measurable space. On  $(\Omega, \mathcal{F})$  we define a probability measure  $\mathbb{P}$ . We call  $\mathbb{P}$  the *physical probability measure* or the *objective probability measure*. We assume that  $\mathbb{P}[s] > 0$  for all  $s \in \Omega$ .

Since our analysis is addressed to the portfolio selection problem, we suppose that  $K + 1$  assets are available for investment and prices at time  $t = 0$  are given by  $q^k$  for  $k = 0, \dots, K$ . The payoff  $A^k \geq 0$  of asset  $k$  ( $k = 0, \dots, K$ ) is a random variable on  $\Omega$  giving the payoff  $A_s^k$  of asset  $k$  in each each state of nature  $s \in \Omega$ .

Let  $R^k = \frac{A^k}{q^k} - 1$  for  $k = 1, \dots, K$  and  $\mathbf{R} = (R^0, \dots, R^K)'$ .  $R^k$  gives the return of asset  $k$  at time  $T$ . We suppose that asset  $k = 0$  is a risk-free asset and  $\mathbb{P}[R^k < R^0] > 0$  for all  $k = 1, \dots, K$ : else a risky asset  $k'$  exists, which is still preferred to the risk-free asset 0 by every rational investor and thus asset 0 will never be selected.

We suppose that investor has decided to fully invest the initial wealth  $w_0$ . Consumption at  $t = 0$  has already passed. The portfolio selection problem consists in finding the number  $z^k$  of assets of type  $k$ , such that the portfolio  $\mathbf{z} = (z^0, \dots, z^K)'$  is *feasible* and is *optimal* with respect to some criterion. The final wealth of an investor with portfolio  $\mathbf{z}$  is given by

$$W = \sum_{k=0}^K z^k A^k, \quad (1)$$

and is a random variable on the state space  $\Omega$ . The portfolio selection problem has not been fully characterized, since it still depends on the two concept of “feasible” and “optimal”. Under the assumption that all the initial wealth is invested in the financial market, feasibility means that the cost at time  $t = 0$  of the portfolio  $\mathbf{z}$ , i.e.  $\sum_{k=0}^K q^k z^k$ , corresponds exactly to the initial wealth  $w_0$ , i.e.

$$\mathbf{z} \in \mathcal{B} = \{\mathbf{z} = (z^0, \dots, z^K)' \in \mathbb{R}^{K+1} \mid \mathbf{q}'\mathbf{z} = \sum_{k=0}^K q^k z^k = w_0\}, \quad (2)$$

where  $\mathbf{q} = (q^0, q^1, \dots, q^K)'$ . Moreover one can additionally introduce other constraints  $\mathcal{Z}$  (into the budget constraint (2)), as for example short-sale constraints

$$\mathcal{Z} = \{z = (z^0, \dots, z^K)' \in \mathbb{R}^{K+1} \mid z^k \geq 0 \text{ for } k = 0, \dots, K\},$$

which ensure that short-selling is not allowed, or diversification constraints

$$\mathcal{Z} = \{z = (z^0, \dots, z^K)' \in \mathbb{R}^{K+1} \mid z^k \in (\underline{\delta}_k, \bar{\delta}_k) \text{ for } k = 0, \dots, K\},$$

which give the minimal and maximal number of asset of each type to be purchased or sold. More general the portfolio  $\mathbf{z}$  is called feasible if  $\mathbf{z} \in \mathbb{B} \cap \mathcal{Z}$  for some  $\mathcal{Z}$ .

Optimality is a less obvious concept and the debate on the choice of a criterion with respect to which one should optimize the portfolio selection is still open, as discussed in the introduction. We briefly consider here two methodologies.

Before proceeding we want to rewrite the portfolio choice problem in an other way. We introduce the vector  $\boldsymbol{\lambda} = (\lambda^0, \lambda^1, \dots, \lambda^K)' \in \mathbb{R}^{K+1}$  giving the budget shares  $\lambda^k$  ( $k = 0, \dots, K$ ) invested in asset  $k$ . Formally we have

$$\lambda^k = \frac{q^k z^k}{w_0}. \quad (3)$$

$\boldsymbol{\lambda}$  is also called strategy, since it gives independently of  $w_0$ , the proportion of wealth to be invested in each asset. The feasibility or budget constraint (2) is now given as follows

$$\sum_{k=0}^K \lambda^k = 1,$$

thus  $\boldsymbol{\lambda} \in \Delta^K = \{x \in \mathbb{R}^{K+1} \mid \sum_{k=1}^{K+1} x_k = 1\}$ . The final wealth can be expressed by

$$W = w_0 \left( 1 + \sum_{k=0}^K R^k \lambda^k \right).$$

and thus the absolute change in wealth is given by

$$w_0 \sum_{k=0}^K R^k \lambda^k.$$

## 2.1 Preferences

In the general setting we suppose that an investor have some preference  $\{\succ\}$  on a subset  $\mathcal{H}$  of  $\mathcal{G} = \{X : \Omega \rightarrow \mathbb{R} \mid X \text{ is } \mathcal{F} \text{ - measurable}\}$ , the space of real-valued random variables on  $(\Omega, \mathcal{F}, \mathbb{P})$ . When an investor is asked to choose between to random outcomes  $X, Y \in \mathcal{H}$ , then one of the following can occur

- (i)  $X \succ Y$ , i.e.  $X$  is preferred to  $Y$ ;
- (ii)  $Y \succ X$  (or  $X \prec Y$ ), i.e.  $Y$  is preferred to  $X$ ;
- (iii)  $X \sim Y$ , i.e. investor is indifferent or indefinite between  $X$  and  $Y$ .

Note that  $\succ$ ,  $\prec$  and  $\sim$  are operators, while the corresponding preference relations  $\{\succ\}$ ,  $\{\prec\}$  and  $\{\sim\}$  are subsets of  $\mathcal{H} \times \mathcal{H}$ , e.g.  $\{\succ\} = \{(X, Y) \in \mathcal{H} \times \mathcal{H} \mid X \succ Y\}$ . Similarly,  $\{\prec\} = \{(X, Y) \in \mathcal{H} \times \mathcal{H} \mid (Y, X) \in \{\succ\}\}$  and  $\{\sim\} = \mathcal{H} \times \mathcal{H} \setminus (\{\succ\} \cup \{\prec\})$ . Thus, given  $\{\succ\}$  ( $\{\prec\}$ ) one can uniquely determine  $\{\prec\}$  ( $\{\succ\}$ ) and  $\{\sim\}$ . We define  $\{\succeq\} = \{\succ\} \cup \{\sim\}$  and we say that  $X$  is weakly preferred to  $Y$  when  $(X, Y) \in \{\succeq\}$ . More generally, a binary relation  $R$  on  $\mathcal{H}$  is a subset of  $\mathcal{H} \times \mathcal{H}$ .

**Definition 2.1.** *The binary relation  $R$  on  $\mathcal{H}$  is said to be*

- (i) reflexive if for  $X \in \mathcal{H}$  we have  $(X, X) \in R$ ;
- (ii) transitive if for  $X, Y, Z \in \mathcal{H}$  we have: if  $(X, Y) \in R$  and  $(Y, Z) \in R$  then  $(X, Z) \in R$ ;

- (iii) complete if for  $X, Y \in \mathcal{H}$ ,  $X \neq Y$  we have either  $(X, Y) \in R$  or  $(Y, X) \in R$ ;
- (iv) an equivalence if  $R$  is reflexive, transitive and complete.

The preference relation  $\{\succ\}$  is usually assumed to be transitive (rationality).  $\{\succ\}$  is not reflexive and usually also not complete. The preference relation  $\{\succeq\}$  is usually assumed to be transitive and reflexive, but not complete.

**Definition 2.2.** *The preference relation  $\{\succ\}$  is called*

- (i) a partial order if  $\{\succ\}$  is transitive;
- (ii) a weak order if  $\{\succ\}$  is transitive and  $\{\succeq\}$  is also transitive;
- (iii) a strict order if  $\{\succ\}$  is a complete weak order.

When  $\{\succ\}$  is a weak order, then  $\{\sim\}$  is an equivalence. Depending on the properties of the preference relations, one can represent these relations numerically. As suggested above, we propose here two methodology which has been used for portfolio selection purposes.

## 2.2 Value function

Following Yu (1985) we give the conditions such that a preference relation  $\{\succ\}$  on a set  $\mathcal{H}$  can be represented by a value function  $v : \mathcal{H} \rightarrow \mathbb{R}$ , i.e.  $(X, Y) \in \{\succ\} \Leftrightarrow v(X) > v(Y)$ . This case is interesting since the expected utility preference relations belong to this class of preferences, where the value function, called expected utility function, has some special form (see Example below). We assume that  $\{\succ\}$  is a weak order. Then it follows that  $\{\sim\}$  is an equivalence and one can define equivalence classes on  $\mathcal{H}$  by

$$[X] = \{Y \in \mathcal{H} \mid X \sim Y\}. \quad (4)$$

$[\mathcal{H}]$  denotes the set of equivalence classes on  $\mathcal{H}$  induced by  $\{\sim\}$ . We define an induced preference relation  $\{\succ_{\text{ind}}\}$  on  $[\mathcal{H}]$  by

$$[X] \succ_{\text{ind}} [Y] \Leftrightarrow X \succ Y, \quad \forall X \in [X], Y \in [Y].$$

One can easily verify that  $\{\succ_{\text{ind}}\}$  defines a strict order on  $[\mathcal{H}]$ .

**Definition 2.3.** *Let  $R$  be a binary relation on  $\mathcal{H}$ . Let  $\mathcal{K} \subset \mathcal{H}$ . Then  $\mathcal{K}$  is said to be  $\succ$ -dense in  $\mathcal{H}$  iff for  $X, Y \in \mathcal{H} \setminus \mathcal{K}$  with  $X \succ Y$ , there exists  $Z \in \mathcal{K}$  such that  $X \succ Z$  and  $Z \succ Y$ .*

**Proposition 2.1.** *The weak order  $\{\succ\}$  on  $\mathcal{H}$  has a value function representation if and only if there is a countable subset of  $[\mathcal{H}]$  that is  $\succ_{\text{ind}}$ -dense in  $\mathcal{H}$ .*

Remember, a weak order is a preference relation which is transitive and the corresponding preference relation  $\{\succ\}$  is also transitive.

### Example

Let  $\mathbb{P}$  be a probability measure on  $(\Omega, \mathcal{F})$ . Let  $u : \mathcal{H} \rightarrow \mathbb{R}$  be a real-valued function on  $\mathcal{H}$  such that  $u(X) \in L^1(\Omega, \mathcal{F}, \mathbb{P}) \forall X \in \mathcal{H}$ . We define a preference relation  $\{\succ_u\}$  on  $\mathcal{H}$  by

$$X \succ_u Y \Leftrightarrow \mathbb{E}[u(X)] > \mathbb{E}[u(Y)]. \quad (5)$$

By definition,  $\{\succ_u\}$  has a value function representation, where  $v(\cdot) = \mathbb{E}[u(\cdot)]$ . The representation (5) is called expected utility representation and  $u$  is the utility function.

For a given preference relation  $\{\succ\}$  on  $\mathcal{H}$  to possess an expected utility representation, strong assumptions are required, i.e.  $\{\succ\}$  has to be a weak order, with a countable  $\succ_{\text{ind}}$ -dense subset for  $[\mathcal{H}]$  and a very strong  $\succ$ -separability condition should be satisfied see (see Yu 1985, Chapter 5). Even when we assume that a preference relation possesses an expected utility representation, one should define the utility function  $u$  and this is not a trivial task. Fortunately, we can characterize the utility function  $u$  for a typology of investors and moreover, independently of the specific form of  $u$  within this typology of investors, some preferences hold.

The first assumption on  $u$  is that  $u$  is increasing, meaning that decision makers prefer more than less (rationality). The second assumption on  $u$  is that  $u$  is concave, i.e. a decision maker is risk averse (see Eichberger and Harper 1997).

Let  $\mathcal{U}_2$  denotes the set of increasing, concave utility functions.

**Definition 2.4 (Second order stochastic dominance).** *Let  $X, Y \in \mathcal{G}$ . Then  $X$  dominates  $Y$  by second order stochastic dominance (written  $X \succ_2 Y$ ) iff*

$$\mathbb{E}_{\mathbb{P}}[u(X)] \geq \mathbb{E}_{\mathbb{P}}[u(Y)] \quad (6)$$

for all  $u \in \mathcal{U}_2$  and there exists at least one element in  $\mathcal{U}$  such the inequality is strict.

The concept of second order stochastic dominance was first introduced by Rothschild and Stiglitz (1970). If  $X$  dominates  $Y$  by second order stochastic dominance, then  $X$  will be preferred to  $Y$  by any expected utility maximizers, who is risk averse and satisfies the ‘‘more than less’’ assumption. The following result holds.

**Lemma 2.1.** *For  $X, Y \in \mathcal{G}$  and  $\mathbb{P}$  a probability measure on  $(\Omega, \mathcal{P})$  are equivalent*

(i)  $X \succ_2 Y$ ;

(ii)  $\int_{-\infty}^x F_X(t) dt \leq \int_{-\infty}^x F_Y(t) dt, \forall x \in \mathbb{R}$  and the inequality is strict for at least one  $x \in \mathbb{R}$ , where  $F_X(t) = \mathbb{P}[X \leq t]$  and  $F_Y(t) = \mathbb{P}[Y \leq t]$ ;

(iii)  $\int_0^t F_X^{-1}(x) dx \geq \int_0^t F_Y^{-1}(t) dt, \forall t \in (0, 1]$  and the inequality is strict for at least one  $t \in (0, 1]$ , where  $F_X^{-1}(t) = \inf\{x \in \mathbb{R} | F_X(x) \geq t\}$  and analogously for  $F_Y^{-1}$ .

*Proof.* Levy (1998, Theorem 3.2, Theorem 4.2). □

### 2.3 Multiple-criteria decision

To describe a preference relation  $\{\succ\}$  on  $\mathcal{H}$ , one could assume that the preference relation is characterized by a set of  $m$  criteria. This means that one can find a function  $f : \mathcal{H} \rightarrow \mathbb{R}^m$  attributing to each element  $X \in \mathcal{H}$  a set of  $m$  criteria  $f_1(X), \dots, f_m(X)$  in  $\mathbb{R}^m$ . The decision problem is now transferred to the subset  $\mathcal{X} = f(\mathcal{H})$  of  $\mathbb{R}^m$ , i.e. one should define a preference relation  $\{\succ_f\}$  on  $\mathcal{X}$  such that

$$X \succ Y \Leftrightarrow f(X) \succ_f f(Y). \quad (7)$$

In our discrete setting, we require that  $m < S$ . In fact under our assumption on  $\Omega$  we can identify  $\mathcal{G}$  with  $R^S$  and consider a realization  $X(s)$  of  $X \in \mathcal{G}$  for  $s = 1, \dots, S$  as a criterion for the decision, i.e. one could define  $\tilde{f} : \mathcal{G} \rightarrow \mathbb{R}^S, X \mapsto (X(1), \dots, X(S))$ . Since  $\mathcal{H} \subseteq \mathcal{G}, \tilde{f}(\mathcal{H}) \subseteq \mathbb{R}^S$ .

The value function approach is a special case of a ‘‘multi’’-criteria decision, where  $m = 1$ . From now on we assume that  $m > 1$ . Using more then one criterion complicates the decision problem, but on the other side the multiple criteria representation is less restrictive allowing a characterization of a large class of preference relations.

On the subset  $\mathcal{X} = f(\mathcal{H}) \subseteq \mathbb{R}^m$  we define a very natural preference relation, assuming that  $f_1(X), \dots, f_m(X)$  are positive characteristics and therefore appreciated by the decision maker with is non satiable with respect to each  $f_j$  for  $j = 1, \dots, m$ . In other words, “more is better” for each criterion  $f_j$  ( $j = 1, \dots, m$ ). We define  $\{\succ_m\}$  by

$$\mathbf{x} \succ_m \mathbf{y} \Leftrightarrow \mathbf{x} > \mathbf{y}, \quad (8)$$

where  $\mathbf{x}, \mathbf{y} \in \mathbb{R}^m$  and  $\mathbf{x} > \mathbf{y}$  means that  $x_j \geq y_j$  for  $j = 1, \dots, m$  and  $\mathbf{x} \neq \mathbf{y}$ . This preference is also called a Pareto preference on  $\mathbb{R}^m$  and we denote it by  $\{\succ_{m,P}\}$  and we drop the index  $m$  if there is no confusion about the dimension of the criteria space. Note that  $\{\succ_P\}$  is not a weak order, since the induced relation  $\{\sim_P\}$  is not an equivalence.

**Definition 2.5.** *Let  $\mathbf{x}^* \in \mathcal{X}$ . Then  $\mathbf{x}^*$  is called a Pareto optimum, if and only there does not exist an element  $\mathbf{x} \in \mathcal{X}$  such that  $\mathbf{x} \succ_P \mathbf{x}^*$ .*

Under weak condition on  $\mathcal{X}$  we are able to prove the existence of Pareto optima on  $\mathcal{X}$ . We would like to find minimal conditions on  $\mathcal{X}$  such that each element  $\mathbf{x} \in \mathcal{X}$  is a Pareto optimum or is dominated by a Pareto optimum: if  $\mathcal{X}$  satisfies this property, we say that  $\mathcal{X}$  is nondominance bounded with respect to  $\{\succ_P\}$ . Nondominance boundedness ensures that the set of Pareto optima is the “good” set to look for. In fact, if  $\mathcal{X}$  is not nondominance bounded with respect to the Pareto preference, then it can occur that an element  $\mathbf{x} \in \mathcal{X}$  is not Pareto optimal but no Pareto optimum dominates it. Therefore, a decision maker considering only Pareto optima, maybe eliminates some “good” solution, which can also be the preferred one when additional information about preferences are introduced.

Note that if  $\mathcal{X}$  is a open set, then no Pareto optimum exists. Another property which could be problematic for the existence of Pareto optima, is the unboundedness of  $\mathcal{X}$ . We introduce the following definitions.

**Definition 2.6.**

$$\Lambda^{\ll} = \{\mathbf{d} \in \mathbb{R}^m \mid \mathbf{d} \ll \mathbf{0}\}, \quad (9)$$

$$\Lambda^{\leq} = \{\mathbf{d} \in \mathbb{R}^m \mid \mathbf{d} \leq \mathbf{0}\}, \quad (10)$$

$$\Lambda^{<} = \{\mathbf{d} \in \mathbb{R}^m \mid \mathbf{d} < \mathbf{0}\}. \quad (11)$$

where  $\mathbf{d} \ll \mathbf{0}$  means that  $d_i < 0 \forall i = 1, \dots, m$ ,  $\mathbf{d} \leq \mathbf{0}$  means that  $d_i \leq 0 \forall i = 1, \dots, m$  and  $\mathbf{d} < \mathbf{0}$  means that  $d_i \leq 0 \forall i = 1, \dots, m$ ,  $\mathbf{d} \neq \mathbf{0}$ .

The advantage of these definitions is that we can express the Pareto preference in the following way

$$\mathbf{x}, \mathbf{y} \in \mathcal{X}, \mathbf{x} \succ_P \mathbf{y} \Leftrightarrow \mathbf{y} \in \mathbf{x} + \Lambda^{<}. \quad (12)$$

Moreover

$$\{\mathbf{y} \in \mathcal{X} \mid \mathbf{x} \succ_P \mathbf{y}\} = \mathcal{X} \cap (\mathbf{x} + \Lambda^{<}). \quad (13)$$

defines the subset of elements in  $\mathcal{X}$  which are dominated by  $\mathbf{x}$ , and

$$\{\mathbf{y} \in \mathcal{X} \mid \mathbf{y} \succ_P \mathbf{x}\} = \mathcal{X} \cap (\mathbf{x} - \Lambda^{<}) \quad (14)$$

defines the subset of elements in  $\mathcal{X}$  which dominate  $\mathbf{x}$ . It follows that  $\mathbf{x} \in \mathcal{X}$  is a Pareto optimum if and only if  $\mathcal{X} \cap (\mathbf{x} - \Lambda^{<}) = \emptyset$ .



**Remark**

One can generalize this characterization and define preference relations on a subset  $\mathcal{X} \subseteq \mathbb{R}^q$  by giving the preferred cone  $\Lambda(\mathbf{x})$  for each  $\mathbf{x} \in \mathcal{X}$ , i.e.

$$\mathbf{y} \succ \mathbf{x} \Leftrightarrow \exists \alpha > 0 : \mathbf{y} - \mathbf{x} \in \alpha \Lambda(\mathbf{x}). \quad (15)$$

For the Pareto preference we have  $\Lambda(\mathbf{x}) = -\Lambda^{\leq}$  for all  $\mathbf{x} \in \mathcal{X}$ .

**Definition 2.7.** Let  $\Lambda$  be a convex cone and  $\bar{\Lambda}$  is its closure<sup>1</sup>.  $\mathcal{X}$  is said to be  $\Lambda$ -compact if for every  $\mathbf{x} \in \mathcal{X}$ ,  $\mathcal{X} \cap (\mathbf{x} - \bar{\Lambda})$  is compact.

**Proposition 2.2.** Assume  $\mathcal{X}$  is nonempty and  $\Lambda^{\leq}$ -compact, then

- (i) there exists a Pareto optimum in  $\mathcal{X}$ ,
- (ii)  $\mathcal{X}$  is nondominance bounded with respect to  $\{\succ_P\}$ .

*Proof.* Yu (1985, Theorem 3.3). □

In the sequel we give a general characterization of Pareto optima under various assumptions. We want to obtain simple methodologies for computing Pareto optima.

**Proposition 2.3.** Let  $\mathbf{x} \in \mathcal{X}$ . Then  $\mathbf{x}$  is a Pareto optimum on  $\mathcal{X}$  if and only if for any  $j \in \{1, \dots, m\}$   $\mathbf{x}$  uniquely maximizes  $y_j$  on  $\mathcal{X}_j(\mathbf{x}) = \{\mathbf{y} \in \mathcal{X} \mid y_k \geq x_k, k \neq j, k = 1, \dots, q\}$ .

*Proof.* Yu (1985, Theorem 3.4). □

**Proposition 2.4.** (i) If  $\mathbf{x} \in \mathcal{X}$  maximizes  $\boldsymbol{\lambda}'\mathbf{x}$  for some  $\boldsymbol{\lambda} \in \Lambda^{\gg}$  over  $\mathcal{X}$ , then  $\mathbf{x}$  is a Pareto optimum ( $\Lambda^{\gg} = -\Lambda^{\ll}$ ).

- (ii) If  $\mathbf{x} \in \mathcal{X}$  uniquely maximizes  $\boldsymbol{\lambda}'\mathbf{x}$  for some  $\boldsymbol{\lambda} \in \Lambda^{\geq}$  over  $\mathcal{X}$ , then  $\mathbf{x}$  is a Pareto optimum ( $\Lambda^{\geq} = -\Lambda^{\leq}$ ).

*Proof.* Yu (1985, Theorem 3.5). □

**Proposition 2.5.**  $\mathbf{x} \in \mathcal{X}$  is a Pareto optimum if for any  $i \in \{1, \dots, m\}$  there exists  $m - 1$  constants  $r(i) = \{r_k \mid k \neq i, k = 1, \dots, m\}$  so that  $\mathbf{x}$  uniquely maximizes  $y_i$  over  $\mathcal{X}(r(i)) = \{\mathbf{y} \in \mathcal{X} \mid y_k \geq r_k, k \neq i, k = 1, \dots, q\}$ .

*Proof.* Yu (1985, Theorem 3.6). □

An additional assumption on  $\mathcal{X}$  is that of  $\Lambda^{\leq}$ -convexity.

**Definition 2.8.**  $\mathcal{X}$  is said to be  $\Lambda^{\leq}$ -convex if and only if  $\mathcal{X} + \Lambda^{\leq}$  is a convex set.

**Lemma 2.2.** If  $f_j, j = 1, \dots, m$  are concave functions on a convex set  $\mathcal{H}$ , then  $\mathcal{X} = f(\mathcal{H})$  is  $\Lambda^{\leq}$ -convex.

*Proof.* Yu (1985, Theorem 3.7). □

**Proposition 2.6.** If  $\mathcal{X}$  is  $\Lambda^{\leq}$ -convex, then a necessary condition for  $\mathbf{x} \in \mathcal{X}$  to be a Pareto optimum, is that  $\mathbf{x}$  maximizes  $\boldsymbol{\lambda}'\mathbf{x}$  over  $\mathcal{X}$  for some  $\boldsymbol{\lambda} \in \Lambda^{\geq}$ .

*Proof.* Yu (1985, Theorem 3.8). □

---

<sup>1</sup> $\bar{A} = \bigcap \{F \mid F \text{ abgeschlossen, } A \subseteq F\}$

**Definition 2.9 (Efficiency on  $\mathcal{H}$ ).** *An random variable  $X \in \mathcal{H}$  is efficient if and only if  $f(X)$  is a Pareto optimum. The efficient-set, is the subset of efficient random variables in  $\mathcal{H}$ .*

**Theorem 2.1.** (i) *A necessary condition for  $X \in \mathcal{X}$  to be efficient is that for any  $j \in \{1, \dots, q\}$ ,  $f_j(X)$  uniquely maximizes  $f_j(Y)$  over  $\mathcal{H}_j(X) = \{Y \in \mathcal{H} \mid f_k(Y) \geq f_k(X), k \neq j, k = 1, \dots, m\}$ .*

*A sufficient condition for  $X$  to be efficient is that  $X$  is the unique element in  $\mathcal{H}$  which solve the above problem.*

(ii)  *$X$  is efficient if  $X$  maximizes  $\lambda' f(X)$  over  $\mathcal{H}$  for some  $\lambda \in \Lambda^>$ , or uniquely maximizes  $\lambda' f(X)$  over  $\mathcal{H}$  for some  $\lambda \in \Lambda^\geq$ .*

(iii) *A necessary condition for  $X \in \mathcal{H}$  to be efficient is that for any  $j \in \{1, \dots, m\}$  there are  $m - 1$  constants  $r(j) = \{r_k \mid k \neq j, k = 1, \dots, m\}$  so that  $f_j(X)$  uniquely maximizes  $f_j(Y)$  over  $\mathcal{H}(r(j)) = \{Y \in \mathcal{H} \mid f_k(Y) \geq r_k, k \neq j, k = 1, \dots, m\}$ .*

*A sufficient condition for  $X$  to be efficient is that  $X$  indeed is the unique maximum point of the above problem.*

*Proof.* Yu (1985, Theorem 3.14). □

**Theorem 2.2.** *If  $\mathcal{H}$  is a convex set and each  $f_i(X)$  ( $i = 1, \dots, m$ ) is concave, then for  $X \in \mathcal{H}$  to be efficient it is necessary that  $X$  maximizes  $\lambda' f(X)$  over  $\mathcal{H}$  for some  $\lambda \in \Lambda^\geq$ .*

*Proof.* Yu (1985, Theorem 3.15). □

### 3 Definition of reward and risk measure

Let  $\mathcal{H}^i = \left\{ w_0^i \sum_{k=0}^K R^k \lambda^{i,k} \mid \sum_{k=0}^K \lambda^{i,k} = 1 \right\}$  be the return set of all feasible portfolios of investor  $i$ ,  $i = 1, \dots, I$ , with initial wealth  $w_0^i$ . The set  $\mathcal{H}^i$  gives the absolute changes in value for each portfolio financed at time 0 by the initial wealth  $w_0^i$ . We consider absolute changes in values and not relative changes, because we believe that an increase in the initial wealth should not necessarily imply that one proportionally increase the ability or willingness to face with risky positions. In fact when we consider relative changes in wealth instead of absolute changes, the relative return of a portfolio remains unchanged with increasing or decreasing initial investment and thus the portfolio decision will be independent of the initial wealth.

We drop the index  $i$  whenever we discuss the general properties of the set  $\mathcal{H}^i$ .

**Remark** (i)  $\mathcal{H}$  is convex.

(ii)  $\mathcal{H}$  remains convex if one adds other *convex* constraints  $\mathcal{Z}$  into the budget constraint  $\mathbb{B}$ .

(iii) If we suppose suppose that the vector of returns  $\mathbf{R}$  has the joint distribution  $F$  under the physical probability measure  $\mathbb{P}$  on  $(\Omega, \mathcal{F})$ , then the set  $\mathcal{H}$  can be considered as a subset of the space of distribution functions, i.e.

$$\mathcal{H} \equiv \left\{ F(\lambda) \left( \frac{\cdot}{w_0} \right) \mid \mathbf{R}'\lambda \sim F(\lambda), \lambda' \mathbf{1} = 1, \mathbf{R} \sim F \right\}. \quad (16)$$

Using this equivalence between set of portfolio return and distribution function, we can define reward and risk measures on random variables or on their distribution functions (see Pflug 1998).

Let  $f : \mathcal{H} \rightarrow \mathcal{X}$  be a  $m$ -dimensional function giving a set of  $m$ -criteria describing the preference relation on  $\mathcal{H}$ , where  $\mathcal{X} = f(\mathcal{H})$ . Note that  $\mathcal{X}$  is  $\Lambda^{\leq}$ -convex (see Definition 2.8), if  $f_i$  is concave for all  $i = 1, \dots, m$  (Lemma 2.2). As we did before, we suppose that  $f$  induces a Pareto preference on the set  $\mathcal{X}$ . We restrict our analysis to the case where investors select only two characteristics ( $m = 2$ ), a reward function  $f_1 = \mu$  and a risk function  $\rho = -f_2$ , that we will define below.  $f_1$  and  $f_2$  are identical for all investors, who only differs by the different “weights” they put on one characteristic or on the other one. Moreover, we assume that  $f_1$  and  $f_2$  are restriction on  $\mathcal{H}$  of functions defined on  $\mathcal{G}$ , the set of all random variable on  $(\Omega, \mathcal{F}, \mathbb{P})$ .

**Definition 3.1 (Reward measure).** A reward measure  $\mu : \mathcal{G} \rightarrow \mathbb{R}$  satisfies the following conditions

(i) Linearity.  $\forall X, Y \in \mathcal{G}, \alpha \in \mathbb{R}$  we have:

$$\mu(X + Y) = \mu(X) + \mu(Y), \quad (17)$$

$$\mu(\alpha X) = \alpha \mu(X). \quad (18)$$

(ii) Risk-free condition

$$\mu(X(\mathbf{e}_j)) = R^0, \quad (19)$$

where  $X(\mathbf{e}_j)$  is the value change in the portfolio where only one unit of initial wealth is invested in asset  $j$ ,  $j = 0, \dots, K$  and asset  $j = 0$  is the risk-free asset,  $\mathbf{e}_j = (0, \dots, 0, \underbrace{1}_{j-te}, 0, \dots, 0)'$ .

(iii) Stochastic dominance. For  $X, Y \in \mathcal{G}$  we have

$$X \succ_2 Y \Rightarrow \mu(X) \geq \mu(Y). \quad (20)$$

We say that  $\mu$  is **isotonic** with respect to second order stochastic dominance (following the nomenclature given by Pflug (1998)).

**Remark (Justification of the axioms)**

Analogously to the mean in the  $(\mu, \sigma)$ -approach of Markowitz, the reward measure is a location measure for the performance of the portfolio. Remember that we consider the absolute changes in value of a position.

The *linearity condition* says that one cannot increase or decrease the reward of portfolios by just separating the single positions. In fact, if the reward measure were not linear, then one can find risks  $X$  and  $Y$  in  $\mathcal{G}$  and  $\alpha \in \mathbb{R}$  such that the reward of the portfolio  $X + \alpha Y$  is greater (or less) than the sum of the reward of the single positions. This means that one can increase or decrease the reward by taking single positions in a portfolio. A reward measure should avoid this behaviour.

The *risk-free condition* states that the reward of a certain payoff, correspond to the difference between this payoff and the initial investment. Suppose that one invests all her wealth  $w_0$  in a risk-free asset with return  $R^0$ . Then the final wealth is given by  $w_0(1 + R^0)$  and thus the absolute change in wealth by  $w_0 R^0$ . Since there is no risk in that position (and thus no dispersion), one would expect that the reward measure gives exactly the same amount  $w_0 R^0$  to describe the location of the portfolio. If this is not the case and linearity holds, one could still normalize a reward measure such that the risk-free condition is satisfied.

Finally, the *isotonicity with respect to second order stochastic dominance* ensures the link to the expected utility approach for portfolio selection. In fact, whenever a risk  $X$  dominates another risk  $Y$  by second order stochastic dominance, then every risk-averse expected utility maximizer, prefers  $X$  to  $Y$ . We want that our reward measure  $\mu$  capture this preference and thus we impose the isotonicity with respect to second order stochastic dominance.

**Example**

The expectation  $\mathbb{E}_{\mathbb{P}}[\cdot]$ , where  $\mathbb{P}$  is the physical measure with respect to which second order stochastic dominance is defined, is a reward measure. We will show that  $\mathbb{E}_{\mathbb{P}}[\cdot]$  is the unique reward measure on  $\mathcal{G}$ .

Analogously to the reward measure we introduce now the definition of a risk measure and the properties we believe that such a measure should satisfy.

**Definition 3.2 (Risk measure).** A risk measure  $\rho : \mathcal{G} \rightarrow \mathbb{R}$  satisfies the following conditions

(i) Convexity.  $\forall X, Y \in \mathcal{G}, \alpha \in (0, 1)$  we have:

$$\rho(\alpha X + (1 - \alpha)Y) \leq \alpha\rho(X) + (1 - \alpha)\rho(Y). \quad (21)$$

(ii) Risk-free condition. For  $X \in \mathcal{G}, \alpha \in \mathbb{R}$  we have

$$\rho(X + \alpha X(\mathbf{e}_0)) = \rho(X). \quad (22)$$

(iii) Zero payoff. We have

$$\rho(0) = 0. \quad (23)$$

(iv) Stochastic dominance. For  $X, Y \in \mathcal{G}$  we have

$$X \succ_2 Y \Rightarrow \rho(X) \leq \rho(Y), \quad (24)$$

i.e.  $\rho$  is isotonic with respect to second order stochastic dominance.

Moreover we say that a risk measure  $\rho$  is positive homogeneous iff

$$\forall X \in \mathcal{G}, \alpha > 0 \Rightarrow \rho(\alpha X) = \alpha\rho(X). \quad (25)$$

**Remark (Justification of the axioms)**

The convexity ensures the diversification effect. In fact, when convexity is not satisfied for some  $X, Y \in \mathcal{G}$  and  $\alpha \in (0, 1)$ , then one could split the portfolio  $\alpha X + (1 - \alpha)Y$  in two parts, hold  $\alpha$  times the position  $X$  and  $1 - \alpha$  times the position  $Y$  and consequently reduce the risk. Similar arguments are provided by Artzner, Delbaen, Eber, and Heath (1997, 1998).

The *risk-free condition* says that adding a risk free position to the portfolio does not change the risk! This is different as in the definition of coherent risk measure introduced by Artzner, Delbaen, Eber, and Heath (1999), where the authors interpret a risk measure as the minimal extra cash one should add to his risky position and allocate “prudently”, to make the investment acceptable. Dealing with portfolio selection, we suggest that the contribution of a risk free position to the portfolio should be captured by the reward measure and not by the risk measure.

The *zero payoff condition* is a natural assumption, which states that “no investment” means no risk.

Finally the *isotonicity with respect to second order stochastic dominance*, analogously to the same property for the reward measure, ensures that an investment which is preferred to another one by all risk averse rational expected utility maximizers, is at most so risky as the dominated investment.

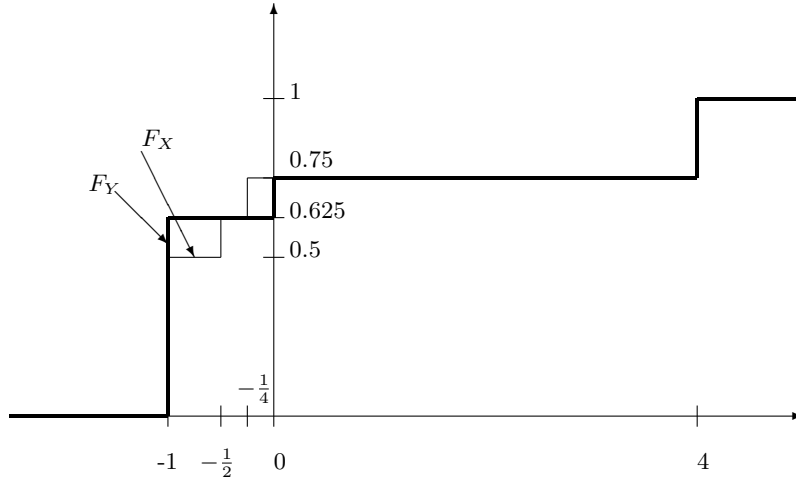
**Example (and counterexample)** (i) A coherent risk measure in the sense of Artzner, Delbaen, Eber, and Heath (1999) is not a risk measure as defined above. First, the risk free condition is not satisfied since for a coherent risk measure  $\tilde{\rho}$ , we have  $\tilde{\rho}(X(\mathbf{e}_0)) = -R^0$ . Second, a coherent risk measure is usually not isotonic with respect to second order stochastic

dominance. Consider the following simple example.

Let  $\Omega = \{1, 2, 3, 4\}$  and  $\mathcal{F} = 2^\Omega$ . Moreover we define  $X$  and  $Y$  on  $\Omega$  by

$$\begin{aligned} X(1) = -1, \quad X(2) = -0.5, \quad X(3) = -0.25 \quad \text{and} \quad X(4) = 4; \\ Y(1) = -1, \quad Y(2) = -1, \quad Y(3) = 0 \quad \text{and} \quad Y(4) = 4. \end{aligned}$$

Under the probability measure  $\mathbb{P} = (0.5, 0.125, 0.125, 0.25)$  on  $\Omega$ ,  $X$  dominates  $Y$  by second order stochastic dominance (see picture).



We generate now a coherent risk measure  $\tilde{\rho}$  such that  $\tilde{\rho}(Y) < \tilde{\rho}(X)$ . A risk measure  $\tilde{\rho}$  is coherent, if and only if there exists a non-empty family  $\mathcal{P}$  of probability measures on  $\Omega$  such that

$$\tilde{\rho}(X) = \sup\{\mathbb{E}_{\mathbb{Q}}[-X] \mid \mathbb{Q} \in \mathcal{P}\},$$

by Artzner, Delbaen, Eber, and Heath (1999, Proposition 4.1).

We choose  $\mathcal{P} = \{\mathbb{P}, \mathbb{P}_1, \mathbb{P}_2\}$  where  $\mathbb{P}_1 = (0.5, 0, 0.5, 0)$  and  $\mathbb{P}_2 = (0, 0, 0.5, 0.5)$  then it follows:

$$\begin{aligned} \tilde{\rho}(X) &= \max\{\mathbb{E}_{\mathbb{P}_i}[-X] \mid i = 0, 1, 2\} = \max\{-0.40625, 0.625, -2.375\} = 0.625; \\ \tilde{\rho}(Y) &= \max\{\mathbb{E}_{\mathbb{P}_i}[-Y] \mid i = 0, 1, 2\} = \max\{-0.375, 0.5, -2.5\} = 0.5. \end{aligned}$$

Naturally, a necessary condition for this results, is that  $\mathcal{P}$  contains at least one measure  $\tilde{\mathbb{Q}}$  such that  $Y$  is not dominated by  $X$  by second order stochastic dominance under  $\tilde{\mathbb{Q}}$ . In our example, the measures  $\mathbb{P}_i$  for  $i = 1, 2$  are such measures. Moreover, under  $\tilde{\rho}$ , the random variable  $Z = Y + 0.5$  has non-positive risk. However, under scenarios  $s = 1$  and  $s = 2$ ,  $Z$  has negative values, i.e. the future net worth is negative.

(ii) The Minkowski-gauge of a random variable  $X \in \mathcal{G}$  is defined as follows.

**Definition 3.3.** Let  $h$  be a convex, monotonic function on  $\mathbb{R}^+$  with  $h(0) = 0$ . Then

$$\|X\|_h = \inf\{a \mid \mathbb{E}_{\mathbb{P}}\left[h\left(\frac{|X|}{a}\right)\right] \leq 1\} \quad (26)$$

is called the Minkowski-gauge of  $X$  under  $h$ .

The Minkowski-gauge is homogeneous and fulfills the triangle inequality.

Let  $\rho(X) = \|(X - \mathbb{E}_{\mathbb{P}}[X])^-\|_h$ , where  $x^- = \min\{0, x\}$ , then  $\rho$  is subadditive, positive homogeneous (and therefore convex). Moreover  $\rho$  is isotonic with respect to second order stochastic dominance.

*Proof.* (a) The positive homogeneity of  $\rho$  follows directly from the homogeneity of the Minkowski-gauge.

(b) Subadditivity and isotonicity with respect to second order stochastic dominance are proved in Pflug (1998, Proposition 2.3). □

Since obviously  $\rho(0) = 0$  and  $\rho(X + \alpha X(\mathbf{e}_0)) = \rho(X)$ ,  $\rho$  defines a risk measure.

**Remark** (i) The risk-free condition implies that

$$\rho(X(\mathbf{e}_0)) = \rho(\alpha X(\mathbf{e}_0) + (1 - \alpha)X(\mathbf{e}_0)) = \rho(\alpha X(\mathbf{e}_0)).$$

For  $\alpha = 0$  we obtain that  $\rho(X(\mathbf{e}_0)) = \rho(0) = 0$ , i.e. the risk-free asset has risk zero.

(ii) For  $X \in \mathcal{G}$  and  $\alpha \in (0, 1)$  we have

$$\rho(\alpha X) \leq \alpha \rho(X).$$

(iii) For  $X \in \mathcal{G}$ ,  $X < 0$  and  $\alpha \in (0, 1)$  we have  $0 \succ_2 \alpha X \succ_2 X$  and thus

$$0 \leq \rho(\alpha X) \leq \alpha \rho(X) \leq \rho(X),$$

by the isotonic property and the convexity.

(iv) For  $X \in \mathcal{G}$ ,  $X > 0$  and  $\alpha \in (0, 1)$  we have  $X \succ_2 \alpha X \succ_2 0$  and thus

$$\rho(X) \leq \rho(\alpha X) \leq \alpha \rho(X) \leq 0,$$

by the isotonic property and the convexity. On the other side for  $\alpha = \max_{s \in \Omega} (X(s))$  and  $\epsilon > 0$ ,  $(\alpha + \epsilon)X(\mathbf{e}_0) \succ_2 X$  and thus  $\rho(X) \geq \rho((\alpha + \epsilon)X(\mathbf{e}_0)) = 0$ . It follows that for  $X > 0$ ,  $\rho(X) = 0$ .

Given a reward measure and a risk measure, we can now define a *reward-risk pair* on  $\mathcal{G}$  as follows:

**Definition 3.4.** Let  $\mu$  be a reward measure on  $\mathcal{G}$  and  $\rho$  be a risk measure on  $\mathcal{G}$ . Then the pair  $(\mu, \rho)$  is called a *reward-risk pair* on  $\mathcal{G}$ .

**Remark** (i) Given a reward-risk pair  $(\mu, \rho)$ , a necessary condition for a portfolio  $X \in \mathcal{H}$  to be efficient is that  $X$  minimizes  $\xi_1 \rho - \xi_2 \mu$  over  $\mathcal{H}$  for some  $(\xi_1, \xi_2)' \in \Lambda^>$  (Theorem 2.1). If  $\xi_1 = 0$ , then an investor is risk neutral and it takes care only on the reward  $\mu$ . We assume that  $\xi_1 \neq 0$ . If  $\xi_2 = 0$ , then the minimization problem has an obvious solution on  $\mathcal{H}$ , i.e.  $X = w_0 X(\mathbf{e}_0)$ . Thus we assume that both  $\xi_1$  and  $\xi_2$  are positive. In this case we can minimize the objective function  $\xi \rho - \mu$  where  $\xi = \frac{\xi_1}{\xi_2} > 0$ . This function can also be interpreted as a utility function depending on  $\mu$  and  $\rho$ , since  $\mu$  is linear and  $\rho$  is convex.

(ii) Given a *reward-risk pair*  $(\mu, \rho)$ , one can easily construct a combined measure of location and dispersion as suggested by Pflug (1998). We define the measure  $\mathcal{R}_\xi$  on  $\mathcal{G}$  by

$$\mathcal{R}_\xi(X) = \xi \rho(X) - \mu(X),$$

where  $\xi > 0$ . Then the following holds:

- (a)  $X \succ_2 Y \Rightarrow \mathcal{R}_\xi(X) \leq \mathcal{R}_\xi(Y)$ , i.e.  $\mathcal{R}_\xi$  is isotonic with respect to the second order stochastic dominance;
- (b)  $\mathcal{R}_\xi(X + R^0) = \xi\rho(X) - \mu(X) - R^0 = \mathcal{R}_\xi(X) - R^0$ , i.e.  $\mathcal{R}_\xi$  is translation-equivariant;
- (c)  $\mathcal{R}_\xi(\alpha X + (1 - \alpha)Y) \leq \alpha\mathcal{R}_\xi(X) + (1 - \alpha)\mathcal{R}_\xi(Y)$ , i.e.  $\mathcal{R}_\xi$  is a convex measure.

If we additionally assume that  $\rho$  is positive homogeneous, then  $\mathcal{R}_\xi$  is a coherent measure in the sense of Artzner, Delbaen, Eber, and Heath (1999). Note that the isotonicity with respect to second order stochastic dominance, is a stronger condition than just monotonicity, as suggested by Artzner, Delbaen, Eber, and Heath (1999): in fact it is well known that for  $X, Y \in \mathcal{G}$  such that  $X > Y$ , then  $X \succ_2 Y$  but the opposite is not true. Moreover, the measure  $\mathcal{R}_\xi$  depends on the preferences of the investor through  $\xi$ . This suggests that set of acceptable risks  $\{X \in \mathcal{G} \mid \mathcal{R}_\xi(X) \geq 0\}$  introduced by Artzner, Delbaen, Eber, and Heath (1999) and which uniquely defines a coherent measure depends on the investor's preferences as one should expect.

## 4 Characterization of reward and risk measures

Given the properties we would like that reward and risk measures satisfy, one can ask the question whether there is a characterization of these measures. As we have already seen (Examples), one can find some functional on  $\mathcal{G}$  which fulfills the conditions defining reward, respectively risk measures. In this section we prove that on  $\mathcal{G}$  we can find only one linear functional satisfying the risk-free condition and the isotonicity with respect to the second order stochastic dominance, namely the expectation with respect to the physical probability measure, which is also the one giving the stochastic order between random variables. For risk measures we cannot prove a uniqueness result, but we can still find a family of functionals on  $\mathcal{G}$ , which are risk measures.

### 4.1 Reward measure

**Theorem 4.1.** *Let  $\mu : \mathcal{G} \rightarrow \mathbb{R}$  be a reward measure and  $\mathbb{P} = (P(1), \dots, P(S))$  be the physical probability measure on  $(\Omega, \mathcal{F})$  (second order stochastic dominance is defined with respect to  $\mathbb{P}$ ). We assume that  $P(s) > 0 \forall s \in \Omega$ . Then*

$$\mu(X) = \mathbb{E}_{\mathbb{P}}[X]. \quad (27)$$

*Proof.* (i) Since  $\Omega$  is finite and  $\text{card}(\Omega) = S$ , we can identify  $\mathcal{G}$  with  $\mathbb{R}^S$  and the reward measure  $\mu$  with a linear functional  $\mathbb{R}^S \rightarrow \mathbb{R}$ . We define the usual scalar product on  $\mathbb{R}^S$  by

$$\langle X, Y \rangle = \sum_{s=1}^S X(s)Y(s). \quad (28)$$

Then  $(\mathcal{G}, \langle \cdot, \cdot \rangle)$  is an Hilbert space and thus by the the Riesz's Representation Theorem there exists a unique  $P^* \in \mathcal{G}$  such that

$$\mu(X) = \langle P^*, X \rangle = \sum_{s=1}^S P^*(s)X(s). \quad (29)$$

The riskfree condition for  $\mu$  implies

$$R^0 = \mu(X((e)^0)) = \sum_{s=1}^S P^*(s)R^0 = R^0 \sum_{s=1}^S P^*(s).$$

Thus  $\sum_{s=1}^S P^*(s) = 1$ .

(ii) Consider now  $X^s \in \mathcal{G}$  for  $s = 1, \dots, S$  defined by

$$X^s(s') = \begin{cases} 1 & \text{if } s' = s, \\ 0 & \text{if } s' \neq s, \end{cases}$$

for  $s' = 1, \dots, S$ . Then  $X^s > 0$  and thus  $X^s \succ_2 0 \forall s$ . By the isotonic property it follows  $\mu(X^s) \geq \mu(0) = 0$ . On the other side we have

$$0 \leq \mu(X^s) = \langle P^*, X^s \rangle = \sum_{s'=1}^S P^*(s') X^s(s') = P^*(s).$$

Therefore, the  $S$ -tuple  $\mathbb{P}^* = (P^*(1), \dots, P^*(S))$  defines a probability measure on  $\Omega$  with

$$\mu(X) = \mathbb{E}_{\mathbb{P}^*} [X].$$

(iii) We show:  $P(s) > P(s') \Rightarrow P^*(s) \geq P^*(s')$ .

For  $s, s' \in \Omega$  such that  $P(s) > P(s')$  we have  $X^s \succ_2 X^{s'}$  and therefore by isotonicity with respect to second order stochastic dominance,  $\mu(X^s) \geq \mu(X^{s'})$ . Since  $P^*(X^s) = \mu(X^s)$  the statement follows.

(iv) We show:  $P(s) = P(s') \Rightarrow P^*(s) = P^*(s')$ .

Suppose that  $P(s) = P(s') > 0$  for  $s, s' \in \Omega$  and  $P^*(s) > P^*(s')$ . Then there exists  $\epsilon > 1$  such that  $P^*(s) > \epsilon P^*(s')$ . Since  $P(s) = P(s')$  we have  $\epsilon X^{s'} \succ_2 X^s$ . Thus by isotonicity and linearity of  $\mu$  we obtain

$$\epsilon P^*(s') = \epsilon \mu(X^{s'}) = \mu(\epsilon X^{s'}) \geq \mu(X^s) = P^*(s) > \epsilon P^*(s').$$

A contradiction. Thus we must have  $P^*(s) \leq P^*(s')$ . By the same argument we can exclude  $P^*(s) < P^*(s')$  and therefore  $P^*(s) = P^*(s')$ .

(v) We show:  $\mathbb{P}^* = \mathbb{P}$ .

Let  $X \in \mathcal{G}$ . Let  $Z : \mathcal{G} \rightarrow \mathbb{R}$  be the density of  $\mathbb{P}^*$  with respect to  $\mathbb{P}$ , i.e.  $Z(s) = \frac{P^*(s)}{P(s)}$  for  $s \in \Omega$  (the density exists since  $P(s) > 0$  for all  $s \in \Omega$ ). Following the idea of Dana (2002), we define

$$\Psi(X, Z) = \arg \min \{ \mathbb{E}_{\mathbb{P}} [ZY] \mid Y \in \mathcal{G}, Y \succ_2 X, Y \sim_d X \}.$$

For  $\mu$  to be consistent with respect to  $\succ_2$ , we should obtain that  $X \in \Psi(X, Z)$ . If this is not true, we find  $Y \in \mathcal{G}$  with  $Y \succ_2 X$  but  $\mu(Y) = \mathbb{E}_{\mathbb{P}^*} [Y] = \mathbb{E}_{\mathbb{P}} [ZY] < \mathbb{E}_{\mathbb{P}} [ZX] = \mathbb{E}_{\mathbb{P}^*} [ZX] = \mu(X)$ , which contradict the isotonicity of  $\mu$ .

For  $Y \in \Psi(X, Z)$  we have

$$Y(s) > Y(s') \Rightarrow Z(s) \leq Z(s'). \quad (30)$$

We say that  $Y$  and  $Z$  are anticomotone. To prove this last statement we proceed by contradiction. We assume that we find  $s, s' \in \Omega$  such that  $Y(s) > Y(s')$  but  $Z(s) > Z(s')$ . Suppose without loss of generality that  $P(s) < P(s')$ . The case  $P(s) > P(s')$  can be treated similarly. The case  $P(s) = P(s')$  couldn't occur. In fact  $P(s) = P(s') \Rightarrow P^*(s) = P^*(s')$  (see above) and thus  $Z(s) = Z(s')$ , which contradicts the assumption about  $Z$ . Consider now the random variable  $\tilde{Y}$  defined by

$$\begin{aligned} \tilde{Y}(s) &= Y(s'), \\ \tilde{Y}(s') &= Y(s') + \frac{P(s)}{P(s')} (Y(s) - Y(s')), \\ \tilde{Y}(t) &= Y(t), \forall t \in \Omega \setminus \{s, s'\}. \end{aligned}$$



Then we obtain

$$\mathbb{E}_{\mathbb{P}} \left[ Z(\tilde{Y} - Y) \right] < 0.$$

Moreover for each increasing, concave function  $u : \mathcal{G} \rightarrow \mathbb{R}$  we have

$$\begin{aligned} \mathbb{E}_{\mathbb{P}} \left[ u(\tilde{Y}) - u(Y) \right] &= \\ &= P(s) \left( u(\tilde{Y}(s)) - u(Y(s)) \right) + P(s') \left( u(\tilde{Y}(s')) - u(Y(s')) \right) \\ &= P(s) \left( u(Y(s')) - u(Y(s)) \right) + P(s') \left( u\left( Y(s') + \frac{P(s)}{P(s')} (Y(s) - Y(s')) \right) - u(Y(s')) \right) \\ &\geq P(s) \left( u(Y(s')) - u(Y(s)) \right) + \\ &\quad + P(s') \left( \frac{P(s') - P(s)}{P(s')} u(Y(s')) + \frac{P(s)}{P(s')} u(Y(s)) - P(s') u(Y(s')) \right) \\ &= 0. \end{aligned}$$

The inequality follows from the concavity of  $u$ . Thus  $\tilde{Y} \succ_2 X$  and  $\mathbb{E}_{\mathbb{P}} \left[ Z(\tilde{Y} - Y) \right] < 0$ , a contradiction to  $Y \in \Psi(X, Z)$ .

The anticomonicity of  $Y$  and  $Z$  it is a necessary condition for  $Y$  to be in  $\Psi(X, Z)$  but also a sufficient condition. For  $\mu$  to be consistent with  $\succ_2$  on  $\mathcal{G}$ , for all  $X \in \mathcal{G}$  we should have  $X \in \Psi(X, Z)$ . This implies that  $X$  and  $Z$  should be anticomotone for all  $X \in \mathcal{G}$ . This is possible if and only if  $Z(s) = z \forall s \in \Omega$ . In fact if for  $s, s'$  we have  $Z(s) > Z(s')$  and  $\mu$  satisfies the isotonicity, then for all  $X \in \mathcal{G}$  we must have  $X(s) \leq X(s')$ . But this naturally false. Since  $Z$  is a density, we have  $z = 1$ . Thus  $\mathbb{P}^* = \mathbb{P}$ .  $\square$

**Remark** (i) The Theorem states that if one defines second order stochastic dominance on  $(\Omega, \mathcal{F})$  with respect to a probability measure  $\mathbb{P}$ , then the unique reward measure which is isotonic with  $\succ_2$  on  $\mathcal{G}$  is the expectation under  $\mathbb{P}$ .

(ii) We have defined the reward measure on the space  $\mathcal{G}$ . If one restrict the space of future value changes, then it might be possible that  $\mathbb{P}^* \neq \mathbb{P}$ . Moreover, in this case, the  $\mathbb{P}^*$  can have negative components.

## 4.2 Risk measures

Let  $\rho$  be a risk measure. We ask the question about a characterization of  $\rho$ . In this work we consider only positive homogeneous risk measures; an extension to other risk measures will be considered in an other work. If we additionally assume that a risk measure  $\rho$  is positive homogeneous, then it follows that  $\rho$  is a subadditive functional on  $\mathcal{G}$ , which is also isotonic with respect to second order stochastic dominance and satisfies the risk free condition, i.e.  $\rho$  is invariant under addition of a constant.

Given a positively homogeneous risk measure  $\rho : \mathcal{G} \rightarrow \mathbb{R}$  and reward measure  $\mu : \mathcal{G} \rightarrow \mathbb{R}$ , i.e.  $\mu(X) = \mathbb{E}_{\mathbb{P}}[X]$  for all  $X \in \mathcal{G}$ , the measure  $\mathcal{R}_{\xi} = \xi\rho - \mu$  for  $\xi > 0$  is a coherent measure, which additionally satisfies the isotonic property. By Artzner, Delbaen, Eber, and Heath (1999) we know that there exists a scenario  $\mathcal{P}$ , i.e. a nonempty family of probability measures on  $\Omega$ , such that

$$\mathcal{R}_{\xi}(X) = \mathcal{R}_{\mathcal{P}}(X) = \sup \{ \mathbb{E}_{\mathbb{Q}}[-X] \mid \mathbb{Q} \in \mathcal{P} \} = - \inf \{ \mathbb{E}_{\mathbb{Q}}[X] \mid \mathbb{Q} \in \mathcal{P} \}. \quad (31)$$

The question is, what conditions should be satisfied by  $\mathcal{P}$  in order that  $\mathcal{R}_{\mathcal{P}}$  is isotonic with respect to second order stochastic dominance.

First some considerations. Suppose that  $P(s) > P(s')$ , then  $X^s \succ_2 X^{s'}$  and thus  $\mathcal{R}_{\mathcal{P}}(X^s) \leq \mathcal{R}_{\mathcal{P}}(X^{s'})$ . This implies that  $\sup\{-Q(s) | \mathbb{Q} \in \mathcal{P}\} \leq \sup\{-Q(s') | \mathbb{Q} \in \mathcal{P}\}$  or,

$$\inf\{Q(s) | \mathbb{Q} \in \mathcal{P}\} \geq \inf\{Q(s') | \mathbb{Q} \in \mathcal{P}\}.$$

Therefore, if  $P(s) > P(s')$  for some  $s, s' \in \Omega$ , then it should exist at least one measure  $\mathbb{Q}^* \in \mathcal{P}$  such that  $Q(s) \geq Q^*(s')$  for all  $\mathbb{Q} \in \mathcal{P}$ , and therefore also for  $\mathbb{Q}^*$ .

Suppose now that  $P(s) = P(s')$ , then  $X^s \succ_2 \epsilon X^{s'}$  for  $\epsilon > 1$ . Thus

$$\inf\{Q(s) | \mathbb{Q} \in \mathcal{P}\} \geq \epsilon \inf\{Q(s') | \mathbb{Q} \in \mathcal{P}\}.$$

It follows that for  $s, s' \in \Omega$  such that  $P(s) = P(s')$ , we find  $\mathbb{Q}^* \in \mathcal{P}$  such that  $Q(s) \geq \epsilon Q^*(s') \geq Q^*(s')$  for all  $\mathbb{Q} \in \mathcal{P}$ , and therefore also for  $\mathbb{Q}^*$ . By symmetry, we find  $\mathbb{Q}^{**} \in \mathcal{P}$  such that  $Q(s') \geq Q^{**}(s)$  for all  $\mathbb{Q} \in \mathcal{P}$ , and therefore also for  $\mathbb{Q}^*$  and  $\mathbb{Q}^{**}$ . We have  $Q^{**}(s) = Q^{**}(s') = Q^*(s) = Q^*(s')$ .

**Lemma 4.1.**

$$\begin{aligned} \exists s, s' \in \Omega \text{ s.t. } P(s) = P(s') &\Rightarrow \exists \mathbb{Q}^* \in \mathcal{P} \text{ s.t. } Q^*(s) = Q^*(s'), \\ \exists s, s' \in \Omega \text{ s.t. } P(s) > P(s') &\Rightarrow \exists \mathbb{Q}^* \in \mathcal{P} \text{ s.t. } Q^*(s) \geq Q^*(s'). \end{aligned}$$

The Lemma suggests that  $\mathcal{P}$  should contain some probability measure, which is a non-decreasing function of  $\mathbb{P}$ . Let us consider the following class of measures on  $(\Omega, \mathcal{F})$

$$\mathcal{D} = \{\nu | \nu(A) = g(\mathbb{P}[A]), \forall A \in \mathcal{F}, g \text{ non-decreasing, } g(0) = 0, g(1) = 1\}. \quad (32)$$

$\nu \in \mathcal{D}$  is called a distortion of  $\mathbb{P}$  or capacity. Note that  $\nu$  is positive,  $\nu(\Omega) = 1$ , but  $\nu$  might not be a probability measure on  $(\Omega, \mathcal{F})$ . For example, for some  $A \in \mathcal{F}$ , following could be true

$$\nu(A^c) = g(\mathbb{P}[A^c]) = g(1 - \mathbb{P}[A]) \neq 1 - g(\mathbb{P}[A]) = 1 - \nu(A).$$

For  $\nu \in \mathcal{D}$  we define the functional  $\mathcal{R}_{\nu}$  on  $\mathcal{G}$  as follows

$$\mathcal{R}_{\nu}(X) = - \int_{-\infty}^0 (\nu(X > x) - 1) dx - \int_0^{\infty} \nu(X > x) dx \quad (33)$$

$$= - \int_{-\infty}^0 (g(\overline{F}_X(x)) - 1) dx - \int_0^{\infty} g(\overline{F}_X(x)) dx, \quad (34)$$

where  $\overline{F}_X(x) = 1 - F_X(x)$  and  $F_X(x) = \mathbb{P}[X \leq x]$  and  $g$  is the non-decreasing function on  $[0, 1]$  associated to  $\nu$ .

**Remark** (i)  $-\mathcal{R}_{\nu}(X)$  is also called the Choquet integral of  $X$  with respect to  $\nu$ .

(ii) The expectation of  $X$  under  $\mathbb{P}$  can be written as

$$\mathbb{E}_{\mathbb{P}}[X] = \int_{-\infty}^{\infty} x dF_X(x) = \int_{-\infty}^0 (\overline{F}_X(x) - 1) dx + \int_0^{\infty} \overline{F}_X(x) dx.$$

If  $g$  is differentiable, the functional  $-\mathcal{R}_{\nu}$  corresponds to

$$\int_{-\infty}^{\infty} x g'(\overline{F}_X(x)) dF_X(x).$$

Therefore,  $-\mathcal{R}_\nu$  can be viewed as a "corrected" mean of  $X$ , where the realization  $x$  receives the weight  $g'(\bar{F}_X(x))$ . Note that

$$\int_{-\infty}^{\infty} g'(\bar{F}_X(x)) dF_X(x) = \int_{-\infty}^{\infty} \frac{d}{dx}(-g(\bar{F}_X(x))) dx = g(1) - g(0) = 1.$$

Suppose that  $g$  is strict convex, differentiable. For  $x > y$ ,  $\bar{F}_X(x) < \bar{F}_Y(y)$ , then

$$g'(\bar{F}_X(x)) < g'(\bar{F}_X(y)).$$

Therefore, the weight assigned to a high outcome  $x$  is less than the weight assigned to a low outcome, meaning that under  $g$  the decision maker assign a higher probability to low outcomes.

- (iii) *Choquet Expected Utility Theory.* The previous remark and equation (33) suggest also a link with the Choquet Expected Utility Theory (CEU) of Schmeidler (1989). For a brief introduction to this theory see also Eisenführ and Weber (1999, Chapter 14) and Karni and Schmeidler (1991, Section 6). In this latter reference, the authors defines the *comonotonic independence axiom*, by restricting the independence axiom from the von Neuman-Morgenstern Utility Theory to comonotonic outcomes<sup>2</sup>. They also show (Theorem 6.4.2) that comonotonic independence together with monotonicity, continuity (Archimedian axiom) and non-degeneracy, are equivalent to the existence of a unique (up to linear transformation) representation as integral with respect to a capacity, which is the Choquet expected utility. The motivation for the CEU comes from the observation by the Ellsberg experiment<sup>3</sup>, that when decision makers are asked to ranks bets, their behaviour is inconsistent with the independence axiom of classical Expected Utility Theory. Therefore, Karni and Schmeidler (1991) proposed a weakening form of this axiom.

**Lemma 4.2.** *Let  $\nu \in \mathcal{D}$  and  $g$  the associated function.  $\mathcal{R}_\nu$  satisfies the following properties:*

- (i)  $\mathcal{R}_\nu(\alpha X) = \alpha \mathcal{R}_\nu(X)$  for  $X \in \mathcal{G}$  and  $\alpha > 0$ ;
- (ii)  $\mathcal{R}_\nu(0) = 0$ ;
- (iii)  $\mathcal{R}_\nu(X(\mathbf{e}_0)) = -R_0$ ;
- (iv) If  $g$  is convex then  $\mathcal{R}_\nu$  is convex.

*Proof.* (i) Note that for  $\alpha > 0$ ,  $F_{\alpha X}(x) = F_X(\frac{x}{\alpha})$ . The assertion follows directly.

(ii) Obvious.

(iii) Obvious.

(iv) If  $g$  is convex, then we have the following representation due to Schmeidler (1986)

$$\mathcal{R}_\nu(X) = -\min\{\mathbb{E}_Q[X] \mid Q \in \text{core}(\nu)\}, \quad (35)$$

where  $\text{core}(\nu) = \{Q \mid \text{additive on } \Omega \text{ s.t. } Q(\Omega) = \nu(\Omega), \forall A \in \mathcal{F} : Q(A) \geq \nu(A)\}$ . Note that if  $g$  is convex,  $\nu$  is convex in the sense of Schmeidler (1986), i.e. for all  $E, F \in \mathcal{F}$ , we have

<sup>2</sup> $X, Y \in \mathcal{G}$  are said to be comonotonic, if  $X(s) > X(s') \Rightarrow Y(s) > Y(s')$  for all  $s, s' \in \Omega$ .

<sup>3</sup>The experiment presents to decision makers an urn with 90 balls, 30 red and 60 black or white. Decision makers are asked to guess the color of one ball drawn at random. When the guess is correct, the get \$100, else \$0.

$\nu(E \cap F) + \nu(E \cup F) \geq \nu(E) + \nu(F)$ . The opposite is also true.

Let  $X, Y \in \mathcal{G}$  and  $\mathbb{Q} \in \text{core}(\nu)$  such that  $\mathcal{R}_\nu(X + Y) = -\mathbb{E}_\mathbb{Q}[X + Y]$  then by (35)

$$\begin{aligned}\mathcal{R}_\nu(X) &\geq -\mathbb{E}_\mathbb{Q}[X], \\ \mathcal{R}_\nu(Y) &\geq -\mathbb{E}_\mathbb{Q}[Y].\end{aligned}$$

It follows

$$\mathcal{R}_\nu(X + Y) = -\mathbb{E}_\mathbb{Q}[X + Y] = -\mathbb{E}_\mathbb{Q}[X] - \mathbb{E}_\mathbb{Q}[Y] \leq \mathcal{R}_\nu(X) + \mathcal{R}_\nu(Y).$$

The subadditivity and the positive homogeneity imply the convexity of  $\mathcal{R}_\nu$ .  $\square$

The Lemma shows that given a capacity  $\nu \in \mathcal{D}$  for  $\mathbb{P}$ , which is obtained by a non-decreasing, convex function  $g$ , the measure  $\mathcal{R}_\nu$  defines a coherent risk measure.

**Lemma 4.3.** *Let  $\mathcal{R}_\nu$  be the measure associated to the capacity  $\nu \in \mathcal{D}$ . Suppose that  $\mathcal{R}_\nu$  is convex, then  $g$  is convex.*

*Proof.* Schmeidler (1986, Proposition 3)  $\square$

The next step consists in identifying the condition such that  $\mathcal{R}_\nu$  is isotonic with respect to second order stochastic dominance. The following Lemma answer this question.

**Lemma 4.4.** *If  $g$  is increasing and convex, then  $\mathcal{R}_\nu$  is isotonic with respect to the second order stochastic dominance.*

*Proof.* The following proof is based on the argument used by Wang (1996).

Let  $X, Y \in \mathcal{G}$  such that  $X \succ_2 Y$ .

- (i) Suppose that  $F_X(x) \leq F_Y(x) \forall x$ , and the inequality is strict for at least one  $x_0$ . Then  $\bar{F}_X(x) \geq \bar{F}_Y(x) \forall x$ , and the inequality is strict for at least one  $x_0$ . Since  $g$  is non-decreasing and equation (33), it follows

$$\begin{aligned}\mathcal{R}_\nu(X) &= -\int_{-\infty}^0 (g(\bar{F}_X(x)) - 1) dx - \int_0^\infty g(\bar{F}_X(x)) dx \\ &\leq -\int_{-\infty}^0 (g(\bar{F}_Y(x)) - 1) dx - \int_0^\infty g(\bar{F}_Y(x)) dx \\ &= \mathcal{R}_\nu(Y).\end{aligned}$$

- (ii) Suppose that  $F_X$  and  $F_Y$  cross once. Then  $\bar{F}_X$  and  $\bar{F}_Y$  also cross once. Thus we find  $x_0$  such that

$$\begin{aligned}\bar{F}_X(x) &\geq \bar{F}_Y(x), \quad \text{for } x < x_0, \\ \bar{F}_X(x) &\leq \bar{F}_Y(x), \quad \text{for } x \geq x_0.\end{aligned}$$

Moreover since  $X \succ_2 Y$  at least one the following inequalities should be satisfied

$$\begin{aligned}\bar{F}_X(x) &> \bar{F}_Y(x), \quad \text{for some } x < x_0, \\ \bar{F}_X(x) &< \bar{F}_Y(x), \quad \text{for some } x > x_0.\end{aligned}$$

Let  $\bar{F}_Z = \max\{\bar{F}_X, \bar{F}_Y\}$  be the survival distribution function of a random variable  $Z$ . Suppose first that  $x_0 > 0$ , then

$$\begin{aligned} \mathcal{R}_\nu(Z) - \mathcal{R}_\nu(X) &= - \int_{x_0}^{\infty} [g(\bar{F}_Y(x)) - g(\bar{F}_X(x))] dx \\ &\geq - \frac{1 - g(\bar{F}_X(x_0))}{1 - \bar{F}_X(x_0)} \int_{x_0}^{\infty} [\bar{F}_Y(x) - \bar{F}_X(x)] dx. \end{aligned}$$

We use that  $m_y(x) = \frac{g(x) - g(y)}{x - y}$  is non-decreasing in  $x$  for  $y$  fix and non-decreasing in  $y$  for  $x$  fix, where  $x, y \in (0, 1)$  since  $g$  is convex (see also Remark below). Analogously

$$\begin{aligned} \mathcal{R}_\nu(Z) - \mathcal{R}_\nu(Y) &= - \int_{-\infty}^{x_0} [g(\bar{F}_X(x)) - g(\bar{F}_Y(x))] dx \\ &\leq - \frac{g(\bar{F}_X(x_0))}{\bar{F}_X(x_0)} \int_{-\infty}^{x_0} [\bar{F}_X(x) - \bar{F}_Y(x)] dx \\ &\leq - \frac{1 - g(\bar{F}_X(x_0))}{1 - \bar{F}_X(x_0)} \int_{-\infty}^{x_0} [\bar{F}_X(x) - \bar{F}_Y(x)] dx. \end{aligned}$$

Subtracting these inequalities, we obtain

$$\begin{aligned} \mathcal{R}_\nu(Y) - \mathcal{R}_\nu(X) &= \\ &= [\mathcal{R}_\nu(Z) - \mathcal{R}_\nu(X)] - [\mathcal{R}_\nu(Z) - \mathcal{R}_\nu(Y)] \\ &\geq - \frac{1 - g(\bar{F}_X(x_0))}{1 - \bar{F}_X(x_0)} \int_{x_0}^{\infty} [\bar{F}_Y(x) - \bar{F}_X(x)] dx + \frac{1 - g(\bar{F}_X(x_0))}{1 - \bar{F}_X(x_0)} \int_{-\infty}^{x_0} [\bar{F}_X(x) - \bar{F}_Y(x)] dx \\ &= \frac{1 - g(\bar{F}_X(x_0))}{1 - \bar{F}_X(x_0)} \int_{-\infty}^{\infty} [\bar{F}_X(x) - \bar{F}_Y(x)] dx \\ &= \frac{1 - g(\bar{F}_X(x_0))}{1 - \bar{F}_X(x_0)} \left\{ \int_{-\infty}^0 [(\bar{F}_X(x) - 1) - (\bar{F}_Y(x) - 1)] dx + \int_0^{\infty} [(\bar{F}_X(x) - \bar{F}_Y(x))] dx \right\} \\ &\geq \mathbb{E}_{\mathbb{P}}[X] - \mathbb{E}_{\mathbb{P}}[Y] \geq 0. \end{aligned}$$

The last but one inequality uses that  $g(x) \leq x$  for  $x \in (0, 1)$  (see Remark below). For  $x_0 < 0$  the proof is analog. Thus, for  $X \succ_2 Y$  we have  $\mathcal{R}_\nu(X) \leq \mathcal{R}_\nu(Y)$ , if the corresponding cumulative distribution functions cross only once.

- (iii) The general case. If  $X \succ_2 Y$ , then by Müller (1985, Corollary 4.4), there exists a sequence of distribution functions  $(G_i)_{i \geq 1}$ , such that  $G_1 = F_Y$ ,  $G_i$  and  $G_{i+1}$  cross only once,  $G_{i+1}$  dominates  $G_i$  by second order stochastic dominance, and  $G_i \rightarrow F_X$  weakly. It follows that

$$\begin{aligned} \mathcal{R}_\nu(Y) - \mathcal{R}_\nu(X) &= \mathcal{R}_\nu(G_1) - \mathcal{R}_\nu(G_\infty) \\ &= \sum_{i=1}^{\infty} \mathcal{R}_\nu(G_i) - \mathcal{R}_\nu(G_{i+1}) \\ &\geq \sum_{i=1}^{\infty} \mathbb{E}_{\mathbb{P}}[G_i] - \mathbb{E}_{\mathbb{P}}[G_{i+1}] = \mathbb{E}_{\mathbb{P}}[X] - \mathbb{E}_{\mathbb{P}}[Y]. \end{aligned}$$

□

**Corollary 4.1.** *Let  $\nu \in \mathcal{D}$  and  $g$  the associated function. Let  $X, Y \in \mathcal{G}$ . If  $g$  is non-decreasing and convex, then*

$$X \succ_2 Y \Rightarrow \mathcal{R}_\nu(Y) + \mathbb{E}_\mathbb{P}[Y] \geq \mathcal{R}_\nu(X) + \mathbb{E}_\mathbb{P}[X]. \quad (36)$$

**Remark** (i) Let  $\mathbb{Q} \in \text{core}(\nu)$ . Then  $\mathbb{Q}$  is absolutely continuous with respect to  $\mathbb{P}$ . In fact for  $A \in \mathcal{F}$  such that  $\mathbb{P}[A] = 0$  we have  $\nu(A^c) = g(\mathbb{P}[A^c]) = g(1) = 1$  and thus for  $\mathbb{Q} \in \text{core}(\nu)$ ,  $\mathbb{Q}(A^c) \geq \nu(A^c) = 1$ , i.e.  $\mathbb{Q}(A^c) = 1$ .

It follows that for  $\mathbb{Q} \in \text{core}(\nu)$ , there exists a random variable  $Z \geq 0$  on  $(\Omega, \mathcal{F})$  such that  $\mathbb{E}_\mathbb{P}[Z] = 1$  and  $\mathbb{E}_\mathbb{Q}[X] = \mathbb{E}_\mathbb{P}[ZX]$  ( $Z$  is the density of  $\mathbb{Q}$  with respect to  $\mathbb{P}$ ). When we write  $Z \in \text{core}(\nu)$  we mean that  $Z$  is the density of  $\mathbb{Q} \in \text{core}(\nu)$  with respect to  $\mathbb{P}$ .

(ii) If  $g$  is convex and  $g(0) = 0$ ,  $g(1) = 1$ , then  $g(x) \leq x \forall x \in [0, 1]$ .

*Proof.* Since  $g$  is convex, the function  $m_y(x) = \frac{g(x)-g(y)}{x-y}$  is non-decreasing in  $x$  for  $y$  fix. Let  $y = 0$ , we obtain  $m_0(x) = \frac{g(x)}{x}$  since  $g(0) = 0$  and therefore  $m_0(1) \geq m_0(x) \forall x \in (0, 1]$ . Since  $m_0(1) = 1$  we obtain  $g(x) \leq x \forall x \in (0, 1]$ .  $\square$

It follows that  $\nu(A) = g(\mathbb{P}[A]) \leq \mathbb{P}[A]$  and thus  $\mathbb{P} \in \text{core}(\nu)$ . Therefore

$$\mathcal{R}_\nu(X) = -\min\{\mathbb{E}_\mathbb{Q}[X] \mid \mathbb{Q} \in \text{core}(\nu)\} \geq -\mathbb{E}_\mathbb{P}[X],$$

i.e.

$$\mathcal{R}_\nu(X) + \mathbb{E}_\mathbb{P}[X] \geq 0.$$

**Proposition 4.1.** *Let  $\nu \in \mathcal{D}$  be a capacity with respect to  $\mathbb{P}$  and  $g$  the associated function. If  $g$  is non-decreasing and convex, then  $\mathcal{R}_\nu$  defined by equation (33) is a coherent measure and moreover, it satisfies the isotonicity with respect to second order stochastic dominance and*

$$\mathcal{R}_\nu(X) = -\min\{\mathbb{E}_\mathbb{Q}[X] \mid \mathbb{Q} \in \text{core}(\nu)\}, \quad (37)$$

where  $\text{core}(\nu) = \{\mathbb{Q} \mid \text{additive on } \Omega \text{ s.t. } \mathbb{Q}(\Omega) = \nu(\Omega), \forall A \in \mathcal{F} : \mathbb{Q}(A) \geq \nu(A)\}$ .

**Theorem 4.2.** *Let  $\nu \in \mathcal{D}$  and  $g$  the associated function. If  $g$  is non-decreasing and convex then*

$$\rho_\nu(X) = -\min\{\mathbb{E}_\mathbb{Q}[X - \mathbb{E}_\mathbb{P}[X]] \mid \mathbb{Q} \in \text{core}(\nu)\}, \quad (38)$$

is a risk measure.

*Proof.* Proposition 4.1 and Corollary 4.1.  $\square$

### 4.3 Examples

From Theorem 4.2 one can define a risk measure  $\rho$  on  $\mathcal{G}$  with physical probability measure  $\mathbb{P}$ , by choosing a non-decreasing and convex function  $g$ , with  $g(0) = 0$ ,  $g(1) = 1$  and the corresponding Coquet integral given by equation (33).

#### Example (Expected Value)

Let  $g(x) = x$ .  $g$  is obviously non-decreasing and convex and  $g(0) = 0$ ,  $g(1) = 1$ . Let  $\nu$  be the corresponding capacity. We have  $\nu \equiv \mathbb{P}$  and  $\mathcal{R}_\nu(X) = -\mathbb{E}_\mathbb{P}[X]$ , i.e.  $\rho_\nu \equiv 0$ . In this case the decision maker is risk-neutral and care only about the reward.

**Example (Conditional Value-at-Risk)**

Let  $\beta \in (0, 1)$ . We define the function  $g_\beta$  by

$$g_\beta(x) = \begin{cases} 0 & \text{if } 0 < x \leq 1 - \beta, \\ -\frac{1-\beta}{\beta} + \frac{x}{\beta} & \text{if } 1 - \beta < x \leq 1. \end{cases}$$

$g_\beta$  is non-decreasing and convex (see Figure 1).  $\nu_\beta$  denotes the corresponding capacity on  $(\Omega, \mathcal{F})$ .

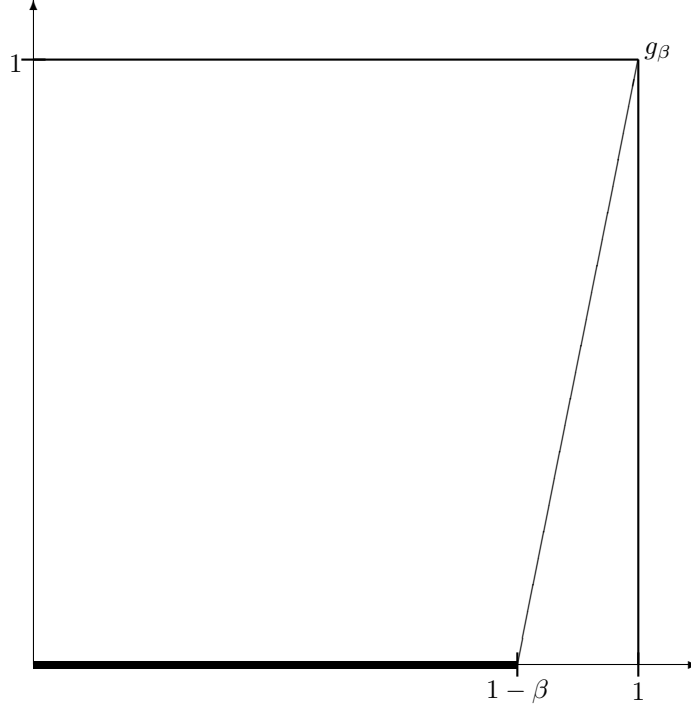


Figure 1: Convex function, which generates the Conditional Value-at-Risk with level  $\beta$ .

Let  $x^\beta = F_X^{-1}(\beta) = \inf\{x \in \mathbb{R} \mid F_X(x) \geq \beta\} = \inf\{x \in \mathbb{R} \mid \bar{F}_X(x) \leq 1 - \beta\}$ . Suppose first that  $x^\beta < 0$ , then we have

$$\begin{aligned} \mathcal{R}_{\nu_\beta}(X) &= - \int_{-\infty}^0 (g_\beta(\bar{F}_X(x)) - 1) dx - \int_0^\infty g_\beta(\bar{F}_X(x)) dx \\ &= - \int_{-\infty}^{x^\beta} \left( -\frac{1-\beta}{\beta} + \frac{\bar{F}_X(x)}{\beta} - 1 \right) dx \\ &= \int_{-\infty}^{x^\beta} \left( \frac{1 - \bar{F}_X(x)}{\beta} \right) dx = \frac{1}{\beta} \int_{-\infty}^{x^\beta} F_X(x) dx. \end{aligned}$$

For  $x^\beta > 0$ , we obtain

$$\mathcal{R}_{\nu_\beta}(X) = \frac{1}{\beta} \int_{-\infty}^{x^\beta} F_X(x) dx - x^\beta.$$

Therefore,  $\mathcal{R}_{\nu_\beta}$ , gives the Conditional Value-at-Risk for  $\beta \in (0, 1)$  (see Vanini and Vignola 2001, Rockafellar and Uryasev 1999). The measure

$$\rho_{\nu_\beta}(X) = \mathcal{R}_{\nu_\beta}(X) + \mathbb{E}_{\mathbb{P}}[X]$$

is a risk measure. Note that  $\rho_{\nu_\beta}$  is the usual Conditional-Value-at-Risk shifted to the left (or to the right) by the expectation over  $\mathbb{P}$ . This measure has also been studied by Hürlimann (2002).

**Example (Bickel-Lehman Order)**

Let  $\mathbb{Q}$  be absolutely continuous with respect to  $\mathbb{P}$  and  $Z = \frac{d\mathbb{Q}}{d\mathbb{P}}$  be the density of  $\mathbb{Q}$  with respect to  $\mathbb{P}$ . Let  $F_Z$  be the cumulative distribution function of  $Z$  and  $F_Z^{-1}(t) = \inf\{z | F_Z(z) \geq t\}$  its generalized inverse. Then

$$g_Z(x) = \frac{\int_0^x F_Z^{-1}(t) dt}{\int_0^1 F_Z^{-1}(t) dt}$$

is a non-decreasing, convex function on  $[0, 1]$  with  $g_Z(0) = 0$ ,  $g_Z(1) = 1$ . Thus the corresponding capacity  $\nu_Z = g_Z \circ \mathbb{P}$  defines a measure  $\mathcal{R}_{\nu_Z}$  through equation (33), which is coherent and isotonic with respect to second order stochastic dominance. The corresponding risk measure is given by  $\mathcal{R}_{\nu_Z} + \mathbb{E}_{\mathbb{P}}[\cdot]$ . When the market is complete, then one could define unique state prices  $Z(s)$  for  $s = 1, \dots, S$ . These prices define a density  $Z$  and for  $k = 0, \dots, K$ ,  $q^k = \sum_{s=1}^S Z(s)R^k(s)$ . As shown by Dana (2002, Proposition 5), the risk measure  $\mathcal{R}_{\nu_Z} + \mathbb{E}_{\mathbb{P}}[\cdot]$  is consistent with the Bickel-Lehman order<sup>4</sup>.

**Example (Spectral Risk Measure)**

Acerbi (2002) and Acerbi and Simonetti (2002) define a spectral measures as any measure of the form

$$\mathcal{R}_\phi(X) = - \int \phi(t) F_X^{-1}(t) dt$$

for a real function  $\phi$  on  $[0, 1]$ . Moreover, they say that  $\phi$  is an acceptable (risk) spectrum if  $\phi$  is “positive”, “decreasing” and  $\int_0^1 |\phi(t)| dt = 1$ . Note that the properties of monotonicity and positivity are defined by Acerbi (2002, Definition 2.3) on the space of integrable functions on  $[0, 1]$ ,  $\mathcal{L}^1([0, 1])$  and not pointwise. Positive and decreasing functions in the “pointwise”-sense are also positive and decreasing in the sense of Acerbi (2002).

Let  $g$  be an increasing, strict convex function with  $g(0) = 0$  and  $g(1) = 1$  such that the first and the second derivative exist. Let  $\phi(t) = g'(1-t)$ . Then obviously  $\phi$  is positive (since  $g$  is increasing), decreasing (since  $g$  is strict convex) and  $\int_0^1 |\phi(t)| dt = \int_0^1 g'(1-t) dt = -[g(0) - g(1)] = 1$ . Therefore,  $\phi = g'$  is an acceptable spectrum in the sense of Acerbi (2002). Let  $\nu$  be the capacity corresponding to  $g$  and  $\mathcal{R}_\nu$  its Choquet integral (equation (33)). Then we have

$$\mathcal{R}_\nu(X) = - \int_0^1 \phi(t) F_X^{-1}(t) dt = \mathcal{R}_\phi(X),$$

i.e.  $\mathcal{R}_\nu$  is a spectral measure with acceptable spectrum. The corresponding risk measure is  $\rho_\phi = \mathcal{R}_\phi + \mathbb{E}_{\mathbb{P}}[\cdot]$ . Note that

$$\rho_\phi(X) = - \int_0^1 \phi(t) F_X^{-1}(t) dt + \int_0^1 F_X^{-1}(t) dt = - \int_0^1 (\phi(t) - 1) F_X^{-1}(t) dt,$$

which is still a spectral measure, but the corresponding spectrum is not acceptable in the sense of Acerbi (2002) (see Acerbi and Simonetti 2002, Proposition 4.1).

<sup>4</sup>Let  $X, Y \in \mathcal{G}$ . We say that  $X$  dominates  $Y$  by for Bickel-Lehman order denoted  $X \succ_{\text{BL}} Y$  iff the map  $t \mapsto F_X^{-1}(t) - F_Y^{-1}(t)$  is non-increasing on  $(0, 1)$ .



## 5 Conclusion

In this paper we have considered the reward-risk approach for portfolio selection and we have defined reward and risk measures through a set of axioms, following the approach proposed by Artzner, Delbaen, Eber, and Heath (1999). To maintain a link between our reward-risk framework and expected utility decision theory we have imposed the isotonicity axioms for both the reward and the risk measure. The consequences of this axiom is that a reward-risk decision maker is consistent with the expected utility maximizer, at least for all the situation where the rational, risk averse expected utility decision does not depend on the shape of the utility index, i.e. when one risk dominates another risk by second order stochastic dominance.

For a finite state of the world, we have shown that on the space of random variables there exists only one reward measure, which is given by expectation under the physical probability. For risk measures we obtain a nice characterization through Choquet integral, which can be interpreted as a “weighted mean”, where the weights are non-increasing in the outcomes.

We do not see any reason for suggesting one particular risk measure in the class of measure we have characterized. The choice of the “risk aversion” function  $g$  should depend on the risk characteristics of the investor.

## References

- ACERBI, C. (2002): “Spectral measure of risk: a coherent representation of subjective risk aversion,” Working paper, Abaxbank, Milano.
- ACERBI, C., AND P. SIMONETTI (2002): “Portfolio optimization with spectral measure of risk,” Working paper, Abaxbank, Milano.
- ACERBI, C., AND D. TASCHE (2001): “Expected Shortfall: a natural coherent alternative to Value at Risk,” Working Paper, Abaxbank, Milano and Zentrum Mathematik, TU München, München.
- ARTZNER, P., F. DELBAEN, J.-M. EBER, AND D. HEATH (1997): “Thinking coherently,” *Risk*, 10(11), 68–71.
- (1999): “Coherent measure of risk,” *Mathematical Finance*, 9(3), 203–228.
- CARLIER, G., AND R. DANA (2001): “Core of convex distortions of a probability on a non atomic space,” Working Paper, Ceremade.
- DANA, R. (2002): “Stochastic dominance, risk, dispersion and equilibrium asset pricing,” Working Paper, Ceremade.
- DE GIORGI, E. (2002): “A note on portfolio selection under various risk measures,” Working Paper 122, Insitute for Empirical Research in Economics, University of Zurich.
- DUFFIE, D., AND J. PAN (1997): “An overview of Value-at-Risk,” *Journal of Derivatives*, 4, 7–49.
- EICHBERGER, J., AND I. HARPER (1997): *Financial Economics*. Oxford University Press.
- EISENFÜHR, F., AND M. WEBER (1999): *Rationales entscheiden*. Springer, Berlin-New York.
- HÜRLIMANN, W. (1999): “Change of measure and stop-loss order,” Unpublished note.
- (2002): “An alternative approach to portfolio selection,” in *Proceedings of the 12th international AFIR Colloquium*, Cancun, Mexico.

- JASCHKE, S., AND U. KÜCHLER (2000): "Coherent risk measures, valuation bounds, and  $(\mu, \sigma)$ -portfolio optimiation," Working Paper, Weierstrass-Institut für Angewandte Analysis und Stochastic and Humboldt-Universität zu Berlin.
- JORION, P. (1997): *Value at Risk*. Irwin, Burr Ridge, IL.
- KARNI, E., AND D. SCHMEIDLER (1991): "Utility theory with uncertainty," in *Handbook of Mathematical Economics*, ed. by W. Hilderbrand, and H. Sonnenschein, vol. IV, chap. 33. Elsevier, Amsterdam.
- LEVY, H. (1998): *Stochastic dominance*. Kluwer Academic Publisher, Boston-Dordrecht-London.
- MARKOWITZ, H. (1952): "Portfolio Selection," *Journal of Finance*, 7, 77–91.
- (1959): *Portfolio Selection*. John Wiley & Sons, New York.
- MÜLLER, M. (1985): "Ordering of risks: a comparative study via stop-loss transforms," *Insurance: Mathematics and Economics*, 17(3), 215–222.
- MGRYZAK, W., AND A. RUSZCZYNSKI (1997): "From stochastic dominance to mean-risk models: semideviations as risk measures," *IIASA Interim Report, IR-97-027*.
- PFLUG, C. (1998): "How to measure risk?," Working Paper, IIASA and University of Vienna.
- ROCKAFELLAR, R., AND S. URYASEV (1999): "Optimization of Conditional Value-at-Risk," Working Paper, University of Washington and University of Florida.
- ROTHSCHILD, M., AND J. STIGLITZ (1970): "Increasing risk I: a definition," *Journal of Economic Theory*, 2(3), 225–243.
- (1971): "Increasing risk II: its economic consequences," *Journal of Economic Theory*, 3(1), 66–84.
- SCHMEIDLER, D. (1986): "Integral representation without additivity," *Proceeding of the American Mathematical Society*, 97(2), 255–261.
- (1989): "Subjective probability and expected utility without additivity," *Econometrica*, 57, 571–587.
- SHAKED, M., AND J. SHANTHIKUMAR (1994): *Stochastic orders and their applications*. Academic Press, San Diego.
- VANINI, P., AND L. VIGNOLA (2001): "Portfolio Selection," Lecture Notes ETH Zurich.
- WANG, S. (1996): "Premium calculation by transforming the layer premium density," *ASTIN Bulletin*, 26, 71–92.
- YU, P.-L. (1985): *Multiple-criteria decision making*. Plenum Press, New York-London.