AN INTRODUCTION TO

PORTFOLIO MANAGEMENT

J. Gáll [1]

Radboud University (RU), Nijmegen, The Netherlands,
Faculty of Science,
Institute for Mathematics, Astrophysics and Particle Physics (IMAPP)

University of Debrecen, Hungary,
Faculty of Economics and Business Administration
Department for Economic Analysis and Business Informatics

E-mail: j.gall@math.ru.nl, jozsef.gall@econ.unideb.hu

G. Pap

University of Debrecen, Hungary,
Faculty of Informatics,
Department of Applied Mathematics and Probabiliy Theory

E-mail: papgy@inf.unideb.hu

M.C.A. van Zuijlen

Radboud University (RU), Nijmegen, The Netherlands,
Faculty of Science,
Institute for Mathematics, Astrophysics and Particle Physics (IMAPP)

E-mail: m.vanzuijlen@math.ru.nl

[1]The author worked for the Faculty of Informatics, University of Debrecen (Hungary) during the development of the first five chapters of the lecture notes.

Contents

1 Preface

Our aim was to write lecture notes for a one-academic-term course. It presents an introduction to portfolio management for students in mathematics and economics as well. For this, elementary course on calculus and probability theory are prerequisite. Since the economic notions are explained in detail, this manuscript is self-contained from economic point of view.

This material was developed in the University of Nijmegen (the Netherlands) and Lajos Kossuth University of Debrecen (Hungary) in the framework of a three-year TEMPUS project [JEP 09269-95].

This material was produced with the aid of LaTeX. We mention that an HTML version is also available which can be viewed by a Netscape Browser, which makes the material more conveniently readable.

Our thank also goes to Tjacco van der Meer for playing an important role at the early stage of the development of the course. Preparatory discussions have been made with several members of the scientific staff of universities and institutions.

2 Introduction

There has been a great development of econometrics and, in particular, in financial mathematics in the past decades. Managing portfolios, making financial decisions under uncertain circumstances have taken an important role either in research of economics and mathematics or in practice.

One of the classical problems of the theory of finance and financial mathematics is the optimal portfolio selection. Imagine a market where some financial assets (securities) are available like treasury bonds, zero coupon bonds, stocks, options, futures. Let us suppose that one is given a certain amount of capital which shall be invested in the market. The capital can be allocated among the securities in many ways. Some allocations (that is portfolios) may be more promising (e.g. in the sense that they have large expected value in the future), whereas others can be less risky (e.g. in the sense that they have small variance of future value).

The next questions arise naturally. Which is the portfolio one would like to choose in such a market? What is the difference of the different individuals' decision making and what is it caused by? What measures or means could help us to characterize the individual's decision making, the risk aversion of the individual in such situations? These are some of the problems we shall deal with in this material.

The main issue to study is the problem of financial decision making under uncertainty and certainly our focus is mainly on portfolio choice problems. Thus, our settings are based on utility theory which is the subject of the first part of the material. We discuss here the concept of the underlying theory and study its main results which are fundamental for our further work. The concept of Neumann-Morgenstern utility theory, decision making and choice problem under certainty, ordinality and cardinality are the main topics of this part of the material.

We discuss the important features of the individual's utility function (i.e. of the individual's preferences) regarding to the individual's decision making under uncertainty. We study the possible measures of risk aversion of the individuals based on their system of preferences (risk aversion, absolute risk aversion and their characterizations are contained). Furthermore, optimal portfolio selections and their relation with the demand for financial assets are discussed in securities markets, which is the main object of the course.

The concept of stochastic dominance of financial assets (either of type first order or of type second order) is given. Another important issue we cover is the measurement of the riskiness of financial assets. For this, we study coherent measures, we focus on two very widespreadly used risk measures, namely on Value at Risk and expected shortfall.

Several examples are given in the material. We remark that another important aim was to give a precise mathematical formulation of the results which are usually omitted in many financial books. The examples are also to emphasize the importance of the mathematical and technical conditions of the certain statements and also to show that counterexamples (with very realistic parameters) can be found to refute some relations which look very natural from an intuitive point of view and therefore one would erroneously expect them to become true in the theory.

Some of the important mathematical theorems are contained in Section Appendix to help the better understanding and the easy use of the material. The Bibliographic Notes provide more information about the literature of the underlying theory, its related areas and references for the interested reader.

3 Utility Theory

More than a century ago the economic theory had the years of the marginal revolution. One of its great breakthroughs was the beginning of the development of utility theory, which has been playing a basic role in economics, especially in microeconomics, since that time. Many models and problems are based on this theory. Such a problem is the individual's decision making under either certainty or uncertainty. First we present the utility settings needed for our further purposes. However, it is not our purpose to present the whole concept of utility with the most generality.

3.1 Preference Ordering

Let us assume that there are n (different) goods. The individuals are supposed to be able to value the possible baskets of goods. A basket of goods is a vector $\mathbf{x} = (x_1, \ldots, x_n)$ where x_i $(i = 1, \ldots, n)$ indicates the amount of good i consumed if basket $\mathbf{x}$ is chosen by the individual. Denote by $\mathcal{B}$ the set of all feasible baskets. Clearly, $\mathcal{B} \subset \mathbb{R}^n$.

3.1 Remark. In many cases it will be realistic to assume for all $\mathbf{x} \in \mathcal{B}$ that $\mathbf{x} \geq \mathbf{b}$, $\mathbf{b} \in \mathbb{R}^n$, where the inequality is meant coordinate-wise. If it is valid than $\mathcal{B}$ is said to be bounded below (coordinate-wise). This assumption taken about the lower bound $\mathbf{b}$ means that though a basket can have negative entries representing selling goods, e.g. labour services, this opportunity is limited (bounded below). One can take $\mathcal{B} = \{(x_1, \ldots, x_n) \mid x_i \geq 0,\ i = 1, \ldots, n\}$ though $\mathcal{B}$ is usually bounded because of the scarcity of goods or the limited income of the individual.

3.2 Remark. Moreover, it is also usually supposed in utility theory that $\mathcal{B}$ is convex and $\mathbf{0} = (0, \ldots, 0) \in \mathcal{B}$. Thus goods can be 'infinitely divided' into parts, for instance, 1/7 of good i can be consumed as well as 1, 2 or $\sqrt{2}$.

The preferences (or preference system) of the individual are described by an ordering relation $\preceq$. The statement

$$\mathbf{x} \preceq \mathbf{y}, \qquad \mathbf{x}, \mathbf{y} \in \mathcal{B},$$

is read "$\mathbf{y}$ is preferred to $\mathbf{x}$". We may write $\mathbf{y} \succeq \mathbf{x}$ meaning the same as $\mathbf{x} \preceq \mathbf{y}$. The baskets $\mathbf{y}$ and $\mathbf{x}$ are said to be indifferent or the same, denoted by $\mathbf{x} \approx \mathbf{y}$, if both $\mathbf{x} \preceq \mathbf{y}$ and $\mathbf{y} \preceq \mathbf{x}$ are satisfied. The basket $\mathbf{y}$ is strictly preferred to basket $\mathbf{x}$, write $\mathbf{y} \succ \mathbf{x}$ or $\mathbf{x} \prec \mathbf{y}$ equivalently, if $\mathbf{x} \preceq \mathbf{y}$ but $\mathbf{y}$ and $\mathbf{x}$ are not indifferent.

We suppose that the preference ordering has the following properties.

(1) *reflexivity*, i.e. $\mathbf{x} \preceq \mathbf{x}$ for all $\mathbf{x} \in \mathcal{B}$,

(2) *transitivity*, i.e. $\mathbf{x} \preceq \mathbf{y}$ and $\mathbf{y} \preceq \mathbf{z}$ imply together $\mathbf{x} \preceq \mathbf{z}$, for $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathcal{B}$,

(3) *linearity*, i.e. for every pair $(\mathbf{x}, \mathbf{y}) \in \mathcal{B} \times \mathcal{B}$ either $\mathbf{y} \preceq \mathbf{x}$ or $\mathbf{x} \preceq \mathbf{y}$ and finally

(4) *continuity* which means that for every $\mathbf{x} \in \mathcal{B}$ the set of strictly preferred baskets and the set of strictly worse (not preferred) baskets are both open.

3.2 Utility Functions

So far we have built up a relation in order to represent the individual's preferences. The preference ordering enables the individual to make decisions about consumption plans. Obviously, if a basket $\mathbf{x}$ is preferred to basket $\mathbf{y}$ then he or she[2] is willing to choose or buy basket $\mathbf{x}$ of these two baskets. In other words, $\mathbf{x}$ is found to be more useful for him.

Now, having the concept of preference ordering, it seems natural to value each basket such that the value of a basket would show the utility of the basket according to the individual's preferences: a large value would correspond to a higher level of utility than a smaller

[2]For what follows, we will sometimes write 'he' instead of 'he or she' or instead of 'individual' for simplicity.

value. In term of mathematics, we look for a function

$$U : \mathcal{B} \mapsto \mathbb{R}$$

which satisfies the following properties:

$$\mathbf{x} \preceq \mathbf{y} \iff U(\mathbf{x}) \leq U(\mathbf{y}) \qquad \forall \mathbf{x}, \mathbf{y} \in \mathcal{B}.$$

Such a function is called utility function.

It is well known in microeconomics that such a function exists provided that our assumptions above are valid. The precise statement is the following.

3.3 Theorem. *Suppose that $\mathcal{B}$ is a separable and connected set in $\mathbb{R}^n$ and we have an ordering relation on it which is denoted by $\preceq$.*

Then $\preceq$ is a preference relation on $\mathcal{B}$, i.e. it satisfies properties (1)-(4) defined in Section 3.1, if and only if there exists a continuous function $U : \mathcal{B} \mapsto \mathbb{R}$ such that

$$\mathbf{x} \preceq \mathbf{y} \iff U(\mathbf{x}) \leq U(\mathbf{y}), \qquad \forall \mathbf{x}, \mathbf{y} \in \mathcal{B}. \tag{1}$$

Proof. *First suppose that the preference relation satisfies properties (1)-(4).* Then we construct a function U which is appropriate.

For this let Y be a countable and dense set of baskets in $\mathcal{B}$. The existence of such set follows from the separability of $\mathcal{B}$. Now due to countability we can consider Y as a sequence, i.e. write $Y = \{\mathbf{y_i}\}_{i \in \mathbb{N}}$. First we shall define a function $\bar{U} : Y \mapsto \mathbb{R}$ with property (1). We give the definition of $\bar{U}$ by induction.

Set $\bar{U}(\mathbf{y_0}) = 1/2$. If $\bar{U}(\mathbf{y_0}), \bar{U}(\mathbf{y_1}), \ldots, \bar{U}(\mathbf{y_m})$ are all defined then we take $\mathbf{y_{m+1}}$ and consider the following four cases.

(1) $\mathbf{y_{m+1}}$ is indifferent to a basket which utility value has already been defined, i.e. there is a index $0 \leq j \leq m$ such that $\mathbf{y_{m+1}} \approx \mathbf{y_j}$, or

(2) $\mathbf{y_{m+1}}$ is not indifferent to any of the previous baskets in Y such that there are two indices $0 \le i_1, i_2 \le m$ with $\mathbf{y_{i_1}} \prec \mathbf{y_{m+1}} \prec \mathbf{y_{i_2}}$. In this case let $0 \le k_1 \le m$ and $0 \le k_2 \le m$ be two indices with $\bar{U}(\mathbf{y_{k_1}}) = \max\{\bar{U}(\mathbf{y_i}) \mid \mathbf{y_{m+1}} \succ \mathbf{y_i}, \quad 0 \le i \le m\}$ and $\bar{U}(\mathbf{y_{k_2}}) = \min\{\bar{U}(\mathbf{y_i}) \mid \mathbf{y_{m+1}} \prec \mathbf{y_i}, \quad 0 \le i \le m\}$ respectively.

(3) Next take the case where $\mathbf{y_{m+1}}$ is not indifferent to any of the previous baskets in Y such that there is an index $0 \le r_1 \le m$ with $\mathbf{y_{m+1}} \prec \mathbf{y_{r_1}} \prec \mathbf{y_i}$ for all $0 \le i \le m$.

(4) Finally $\mathbf{y_{m+1}}$ may not be indifferent to any of the previous baskets in Y such that there is an index $0 \le r_2 \le m$ with $\mathbf{y_{m+1}} \succ \mathbf{y_{r_2}} \succ \mathbf{y_i}$ for all $0 \le i \le m$.

Now set

$$\bar{U}(\mathbf{y_{m+1}}) := \begin{cases} \bar{U}(\mathbf{y_j}) & \text{in case (1),} \\ \frac{\bar{U}(\mathbf{y_{k_1}})+\bar{U}(\mathbf{y_{k_2}})}{2} & \text{in case (2),} \\ \frac{\bar{U}(\mathbf{y_{r_1}})}{2} & \text{in case (3),} \\ \frac{\bar{U}(\mathbf{y_{r_2}})+1}{2} & \text{in case (4).} \end{cases}$$

Note that the set $A = \{\bar{U}(\mathbf{y_i}) \mid i \in \mathbb{N}\}$ is dense in $[a, b]$ where a and b are the infimum and supremum of A, respectively. To see this it is sufficient to show that

$$B = \left\{ \frac{k}{2^n} \in (a, b) \mid k, n \in \mathbb{N} \right\} \subset \bar{U}(Y).$$

By definition $1/2$ is equal to $\bar{U}(\mathbf{y_0})$. Then by the axiom of continuity the sets $A_1 = \{\mathbf{x} \in \mathcal{B} \mid \mathbf{x} \prec \mathbf{y_0}\}$ and $A_2 = \{\mathbf{x} \in \mathcal{B} \mid \mathbf{x} \succ \mathbf{y_0}\}$ are both open and combining it with the density of Y we get that there exist $\mathbf{y_{i_1}}$ and $\mathbf{y_{i_2}}$ in Y such that $\mathbf{y_{i_1}} \in A_1$ and $\mathbf{y_{i_2}} \in A_2$. Hence, $\bar{U}(\mathbf{y_k}) = 1/4$ and $\bar{U}(\mathbf{y_r}) = 3/4$ for some k and r $(k, l \in \mathbb{N})$. If $k/2^n = \bar{U}(\mathbf{y_{r_1}})$ and $(k+1)/2^n \in \bar{U}(\mathbf{y_{r_2}})$ than the set $\{\mathbf{x} \in \mathcal{B} \mid \mathbf{y_{r_1}} \prec \mathbf{x} \prec \mathbf{y_{r_2}}\}$ is open as well which implies that there is a basket $\mathbf{y} \in Y$ with $\bar{U}(\mathbf{y}) = (2k+1)/2^{n+1}$. Similarly argument shows that if $k \in \mathbb{N}$ is the smallest (or largest) for a fixed $n \in \mathbb{N}$ with $k/2^n = \bar{U}(\mathbf{y_r})$ for some $\mathbf{y_r} \in Y$

then either $C_r = \{\mathbf{x} \in \mathcal{B} \mid \mathbf{x} \prec \mathbf{y_r}\}$ ($C_r = \{\mathbf{x} \in \mathcal{B} \mid \mathbf{x} \succ \mathbf{y_r}\}$) is empty and therefore $a = k/2^n$ ($b = k/2^n$), or C_r is non-empty and open which implies that $(2k-1)/2^{n+1} \in \bar{U}(\mathbf{y_0})$ ($(2k+1)/2^{n+1} \in \bar{U}(\mathbf{y_0})$). Thus A is dense in $[a, b]$.

Moreover, it is clear that $\bar{U}$ satisfies property (1).

With the aid of $\bar{U}$ we define U as follows. Write $B_{\mathbf{x}} := \{\bar{U}(\mathbf{y}) \mid \mathbf{y} \in Y,\ \mathbf{y} \preceq \mathbf{x}\}$, $\mathbf{x} \in \mathcal{B}$ and then define

$$U(\mathbf{x}) := \begin{cases} \sup B_{\mathbf{x}} & \text{if } B_{\mathbf{x}} \neq \emptyset \\ \\ a & \text{otherwise.} \end{cases}$$

Furthermore, $U = \bar{U}$ over Y. It is clear that $U(\mathbf{x}) \geq \bar{U}(\mathbf{x})$ since $\mathbf{x} \in B_{\mathbf{x}}$. On the other hand $U(\mathbf{x}) > \bar{U}(\mathbf{x})$ would imply the existence of a basket $\mathbf{y} \in B_{\mathbf{x}}$ with $\bar{U}(\mathbf{x}) < \bar{U}(\mathbf{y})$ which leads us to the contradiction $\mathbf{x} \prec \mathbf{y}$.

Now we shall check whether (1) holds for U. If $\mathbf{x} \preceq \mathbf{z}$ than $B_{\mathbf{x}} \subset B_{\mathbf{z}}$ and hence $U(\mathbf{x}) \leq U(\mathbf{z})$.

To prove the other implication, first note that if $\mathbf{x} \prec \mathbf{y}$ then there exist baskets $\mathbf{y_1}, \mathbf{y_2} \in Y$ such that $\mathbf{x} \prec \mathbf{y_1} \prec \mathbf{y_2} \prec \mathbf{z}$ (use the same argument as above) which implies that $U(\mathbf{x}) \leq U(\mathbf{y_1}) < U(\mathbf{y_2}) \leq U(\mathbf{z})$. Thus, if $U(\mathbf{x}) \leq U(\mathbf{z})$ were valid together with $\mathbf{x} \succ \mathbf{z}$ than the latter would also imply $U(\mathbf{x}) > U(\mathbf{z})$ which is a contradiction.

It is only left to prove that U is continuous. The continuity of the utility function follows from the continuity of the preference ordering. Indeed, if $u \in U(\mathcal{B})$ then there is a basket $\mathbf{x_0} \in \mathcal{B}$ with $U(\mathbf{x_0}) = u$. Let $A = \{\mathbf{x} \in \mathcal{B} \mid \mathbf{x} \succ \mathbf{x_0}\}$ and $B = \{\mathbf{x} \in \mathcal{B} \mid \mathbf{x} \prec \mathbf{x_0}\}$. Then it is clear that

$$U^{-1}\big((u, \infty)\big) = A \quad \text{and} \quad U^{-1}\big((-\infty, u)\big) = B$$

and thus these are both open. (Note that by the structure theorem of open sets in $\mathbb{R}$ it is sufficient to check whether the inverse images of open intervals of type (u, ∞) or $(-\infty, u)$ are open in $\mathcal{B}$.)

Turning to the opposite implication of the theorem, suppose now that we are given a continuous function $U : \mathcal{B} \mapsto \mathbb{R}$ which satisfies property (1).

It is easy to check the validity of the preference properties (1)-(4) defined in Section 3.1.

Reflexivity: since $U(\mathbf{x}) \leq U(\mathbf{x})$, $\mathbf{x} \in \mathcal{B}$, thus $\mathbf{x} \preceq \mathbf{x}$. Transitivity: $U(\mathbf{x}) \leq U(\mathbf{y}) \leq U(\mathbf{z})$ implies $\mathbf{x} \preceq \mathbf{y} \preceq \mathbf{z}$ for all $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathcal{B}$. Linearity: given $\mathbf{x}, \mathbf{y} \in \mathcal{B}$ we have either $U(\mathbf{x}) \leq U(\mathbf{y})$ or $U(\mathbf{y}) \leq U(\mathbf{x})$, hence $\mathbf{x} \preceq \mathbf{y}$ or $\mathbf{y} \preceq \mathbf{x}$ is held. Continuity: taking a basket $\mathbf{x} \in \mathcal{B}$ and recalling that

$$U^{-1}\big((u, \infty)\big) = A \quad \text{and} \quad U^{-1}\big((-\infty, u)\big) = B,$$

where $A = \{\mathbf{x} \in \mathcal{B} \mid \mathbf{x} \succ \mathbf{x_0}\}$ and $B = \{\mathbf{x} \in \mathcal{B} \mid \mathbf{x} \prec \mathbf{x_0}\}$ it follows from the continuity of U that A and B are open.

Thus the proof is complete. □

3.4 Remark. Note that U is not unique since any strictly monotone increasing transform of U would be a utility function again, i.e. if $\Phi : \mathbb{R} \mapsto \mathbb{R}$ is strictly monotone increasing then $\Phi(U) : \mathcal{B} \mapsto \mathbb{R}$ would satisfy Theorem 3.3 as well.

In the following remark we collect some properties of the utility functions which are useful to characterize the main features of the individual's preferences.

3.5 Remark. Given a preference relation of the individual, we say that a function U is a utility function representing the preference relation if it satisfies property (1).

In most cases it is fairly natural to assume that the individual prefers more to less, i.e. given $\mathbf{x} > \mathbf{y}$ (coordinate-wise), $\mathbf{x}, \mathbf{y} \in \mathcal{B}$, we have $\mathbf{x} \succ \mathbf{y}$. This property is called *the* *principle of dominance*. In this case the utility function is strictly increasing. Furthermore, in case of differentiability it has non-negative partial derivatives.

Since the information contained by the utility function about the preferences is not more than equivalence (1), the preference ordering could be also characterized by the sets $IC_{\mathbf{x_0}} = \{\mathbf{x} \in \mathcal{B} \mid \mathbf{x} \approx \mathbf{x_0}\} = \{\mathbf{x} \in \mathcal{B} \mid U(\mathbf{x}) = U(\mathbf{x_0})\}$, $(\mathbf{x_0} \in \mathcal{B})$. Such a set represents the class of baskets indifferent to a basket and called indifferent curve or surface in economics.

Now suppose that U is a twice differentiable, increasing utility function. Let $\mathbf{x^0} \in \mathcal{B}$ and let us assume, furthermore, that U is strictly concave in a neighbourhood V of $\mathbf{x^0}$ $(V \subset \mathcal{B})$. (The economic explanation of concavity will be discussed later.) Note that $\frac{\partial}{\partial x_i} U(\mathbf{x^0}) > 0$. If $\mathbf{x} \in V \bigcap IC_{\mathbf{x^0}}$ then any coordinate of $\mathbf{x}$ can be expressed by the remaining coordinates. Indeed, by the implicit function theorem (see Appendix, Theorem 8.2) there exists a function g on a neighbourhood V^* of $\mathbf{x_i^0} = (x_1^0, \ldots, x_{i-1}^0, x_{i+1}^0, \ldots, x_n^0)$ in $\mathbb{R}^{n-1}$ such that

$$U\big(x_1, \ldots, x_{i-1}, g(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n), x_{i+1}, \ldots, x_n\big) = U(\mathbf{x^0}) \tag{2}$$

where $(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n) \in V^*$. In other words, if we fix the utility level (in our case it is $u = U(\mathbf{x^0})$) and we are only moving on the indifferent surface corresponding to this utility level (i.e. the indifferent surface lying on $\mathbf{x^0}$) then the point of our location can be identified by only $n-1$ of the n coordinates, that is, any coordinate can be written in terms of the others at least in a neighbourhood of $\mathbf{x^0}$. Thus g is nothing else but x_i (uniquely determined by $x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n$), so we will simply write x_i instead of g (e.g. in (3)) to make the meaning of our setting more intuitive.

Hence it is clear now that the change of the consumption of $n-1$ goods in the basket will uniquely determine the change of the consumption of the remaining one good if we shall attain the same level of utility (and of course if the changes are small enough not to get out of V^*). Now taking the partial derivative of (2) with respect to the jth coordinate one can easily get

$$-\frac{\partial x_i}{\partial x_j}(\mathbf{x_i^0}) = \frac{\frac{\partial}{\partial x_j}\,U(\mathbf{x^0})}{\frac{\partial}{\partial x_i}\,U(\mathbf{x^0})} \qquad i \neq j, \quad 1 \leq i, j \leq n. \tag{3}$$

The latter expression is called *the marginal rate of substitution* (between good i and good j) and denoted by $MRS_{i,j}$. Intuitively it shows how many units of good i should be renounced to keep the same utility level if one more unit of good j is to be consumed:

$$MRS_{i,j} \approx -\frac{\Delta x_i}{\Delta x_j}$$

Note that the indifference curves and the marginal rate of substitution are invariant to any strictly increasing transformation of the utility function.

3.3 Ordinality Versus Cardinality

When the notion of utility function was first used, e.g. by Walras, Menger and Jevons, in the 80's of the last century at the time of the marginal revolution, the utility function was expected to give more information about the "level of satisfactory" than property (1). The above authors actually worked with utility functions of one variable (representing only one good), that is they defined a separate utility function for each good. (Here we mention that the utility of a basket was defined by the aggregate utility of the good's utilities in this case.) They assumed that the utility values are suitable not only to choose the better, for instance, of two given baskets of goods but also to express by how much one alternative is better than the other. For example, if $U(x_0) = 2$ and $U(x_1) = 4$ then they said that x_1 is found to be twice as good or useful as x_0 for the individual.

Such a utility function is called *cardinal* in contrast to *ordinal* utility functions which do not give more information about the baskets than what is expressed in (1). Thus Theorem 3.3 states the existence of a continuous ordinal utility function which is consistent with the underlying preference ordering in the sense of property (1).

Thought first cardinal utility appeared in the economic theories and ordinality came later we should not like to suggest that the latter is the right one, not even that it services better the aims of the economic theory in general. It is still an important issue of recent research in economics. We can only say that for certain problems the ordinal approach provides better means but for other problems it does not. For our purposes ordinal utility functions suffice, as we have seen in Theorem 1 and as we will still see it in the following sections.

3.4 Utility Maximization

A classical problem in microeconomics is the determination of the optimal (meaning 'the most preferred') basket if the individual possesses a certain income (or wealth). It is a utility maximization over the set, say $\mathcal{F}$, of the feasible (economically available) baskets. Clearly, $\mathcal{F} = \{\mathbf{x} \in \mathcal{B} \mid \sum_{i=1}^{n} x_i p_i \leq I\}$, where I is the income and p_i is the price of the ith good. So, our aim is to find the maximum of U over $\mathcal{F}$.

3.6 Theorem. *Let $\mathcal{B}$ be closed, convex, bounded below (coordinate-wise) and suppose that $\mathbf{0} \in \mathcal{B}$ and $U : \mathcal{B} \mapsto \mathbb{R}$ is an increasing, continuous, strictly concave utility function. For any price vector $(p_1, \ldots, p_n)$ with $p_i > 0$, $i = 1, \ldots, n$, and income $I > 0$ there is a unique $\mathbf{x}$ in $\mathcal{B}$ such that*

$$U(\mathbf{x}) = \max_{\mathbf{y} \in \mathcal{F}} U(\mathbf{y})$$

where $\mathcal{F} = \{\mathbf{y} \in \mathcal{B} \mid \sum_{i=1}^{n} y_i p_i \leq I\}$. Furthermore, $\sum_{i=1}^{n} x_i p_i = I$.

Proof. The set $\mathcal{F}$ is bounded above (clearly $y_i \leq I/p_i$ for $\mathbf{y} = (y_1, \ldots, y_n) \in \mathcal{F}$), thus $\mathcal{F}$ is bounded, closed and not empty, hence it is compact. Since U is continuous it must attain its global maximum, say at $\mathbf{x}$. If U took its maximum at $\mathbf{x}^* \in \mathcal{B}$, $\mathbf{x}^* \neq \mathbf{x}$, as well then we would have

$$\frac{\mathbf{x} + \mathbf{x}^*}{2} \in \mathcal{F} \qquad \text{with} \qquad U\left(\frac{\mathbf{x} + \mathbf{x}^*}{2}\right) > \frac{1}{2}U(\mathbf{x}) + \frac{1}{2}U(\mathbf{x}^*) = U(\mathbf{x})$$

by the concavity of U. Hence $\mathbf{x}$ is unique. Finally, it is trivial that $\sum_{i=1}^{n} x_i p_i = I$. Otherwise, for instance, the basket

$$\left(x_1 + \frac{I - \sum_{i=1}^{n} p_i x_i}{p_1}, x_2, \ldots, x_n \right)$$

would be strictly preferred to $\mathbf{x}$. □

Solving the optimum problem

To solve such a problem of optimal allocation, one can use the method of Lagrangian multipliers (see Appendix, Theorem 8.1) in case U is differentiable. Taking the partial derivatives of $U(\mathbf{x}) + \lambda(I - \sum_{i=1}^{n} p_i x_i)$ with respect to x_i $(i = 1, \ldots, n)$ and λ we obtain the following system from the first order conditions:

$$\begin{aligned} \frac{\partial}{\partial x_i} U(\mathbf{x}) &= \lambda p_i \qquad i = 1, \ldots, n \\ I - \sum_{i=1}^{n} p_i x_i &= 0. \end{aligned} \tag{4}$$

Since U is concave over the convex set $\mathcal{F}$, a solution of (4) will necessarily be a global maximum of U over $\mathcal{F}$.

The Second Law of Gossen

Now consider the $n-1$ dimensional hyperplane T determined by $I = \sum_{i=1}^{n} p_i x_i$. It is clear that the solution belongs to $T \cap \mathcal{F}$. Furthermore, if the maximum is taken at an inner point of $T \cap \mathcal{F}$ in space T then we can find it with the aid of the Lagrangian multipliers. However, the maximum might not be found with the Lagrangian multipliers if the point where the maximum is achieved belongs to the set[3] $\partial(T \cap \mathcal{F})$ in space T. See also Examples 3.7, 3.8. (Note that the boundary of $T \cap \mathcal{F}$ is meant to be taken with respect to the topology of space T in T and not in the n dimensional space $\mathbb{R}^n$ where it would give back the set itself.)

[3]Given a set A in a topological space, ∂A is to denote the set of boundary points of A.

If the optimum is obtained from (4) then it has got the following property:

$$MRS_{j,i} = \frac{p_i}{p_j} \qquad i \neq j, \quad 1 \leq i, j \leq n,$$

which can be easily seen from (4). Thus, at the optimum the marginal rate of substitution ($MRS_{j,i}$), which is determined by the individual's preferences, equals the (marginal) rate of substitution provided by the market (p_i/p_j). One would certainly substitute some amount of good j with good i if $MRS_{j,i}$ was larger than p_i/p_j. Or equivalently, we have at the optimum

$$\frac{\frac{\partial}{\partial x_i} U(\mathbf{x})}{p_i} = \frac{\frac{\partial}{\partial x_j} U(\mathbf{x})}{p_j} \qquad i \neq j, \quad 1 \leq i, j \leq n, \tag{5}$$

which means that an additional unit of money would cause the same increase of the utility, no matter which good would it be spent for. Formula (5) is known as the Second Law of Gossen.

The demand curve

Now take a good, say i, and suppose that the prices of the remaining goods are fixed in the market as well as the income and assume that the conditions of Theorem 3.6 remain valid. Then by utility maximization there can be found the corresponding basket $\mathbf{x}_{p_i}$ to any price $p_i > 0$. Thus the set $D_i = \{(p, q) \mid p = p_i, \; q = x_{p_i,i}\}$ can be derived and drawn in $\mathbb{R}^2$ (where $x_{p_i,i}$ is the ith entry of the optimal portfolio. The set D_i is called the individual's demand curve for good i since it shows us the relationship between the market price of a good and the individual's consumption from this good (provided all the rest of the model is fixed). To study the demand curve and related problems is out of the scope of this course. The interested readers are referred to introductory books on microeconomics such as [Nordhaus & Samuelson] or [Kreps]. We will, however, investigate the individual's demand under uncertainty (see Section 4.4).

3.5 Commonly Used Utility Functions

In this section we assume that $\mathcal{B} = \{\mathbf{x} \in \mathbb{R}^n \mid x_i \geq 0, \quad i = 1, \ldots, n\}$. Furthermore, $I > 0$ will be the income of the individual and p_i's ($p_i > 0$, $i = 1, \ldots, n$) are to denote the market prices of the goods.

3.7 Example. (Cobb-Douglas Utility Function) The Coob-Douglas type utility function is defined as follows.

$$U(\mathbf{x}) = \prod_{i=1}^{n} x_i^{a_i} \qquad a_i > 0, \quad i = 1, \ldots, n.$$

For convenience, we can use $\ln U$ for the calculations (see Remark 3.4). Now, from (3) we have

$$MRS_{i,j} = \frac{\frac{\partial}{\partial x_j}\left(\sum_{k=1}^{n} a_k \ln x_k\right)}{\frac{\partial}{\partial x_i}\left(\sum_{k=1}^{n} a_k \ln x_k\right)} = \frac{a_j/x_j}{a_i/x_i} = \frac{a_j x_i}{a_i x_j}. \tag{6}$$

Intuitively, $MRS_{i,j} \approx \Delta x_i / \Delta x_j$ and combining it with (6) we obtain

$$\frac{\Delta x_j / x_j}{\Delta x_i / x_i} \approx \frac{a_j}{a_i}$$

which shows us that for a certain increase expressed in terms of proportion of good j, a constant proportion of good i must be renounced everywhere in $\mathcal{B}$ provided that the level of the utility remains the same. So, 'the proportional rate of substitution' is constant over $\mathcal{B}$.

To get the optimal allocation from (4) we obtain $a_i/x_i = \lambda p_i$ $(i = 1, \ldots, n)$ and $I = \sum_{i=1}^{n} a_i/\lambda$. Hence,

$$x_i = \frac{a_i}{\lambda p_i} = \frac{a_i I}{p_i \sum_{k=1}^{n} a_k} .$$

3.8 Example. (Linear Utility Function) Let

$$U(\mathbf{x}) = \sum_{i=1}^{n} a_i x_i \qquad \text{with} \quad a_i > 0, \quad i = 1, \ldots, n.$$

In this case the marginal rate of substitution is constant:

$$MRS_{i,j} = \frac{a_j}{a_i} , \qquad 1 \leq i, j \leq n.$$

This means that the goods can be perfectly substituted with one another and therefore the consumer maximizing utility will buy only one type of good, namely, the cheapest according to his preferences. Let i be the index ($i \in \{1, \dots, n\}$) for which

$$\frac{a_i}{p_i} = \max_{1 \le j \le n} \frac{a_j}{p_j} \, .$$

Then an optimal basket is $\mathbf{x} = (x_1, \dots, x_n)$ where

$$x_j = \delta_{ij} \, \frac{I}{p_i} \qquad j = 1, \dots, n.$$

Note that $\mathbf{x}$ is not necessarily unique. If there is an index $j \in \{1, \dots, n\}$ such that $i \neq j$ and $a_i/p_i = a_j/p_j$ then the set of optimal allocations is infinite.

3.9 Example. (Complementary Goods) Now suppose that the utility function is of the form

$$U(\mathbf{x}) = \min\{a_i x_i \mid i = 1, \dots, n\},$$

where $a_i > 0$. Thus this utility function is continuous but not differentiable. Let $\mathbf{x}$ be a basket and $i \in \{1, \dots, n\}$ an index such that

$$a_i x_i = \min\{a_j x_j \mid j = 1, \dots, n\}.$$

If $a_j x_j > a_i x_i$ ($j \in \{1, \dots, n\}$) then the individual possesses some of good j unnecessarily. To make the meaning of 'unnecessarily' precise define $x_j^* = a_i x_i / a_j$. Then $x_j > x_j^*$ and x_j^* would be enough to attain the same level of utility than with x_j, i.e. the amount $x_j - x_j^*$ is useless. In other words, for a given x_i we need exactly $x_j = \frac{a_i}{a_j} x_i$ of good j ($j = 1, \dots, i-1, i+1, \dots, n$) to attain the utility level $u = a_i x_i$ without wasting any money. That is why such a utility function is used in case of complementary goods. They are consumed in a certain ratio (determined by the a_i's) but they cannot be substituted with one another.

Turning to the problem of utility maximization it is clear from the above argument that $a_i x_i = a_j x_j$ will be satisfied at the optimum for all $i, j \in \{1, \dots, n\}$. Therefore the optimal allocation is the intersection point of the line given by

$$x_i = \frac{1}{a_i} t, \qquad t \in \mathbb{R}, \ i = 1, \dots, n$$

and the hyperplane given by

$$\sum_{i=1}^{n} x_i p_i = I.$$

Hence,

$$x_i = \frac{I}{a_i \sum_{j=1}^{n} \frac{p_j}{a_j}} \qquad \text{for } i = 1, \ldots, n.$$

Further examples will be studied in the forthcoming sections.

3.6 Expected Utility

Utility functions defined in the previous sections give a nice interpretation of the individual's preferences and so (using the utility maximization) some classical problems of microeconomics can be solved with the aid of utility functions, like optimal choice, the derivation of individual's demand function, etc. However, we usually must face choice problems involving risk, that is we must choose from possibilities which have got uncertain outcomes. Take, for instance, a portfolio of different financial assets: cash, bonds, stocks, futures, options, etc. The value or payoff of such a portfolio at a future time point is uncertain (or risky) since it depends on the future state of the world (or economy).

The idea of solving the choice problem in case of uncertainty is the maximization of the expectation of utility. For this we need some further definition first.

3.10 Definition. *Let $X = (\mathbf{x}_1, \ldots, \mathbf{x}_m)$ with $\mathbf{x}_i \in \mathcal{B}$ $(i = 1, \ldots, m)$ and $P = (p_1, \ldots, p_m)$ with $\sum_{i=1}^{m} p_i = 1$, $0 \le p_i \le 1$ $(i = 1, \ldots, m)$. Then the pair (X, P) is called lottery.*
Let $\mathbb{L}$ denote the set of all lotteries.
Given $\mathcal{L}_1 = (X^1, P^1), \ldots, \mathcal{L}_k = (X^k, P^k) \in \mathbb{L}$ and $\alpha_1, \ldots, \alpha_k \in [0, 1]$ with $\sum_{i=1}^{k} \alpha_i = 1$, the

convex linear combination $\sum_{i=1}^{k} \alpha_1 \mathcal{L}_i$ *is defined by* (X, P) *where*

$$X = (\mathbf{x}_1^1, \ldots, \mathbf{x}_{m_1}^1, \mathbf{x}_1^2, \ldots, \mathbf{x}_{m_2}^2, \ldots, \mathbf{x}_1^k, \ldots, \mathbf{x}_{m_k}^k),$$

$$P = (\alpha_1 p_1^1, \ldots, \alpha_1 p_{m_1}^1, \alpha_2 p_1^2, \ldots, \alpha_2 p_{m_2}^2, \ldots, \alpha_k p_1^k, \ldots, \alpha_k p_{m_k}^k).$$

Thus the convex linear combination of lotteries is defined to be a lottery again.

3.11 Remark. A lottery $\mathcal{L}$ can be interpreted as a gamble with m different outcomes such that p_i is the probability of the event that $\mathbf{x}_i$ will be the outcome of the gamble. In other words, each lottery corresponds to a discrete random vector l on a certain probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with $\mathbb{P}(l = \mathbf{x}_i) = p_i$. For simplicity, we will denote this random vector by $\mathcal{L}$ as well.

As we made an ordering on the set of the basket of goods in Section 3.1, now we should like to make an ordering on $\mathbb{L}$.

Suppose that $\mathcal{L}_1 = (X_1, P_1)$ and $\mathcal{L}_1 = (X_2, P_2)$ are two lotteries. Let $\mathcal{L}_1 \preceq \mathcal{L}_2$ mean that the individual prefers $\mathcal{L}_2$ to $\mathcal{L}_1$. Denote by $\mathcal{L}_1 \approx \mathcal{L}_2$ and read '$\mathcal{L}_1$ and $\mathcal{L}_2$ are the same or indifferent' if $\mathcal{L}_1 \preceq \mathcal{L}_2$ and $\mathcal{L}_2 \preceq \mathcal{L}_1$. If $\mathcal{L}_1 \preceq \mathcal{L}_2$ but they are not indifferent then $\mathcal{L}_2$ is said to be strictly preferred to $\mathcal{L}_1$. We require the following properties.

(1) *reflexivity*: $\mathcal{L} \preceq \mathcal{L}$ for all $\mathcal{L} \in \mathbb{L}$,

(2) *transitivity*: if $\mathcal{L}_1 \preceq \mathcal{L}_2$ and $\mathcal{L}_2 \preceq \mathcal{L}_3$ then $\mathcal{L}_1 \preceq \mathcal{L}_3$ for all $\mathcal{L}_1, \mathcal{L}_2, \mathcal{L}_3 \in \mathbb{L}$,

(3) *linearity*: if $\mathcal{L}_1, \mathcal{L}_2 \in \mathbb{L}$ then we have either $\mathcal{L}_1 \preceq \mathcal{L}_2$ or $\mathcal{L}_2 \preceq \mathcal{L}_1$,

(4) *continuity*: if $\mathcal{L}_1, \mathcal{L}_2, \mathcal{L}_3$ are lotteries satisfying $\mathcal{L}_1 \preceq \mathcal{L}_2$ and $\mathcal{L}_2 \preceq \mathcal{L}_3$ then there exist a constant p in $[0, 1]$ such that $p\mathcal{L}_1 + (1 - p)\mathcal{L}_3 \approx \mathcal{L}_2$,

(5) *dominance*: If $\mathcal{L} \in \mathbb{L}$ and $p \in [0, 1]$ then there exists a lottery $\mathcal{L}'$ such that $\mathcal{L}$ is not indifferent with $p\mathcal{L} + (1 - p)\mathcal{L}'$,

(6) *independence*: if $\mathcal{L} \preceq \mathcal{L}'$, $\mathcal{L}, \mathcal{L}' \in \mathbb{L}$, and $p \in [0, 1]$ then for each $\mathcal{K} \in \mathbb{L}$

$$p\mathcal{L} + (1-p)\mathcal{K} \preceq p\mathcal{L}' + (1-p)\mathcal{K}.$$

Properties (1)-(6) might be called axioms, for instance, the sixth one is often referred as the axiom of independence in literature.

We simply state now the basic result and omit the proof of it.

3.12 Theorem. *If we have a preference ordering satisfying properties (1)-(4) in Section 3.1 and also an ordering on the lotteries satisfying properties (1)-(6) of this section then there exists a function* $V : \mathbb{L} \mapsto \mathbb{R}$ *such that for all* $\mathcal{L}_1, \mathcal{L}_2 \in \mathbb{L}$

$$\mathcal{L}_1 \preceq \mathcal{L}_2 \quad \Longleftrightarrow \quad V(\mathcal{L}_1) \leq V(\mathcal{L}_2). \tag{7}$$

such that V *admits the following representation:*

$$V(\mathcal{L}) = \sum_{i=1}^{m} U(\mathbf{x}_i)p_i, \qquad \forall \mathcal{L} \in \mathbb{L} \tag{8}$$

where $\mathcal{L} = (X, P) = ((\mathbf{x}_1, \ldots, \mathbf{x}_m), (p_1, \ldots, p_m))$.

Proof. First take two baskets $\mathbf{x_1}$, $\mathbf{x_2} \in \mathcal{B}$ such that they are not the same. For instance, let us assume that $\mathbf{x_1} \prec \mathbf{x_2}$. If such baskets do not exists then all the baskets were indifferent of course and thus U could be chosen constant on $\mathcal{B}$.

Let $[\mathbf{x_1}, \mathbf{x_2}]_{\preceq} = \{\mathbf{x} \in \mathcal{B} \mid \mathbf{x_1} \preceq \mathbf{x} \preceq \mathbf{x_2}\}$. First we construct a function U on $[\mathbf{x_1}, \mathbf{x_2}]_{\preceq}$ which property 8 is valid for.

Define $U(\mathbf{x_1}) = 0$ and $U(\mathbf{x_2}) = 1$ and $L_{\mathbf{x_1},\mathbf{x_2}} = \{\mathcal{L} \in \mathbb{L} \mid \mathcal{L} \in [\mathbf{x_1}, \mathbf{x_2}]_{\preceq}\}$.

Now take a basket $\mathbf{y} \in [\mathbf{x_1}, \mathbf{x_2}]_{\preceq}$ and denote by $\mathcal{L}_{\mathbf{y}}$ the lottery which has only $\mathbf{y}$ as a possible outcome, i.e. $\mathbb{P}(\mathcal{L}_{\mathbf{y}} = \mathbf{y}) = 1$. Due to the continuity of the preference relation there is a constant $t_{\mathbf{y}} \in [0, 1]$ such that

$$\mathcal{L}_{mathbfy} \approx (1 - t_{\mathbf{y}})X_1 + t_{\mathbf{y}}X_2.$$

Moreover, such a constant is unique which follows from the monotonicity of the relation. Define $U(\mathbf{y}) = t_{\mathbf{y}}$. Hence we have

$$\mathcal{L}_{\mathbf{y}} \approx (1 - U(\mathbf{y}))X_1 + U(\mathbf{y})X_2. \tag{9}$$

If $\mathcal{L} \in L_{\mathbf{x_1},\mathbf{x_2}}$ with $\mathbb{P}(\mathcal{L} = \mathbf{y_i}) = p_i$, $i = 1, \dots, k$, where $\sum_{i=1}^{k} p_i = 1$ then combining (9) with the axiom of independence we get

$$\begin{aligned} \mathcal{L} &\approx \sum_{i=1}^{k} p_i \mathcal{L}_{\mathbf{y_i}} \approx \sum_{i=1}^{k} p_i \Big[\big(1 - U(\mathbf{y_i})\big)X_1 + U(\mathbf{y_i})X_2\Big] \\ &\approx \left[1 - \sum_{i=1}^{k} p_i U(\mathbf{y_i})\right] X_1 + \sum_{i=1}^{k} p_i U(\mathbf{y_i})X_2 \approx \big(1 - \mathbb{E}\, U(\mathcal{L})\big)X_1 + \big(\mathbb{E}\, U(\mathcal{L})\big)X_2. \end{aligned} \tag{10}$$

The monotonicity of the relation together with (10) implies that

$$\begin{aligned} \mathcal{L}_1 \preceq \mathcal{L}_2 &\Longleftrightarrow (1 - \mathbb{E}\, U(\mathcal{L}_1))X_1 + (\mathbb{E}\, \mathcal{L}_1)X_2 \preceq (1 - \mathbb{E}\, U(\mathcal{L}_2))X_1 + (\mathbb{E}\, \mathcal{L}_2)X_2 \\ &\Longleftrightarrow \mathbb{E}\, \mathcal{L}_1 \leq \mathbb{E}\, \mathcal{L}_2. \end{aligned}$$

Note that with values $U(\mathbf{x_1}) = 0$ and $U(\mathbf{x_2}) = 1$ the only possible choice for the value $U(\mathbf{y})$ is $t_{\mathbf{y}}$ if we want to achieve property (8). Since, taking a value different from $U(\mathbf{y})$ equation (9) would not remain valid. Thus we can see that a utility function U satisfying (8) with $U(\mathbf{x_1}) = 0$ and $U(\mathbf{x_2}) = 1$ must be unique over $[\mathbf{x_1}, \mathbf{x_2}]_{\preceq}$.

Now we can extend U. To see this take a basket $\mathbf{x_3}$ which does not belong to $[\mathbf{x_1}, \mathbf{x_2}]_{\preceq}$. Then we have either $\mathbf{x_3} \preceq \mathbf{x_1}$ or $\mathbf{x_2} \preceq \mathbf{x_3}$. Suppose for instance the first case. (The latter case can be handled analogously.) We can construct a utility function, say $\bar{U}$ on $[\mathbf{x_3}, \mathbf{x_2}]_{\preceq}$ in the same way as above. Then define

$$\tilde{U}(\mathbf{y}) = \frac{\bar{U}(\mathbf{y}) - \bar{U}(\mathbf{x_1})}{\bar{U}(\mathbf{x_2}) - \bar{U}(\mathbf{x_1})}, \quad \text{for} \quad \mathbf{y} \in [\mathbf{x_3}, \mathbf{x_2}]_{\preceq}.$$

Thus $\tilde{U}(\mathbf{x_1}) = 0$ and $\tilde{U}(\mathbf{x_2}) = 1$ which means that $\tilde{U}$ and U must be equal over $[\mathbf{x_1}, \mathbf{x_2}]_{\preceq}$ because of the uniqueness of U.

So we have proved that U can be extended over the whole $\mathcal{B}$.

$\square$

Now, let l_1 and l_2 be the corresponding random vector of the lotteries $\mathcal{L}_1$ and $\mathcal{L}_2$ respectively. The above theorem states the existence of an ordinal utility function such that

$$\mathcal{L}_1 \preceq \mathcal{L}_2 \quad \Longleftrightarrow \quad \mathbb{E}\, U(l_1) \leq \mathbb{E}\, U(l_2).$$

It can also be proved that U in Theorem 3.12 is unique apart from a strict monotone increasing linear (i.e. positive affine) transformation. The precise statement is the following.

3.13 Theorem. *Suppose that the conditions of Theorem 3.12 are valid. Let U_1 and U_2 be two ordinal utility function and define for $k = 1, 2$*

$$V_k(\mathcal{L}) = \sum_{i=1}^{m} U_k(\mathbf{x}_i) p_i, \qquad \forall \mathcal{L} \in \mathbb{L},$$

where $\mathcal{L} = (X, P) = ((\mathbf{x}_1, \ldots, \mathbf{x}_m), (p_1, \ldots, p_m))$. If V_1 and V_2 both satisfy property (7) then there exist constants $a, b \in \mathbb{R}$, $a > 0$ such that

$$U_2(\mathbf{x}) = a U_1(\mathbf{x}) + b, \qquad \mathbf{x} \in \mathcal{B}.$$

3.14 Remark. The function U in Theorem 3.12, that is a function which the expected utility property (1) is held for, is called *Neumann-Morgenstern utility function or index.* Although Ramsey was the first who introduced the theory of expected utility first from a modern approach in the 30's, Neumann and Morgenstern, who developed the theory also but separately from Ramsey, are said to be the establishers of the theory at issue.

For what follows, a utility function will be meant to be Neumann-Morgenstern type unless it is defined otherwise.

3.7 Paradoxical Empirical Results

Though the concept of expected utility is commonly used in economics there are some known experimental results which state that some of our assumptions are refuted in practice. Most of these experiments refute the axiom of independence.

The first and best-known empirical refutation was the Allais paradox. The test of Allais has been repeated by many other researchers in different ways, but each contradicts the axiom of independence. Here we present the test of Kahnemann and Tversky.

Students were given several question. They were asked to choose one of the following two lotteries which were offered to them for free:

- *Lottery* A = ((4000 IS, 0 IS), (0.8, 0.2)), that is they can gain 4000 IS with probability 0.8 and gain nothing with probability 0.2. (Here IS denotes the Israeli Shekel)
- *Lottery* B = ((3000 IS), (1)), so this game assures the owner to obtain 3000 IP surely.

The majority of the students, namely 80%, chose the second lottery. Then they were given again a choice problem, this time with the following lotteries:

- *Lottery* C = ((4000 IS, 0 IS), (0.2, 0.8)),
- *Lottery* D = ((3000 IS, 0 IS), (0.25, 0.75)).

Now, 65% of the same students gave answer C to this question. So, there must have been many of them who chose B in the first case and C in the second problem. If those had made their decision on the basis of expected utility of the lotteries then we would have

$$U(4000) \cdot 0.8 < U(3000) \tag{11}$$

according to the first answer where U is their utility function and we assumed $U(0) = 0$ which can always be obtained by a linear transformation (see the previous section). (Note that in this examples we only indicate one variable of the utility function representing the good called 'currency'.) However, their second decision implies

$$U(4000) \cdot 0.2 > U(3000) \cdot 0.25$$

which contradicts inequality (11) and actually the axiom of independence. To see this let Z be the lottery which gives payoff zero surely. Then it is easy to see that $C = 0.25A + 0.75Z$ and similarly $D = 0.25B + 0.75Z$. But, if the axiom of independence was valid then we would have

$$A \preceq B \quad \Longleftrightarrow \quad C \preceq D.$$

In another questionnaire Kahnemann and Tversky first put the next question to some people.

Question I. You are given 1000 IS and then choose either (a) or (b):

(a) you can gain 1000 IS additionally with probability 0.5 or nothing with probability 0.5,

(b) you get 500 IS.

Question II. You are given 2000 IS and then choose either (c) or (d):

(c) you may loose 1000 IS with probability 0.5 or nothing with probability 0.5,

(d) you loose 500 IS for sure.

Surprisingly, 84% of the people turned out to prefer I/(b) to I/(a) but only 27% preferred II/(d) to II/(c). However, I/(a) and II/(c) can be represented by the same lottery, namely by ((2000 IS, 1000 IS), (0.5, 0.5)) and the same is valid for the pair I/(b) and II/(d) which both coincide the lottery ((1500), (1)).

Finally we mention an experiment made by Lichtenstein and Slovic about the so-called preference reversal. They presented two gambles:

I: a small amount of money can be gained with a large probability,

II: a huge amount of money can be gained with a small probability.

The majority of the people preferred the first game. Next they were asked to price the two gambles and mostly a higher price was given to the second gamble.

These empirical results show that people do not always make their decisions in a rational way or they do so but not on the basis of our assumptions. Considering the second example, some scientists suggest that utility functions of the possible gains and losses should be used instead of the utility functions of wealth.

4 Risk Aversion

In this section we shall study further properties of the utility functions. Our interest is focused on the behaviour of the individual with respect to risky assets (assets which have random value). The underlying analysis will show us what circumstances will lead the individual to the acceptance of risky assets and what the relation is between the acceptance and the form of the utility function.

4.1 Risk Aversion

Notational remarks As we mentioned before, in the following the utility function, say U, will always be a Neumann-Morgenstern utility function (see Remark 3.14) unless it is defined otherwise. Although U is a multivariate function (having as many variables as types of goods are given), in most of the financial problems we are going to deal with it is sufficient to study only one variable representing the 'money' as a good. In these cases utility functions will be indicated as univariate real-valued functions for simplicity. Thus the rest of the variables are supposed to be fixed and therefore we omit to indicate them.

4.1 Definition. *Let $U : I \mapsto \mathbb{R}$ be a utility function, where $I \subset \mathbb{R}$ is an interval, and ξ be a random variable on a certain probability space and with values in I. Then ξ is said to be a gamble. Let $P \in \mathbb{R}$ and suppose that $\mathbb{E}\,|\xi| < \infty$. Then the pair (ξ, P) is called a fair gamble or P is called a the fair price for the gamble ξ if $\mathbb{E}\,\xi = P$.*

We call the individual having utility function U risk averse at $P \in I$ if he is unwilling to accept or indifferent to any (ξ, P) fair gamble, i.e. $\mathbb{E}\,U(\xi) \leq U(P)$. If an individual is risk averse at P and he is not indifferent to any fair gambles with price P then he is said to be strict risk averse at P. If he is risk averse or strict risk averse over the whole I then he is said to be (global) risk averse or (global) strict risk averse.

4.2 Theorem. *Suppose that an individual has utility function $U : I \mapsto \mathbb{R}$ where $I \subset \mathbb{R}$ is an interval.*

Then the individual is (strictly) risk averse at $P \in \mathbb{R}$ if and only if U is (strictly) concave at P.

The individual is (strictly) risk averse if and only if U is (strictly) concave.

Proof. We only have to prove the first statement since the second is a direct consequence of it.

Sufficiency. Let U be concave at $P \in I$. Then, by the concavity, there exists a constant c such that

$$U(x) \leq c(x - P) + U(P), \qquad \text{for all} \quad x \in I, \tag{12}$$

and hence for all fair gamble (ξ, P) we have

$$U(\xi) \leq c(\xi - P) + U(P). \tag{13}$$

Taking the expectation of (13) the first term on the right-hand side vanishes and we obtain the desired statement:

$$\mathbb{E}\, U(\xi) \leq U(\mathbb{E}\, \xi).$$

Necessity. Consider a simple gamble ξ which has only two possible outcomes, x and y such that $p = \mathbb{P}(\xi = x)$ and $1 - p = \mathbb{P}(\xi = y)$ and suppose that $px + (1 - p)y = P$. Then the risk aversion at P implies

$$\mathbb{E}\, \big(U(\xi)\big) = p\, U(x) + (1 - p)\, U(y) \leq U(px + (1 - p)y) = U(P), \tag{14}$$

from which the concavity at P follows.

The proof in the case of strict concavity and strict risk aversion remains almost the same apart from minor changes like in (12), (13) and (14) the strict inequalities are satisfied if $x, \xi, y \neq P$. □

4.3 Remark. Note that the first part of the proof of Theorem 4.2 is actually the proof of Jensen's inequality. Moreover, the above theorem states nothing else but the fact that a real function g is concave if and only if $\mathbb{E}\ g(\xi) \leq g(\mathbb{E}\ \xi)$ holds for all random variable taking values in the domain of g.

4.4 Remark. Although our interest is mainly focused on financial problems we mention that the statement of Theorem 4.2 can be generalized by making some modifications in the above setting. In fact we can handle the case of more goods as it might be useful for the reader interested in microeconomics since risk can occur in many problems of decision making.

To see this let $\xi = (\xi_1, \ldots, \xi_n)$ be a random vector with values in $\mathcal{B}$ and suppose that $\mathbb{E}\ |\xi_i| < \infty$ $(i = 1, \ldots, n)$. Recall that U is defined on $\mathcal{B}$. Given income I and the market prices of the goods $p_1, \ldots, p_n$, denote the optimal allocation corresponding to I (see Theorem 3.6) by $\mathbf{x}_I^{opt}$. Then we call the price $P \in \mathbb{R}$ fair for the gamble ξ for the individual if $U(\mathbb{E}\ \xi) = U\left(\mathbf{x}_P^{opt}\right)$. This means that the gamble ξ is indifferent to income P if the expected utility of the gamble is the same as the maximal utility which can be achieved by the optimal basket corresponding to P. We have $\sum_{i=1}^{n} p_i x_{P,i}^{opt} = P$.

It is clear now that one can get back the setting of Theorem 4.2 by choosing $n = 1$ and $p_1 = 1$.

Now the above theorem can be reformulated as follows.

The individual is unwilling to accept any fair gamble (ξ, P) *with* $\mathbb{E}\,\xi = \mathbf{x}$ *if and only if* U *is concave at* $\mathbf{x}$.
The individual is unwilling to accept any fair gamble (that is he is risk averse) if and only if his utility function is concave.

The proof of theorem 4.2 can be rewritten almost literally. The concavity of U at $\mathbb{E}\,\xi$

now implies the existence of a vector $c \in \mathbb{R}^n$ such that

$$U(\mathbf{x}) \leq U(\mathbb{E}\,\xi) + \langle c, \mathbf{x} - \mathbb{E}\,\xi \rangle, \qquad \forall \mathbf{x} \in \mathcal{B},$$

where $\langle \cdot, \cdot \rangle$ denotes the inner product in $\mathbb{R}^n$. Hence we get the sufficiency in the same way again. For the necessity we only need to replace the values x and y in (14) with vectors from $\mathcal{B}$ such that an appropriate convex combination of them equals $\mathbb{E}\,\xi$.

Note that the above definition of fairness of a gamble is made on the base of a particular preference system. However the utility plays no role in Definition 4.1. In case of several goods the optimal allocations with respect to different preference systems might vary, even if they correspond to the same income. Therefore, denoting by P_1 and P_2 the fair prices of a gamble ξ with respect to utility functions U_1 and U_2 respectively, P_i is the market price of basket $\mathbf{x}_{P_i}^{i,opt}$, where $\mathbf{x}_P^{i,opt}$ denotes the optimal allocation for income P with respect to U_i $(i = 1, 2)$. Then $U_i(\mathbb{E}\,\xi) = U_i\left(\mathbf{x}_{P_i}^{i,opt}\right)$ $(i = 1, 2)$, but the relation between $U_1\left(\mathbf{x}_{P_1}^{1,opt}\right)$ and $U_2\left(\mathbf{x}_{P_2}^{2,opt}\right)$ can be anything, as well as between P_1 and P_2.

To get rid of the dependence on the utility one could define the fair price of ξ by $P = \sum_{i=1}^n p_i \mathbb{E}\,\xi_i$ which is the market price of $\mathbb{E}\,\xi$. The difference in the two definitions can be explained as follows. In the first we fix the utility of $\mathbb{E}\,\xi$ and offer the individual the exact amount of money by which the fixed utility can be attained using utility maximization. In the second definition, however, the market value of $\mathbb{E}\,\xi$ is offered the individual for the gamble from which the basket $\mathbb{E}\,\xi$ can be purchased, obviously. These are equivalent only if there is only one good, like in Definition 4.1, or if $\mathbb{E}\,\xi$ is an optimum, that is $\mathbf{x}_V^{opt} = \mathbb{E}\,\xi$ where $V = \sum_{i=1}^n p_i \mathbb{E}\,\xi_i$. However, 'in large' both imply the equivalence of the individual's (global) risk aversion with the (global) concavity of his utility function.

The first answer could be called Hicks-type whereas the latter one Slutsky-type, since in many microeconomic problems it is possible to achieve two different answers by deciding whether to give the consumer, who is having a certain basket, the same purchasing or welfare conditions (i.e. to make the same utility level available for him) or to make the same basket available for him. For our purposes the first alternative seems suitable since the individual is supposed to make his purchasing decisions with the aid of utility.

4.2 Measure of Risk Aversion

Now consider a gamble ξ with $\mathbb{E}\,\xi = P$. Suppose that the individual has a strictly monotone increasing utility function U which is strictly concave at P. Thus he does not accept the fair gamble (ξ, P). By the monotonicity of U there exists a unique positive value P^* such that

$$U(P - P^*) = \mathbb{E}\,U(\xi). \tag{15}$$

Any price in $(P - P^*, \infty)$ would lead the individual not to accept the risky asset ξ, but he would undertake it for less than $P - P^*$. To put it in another way, he is ready to reduce his wealth P by P^* at most in order to avoid the risk of holding ξ. One can call P^* the insurance premium which can be a possible measure of the individual's risk aversion at point (wealth level) P with respect to gamble ξ. The higher the insurance premium is the more the individual is risk averse if (ξ, P) is offered to him. It is trivial that P^* depends on ξ.

Theorem 4.2 explains why people are willing to use insurance services. Having in mind the proof of this theorem and recalling the notations used in it, suppose now that U is convex in a subdomain J. Then it is trivial that any gamble ξ with property $\mathbb{P}(\xi \in J) = 1$ will be accepted by the individual. Thus an individual, for instance having a utility function with convex and concave parts, might refuse some risk but take another. This is a possible explanation of the fact that some people are ready to pay for insurance but on the other hand they also buy lottery tickets regularly.

4.5 Definition. *Let U be a strictly concave, increasing utility function on interval I. Sup-*

pose that ξ *is a random variable having values in* I *with* $\mathbb{E}\,|\xi| < \infty$. *Then* $P(\xi) \in \mathbb{R}$ *is called the insurance premium for gamble* ξ *if* $U(\mathbb{E}\,\xi - P(\xi)) = \mathbb{E}\,U(\xi)$.

4.6 Definition. *Given a twice differentiable utility function* $U : I \mapsto \mathbb{R}$ *we will call*

$$R(P) = -\frac{U''(P)}{U'(P)}, \qquad P \in I,$$

the (relative) risk aversion (at P*) of the individual whose utility function is* U. *The value* $R_A(P) = R(P) \cdot P$ $(P \in I)$ *is said to be the absolute risk aversion at* P.

This measure of risk aversion has been introduced by Arrow and Pratt. It has got two important features which makes it easy to use.

First, it is invariant to the scaling of U. To see this recall that the Neumann-Morgenstern utility function is determined only to the extend of an affine transformation. But having $\tilde{U}(P) = aU(P) + b$ $(a > 0,\, b \in \mathbb{R})$,

$$U''(P)/U'(P) = \tilde{U}''(P)/\tilde{U}'(P), \qquad P \in I,$$

and therefore the absolute risk aversion is uniquely defined at any wealth level (P) in case of a given preference system.

Secondly, it is also clear from the definition that $R(P)$ is derived only from the utility function and it does not depend on the gamble.

However, we can find some approximate relation between the insurance risk premium (P^*) and the absolute risk aversion which explains the choice of Arrow and Pratt. For this, suppose that $U \in \mathcal{C}^3(I)$ and take Taylor expansion of U around P. Then we get

$$\mathbb{E}\,U(\xi) = U(P - P^*) = U(P) - U'(P)P^* + \frac{U''(\tilde{P})}{2}P^{*2} \tag{16}$$

with an appropriate $\tilde{P} \in (P - P^*, P)$ and

$$U(\xi(\omega)) = U(P) + U'(P)(\xi(\omega) - P) + \frac{U''(P)}{2}(\xi(\omega) - P)^2 + \frac{U'''(\bar{P}(\omega))}{6}(\xi(\omega) - P)^3 \tag{17}$$

for all $\omega \in \Omega$, where ξ is defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and $\bar{P}(\omega)$ is in the interval determined by the endpoints $\xi(\omega)$ and P for all $\omega \in \Omega$. Taking the expectation of (17) and combining it with (15) and (16) we have

$$-U'(P)P^* + \frac{U''(\tilde{P})}{2}P^{*2} = \frac{U''(P)}{2}\,Var\xi + \mathbb{E}\left(\frac{U'''(\bar{P})}{6}(\xi - P)^3\right). \tag{18}$$

In equation (18) the last terms of both hand sides are small compared with the other terms if ξ takes its values in a small neighbourhood of P. Hence we get the following approximation.

$$P^* \approx \frac{1}{2}\,R(P)\,Var\xi, \tag{19}$$

which shows us that the insurance premium can be written as a product of two factors. One factor is determined by individual's preferences and the other depends only upon the gamble. An increase of the risk aversion or of the variance of the gamble would both make the individual ready to pay more for avoiding the risk.

In (19) we gave only an approximation which is valid in a certain neighbourhood of P. The following theorem tells us more about the absolute risk aversion and explains why it can be suitable to characterize the risk aversion of the individual indeed.

4.7 Theorem. *Let $U_i : I \mapsto \mathbb{R}$ be a monotone increasing, strictly concave utility function for $i = 1, 2$ and let us suppose that $U_i \in \mathcal{C}^3(I)$ where I is an interval. Then the following three statements are equivalent.*

(1) *$R_1(x) > R_2(x)$ for all $x \in I$, where R_i $(i = 1, 2)$ denotes the relative risk aversion of an individual having utility function U_i,*

(2) *there exists a twice differentiable real-valued function G defined on $U_2(I)$ such that*

$$G'(x) > 0, \quad G''(x) < 0 \qquad x \in U_2(I)$$

and

$$U_1(x) = G\big(U_2(x)\big) \quad \text{holds for} \quad x \in U_2(I), \tag{20}$$

(3) *$P_1(\xi) > P_2(\xi)$ holds for any random variable ξ which has finite expectation and takes values from I, where $P_i(\xi)$ denotes the insurance premium with respect to U_i $(i = 1, 2)$.*

Proof.

(1) $\Longrightarrow$ **(2)** Due to the strict monotonicity of U_2

$$G(x) := U_1(U_2^{-1}(x)), \qquad x \in U_2(I),$$

is a well-defined, twice differentiable function on $U_2(I)$ and (20) holds for G. Taking the first two derivatives of (20) we obtain

$$G'\big(U_2(x)\big)U_2'(x) = U_1'(x), \tag{21}$$

$$G''\big(U_2(x)\big)\big(U_2'(x)\big)^2 + G'\big(U_2(x)\big)U_2''(x) = U_1''(x), \quad x \in I. \tag{22}$$

Now $G'(x) > 0$ $(x \in I)$ follows directly from (21). Furthermore, dividing (22) by (21) and changing its sign we obtain

$$\begin{aligned} R_1(x) &= \frac{-G''\big(U_2(x)\big)\big(U_2'(x)\big)^2 - G'\big(U_2(x)\big)U_2''(x)}{G'\big(U_2(x)\big)U_2'(x)} \\ &= R_2(x) - \frac{G''\big(U_2(x)\big)\big(U_2'(x)\big)}{G'\big(U_2(x)\big)}. \end{aligned} \tag{23}$$

Thus the quotient in the second line of (23) must be positive (because of (1)) which implies $G''(x) < 0$ for $x \in I$.

(2) $\Longrightarrow$ **(3)** Now suppose that ξ is a random variable taking values from I such that $\mathbb{E}\ |\xi| < \infty$. Let $W = \mathbb{E}\ \xi$ and note that the concavity of G follows from **(2)**. Then by Jensen's inequality

$$\begin{aligned} U_1\big(\mathbb{E}\,\xi - P_1(\xi)\big) &= \mathbb{E}\,U_1(\xi) = \mathbb{E}\,G\big(U_2(\xi)\big) < G\big(\mathbb{E}\,U_2(\xi)\big) \\ &= G\Big(U_2\big(\mathbb{E}\,\xi - P_2(\xi)\big)\Big) = U_1\big(\mathbb{E}\,\xi - P_2(\xi)\big), \end{aligned}$$

and hence $P_2(\xi) < P_1(\xi)$.

(3) $\implies$ (1) Now let $P \in I$, $\varepsilon > 0$ and ξ be a random variable such that $\mathbb{P}(|\xi - P| = \varepsilon) = 1$ and $\mathbb{E}\,\xi = P$. Using Taylor expansion for U_1 and U_2 at P as in (16) and (17) we can rewrite (18) as follows.

$$-U_i'(P)P_i(\xi) + \frac{U_i''(\tilde{P}_i)}{2}\,P_i(\xi)^2 = \frac{U''(P)}{2}\,Var\xi + \mathbb{E}\,\frac{U_i'''(\bar{P}_i)}{6}(\xi - P)^3, \qquad \text{for} \quad i = 1, 2, \tag{24}$$

where $|\tilde{P}_i - P| \le P_i(\xi)$ and $\bar{P}_i$ $(i = 1, 2)$ is a random variable such that $\mathbb{P}(|\bar{P}_i - P| \le \varepsilon) = 1$. Writing (24) for $i = 1$ and for $i = 2$ and combining them we obtain

$$\begin{aligned} P_1(\xi) - P_2(\xi) &+ \frac{U_2''(\tilde{P}_2)}{U_2'(P)}P_2(\xi)^2 - \frac{U_1''(\tilde{P}_1)}{U_1'(P)}P_1(\xi)^2 \\ &= \frac{1}{2}\Big(R_1(P) - R_2(P)\Big)\varepsilon^2 + \frac{1}{U_2'(P)}\mathbb{E}\,\frac{U_2'''(\bar{P}_2)}{6}(\xi - P)^3 - \frac{1}{U_1'(P)}\mathbb{E}\,\frac{U_1'''(\bar{P}_1)}{6}(\xi - P)^3. \end{aligned}$$

If ε is small enough then

$$\begin{aligned} sgn(P_1(\xi) - P_2(\xi)) &= sgn\left(P_1(\xi) - P_2(\xi) + \frac{U_2''(\tilde{P}_2)}{U_2'(P)}P_2(\xi)^2 - \frac{U_1''(\tilde{P}_1)}{U_1'(P)}P_1(\xi)^2\right) \\ &= sgn\Bigg(\frac{1}{2}\left(R_1(P) - R_2(P)\right)\varepsilon^2 \\ &\qquad - \left(\frac{1}{U_1'(P)}\mathbb{E}\,\frac{U_1'''(\bar{P}_1)}{6}(\xi - P)^3 - \frac{1}{U_2'(P)}\mathbb{E}\,\frac{U_2'''(\bar{P}_2)}{6}(\xi - P)^3\right)\Bigg) \\ &= sgn\left(\frac{1}{2}\left(R_1(P) - R_2(P)\right)\varepsilon^2\right) = sgn\big(R_1(P) - R_2(P)\big). \end{aligned}$$

Hence $P_1(\xi) > P_2(\xi)$ implies $R_1(P) > R_2(P)$ and thus the proof is complete. □

4.3 Portfolio Selection

The problem of portfolio selection is a basic issue of microeconomics and particularly important in the theory of finance as well as in its application when some financial decisions are needed to be made.

To put it in general, consider an individual or any financial institution possessing a certain amount of capital which is to be invested. There are some financial assets available for the investor in the market, for instance currencies, several kinds of securities like bonds (e.g. treasury or zero coupon ones), stocks, futures, swaps, options. Then the task for the investor is clearly to allocate the capital among the assets and therefore one can immediately see the analogy of this problem to the choice problem studied in Section 3.4. We have again a constraint given by the initial capital to be spent but many sorts of other constraints can also occur such as short sales of stocks might be forbidden or just limited to a certain extend.

In this case, however, an important difference with the market prices of goods is the fact that we have financial assets the prices of which are known at the time when the portfolio is selected but these prices are supposed to be changing randomly after the selection (unlike the goods' prices, see Sections 3.4, 3.5). Therefore we will study a time period $[0, T]$ where 0 indicates the time of selection and T is a time point in the future, say terminal time. The asset prices change randomly over $[0, T]$ but we shall assume that the decision maker is given the law of these random changes. First we study a one period model, i.e. the simple case when the prices change only once after time 0, namely, at time T. More generally, one can consider a multiperiod model with more time points (called trading times) in $[0, T]$ when the assets are supposed to take new prices which are announced in the market. Or one can study the model in which the prices can be changing continuously over the observed time interval.

Having a certain set of feasible allocations what will be called portfolios, we should like to choose the optimal portfolio. The meaning of 'optimal' depends on our purposes here. The value of a portfolio at time 0 is certainly equal to the initial capital but random at the terminal time. Thus one could look for the portfolio which provides the largest expected utility value in the set of the feasible portfolios and call it optimal. Others might determine first a minimal level what the portfolio value should exceed at time T surely and then look for the optimal one of those as in the previous case. Another problem could be to minimize the variance of the portfolio value of those which have larger expected utility value than a certain level.

In the following we will deal with the problem of maximizing the expected utility of the portfolio. We will assume that there is an asset, indicated by index 0, which has non-random rate of return, say r_0. In other words, r_0 is a fixed interest rate which means that if we invest β_0 amount of money in this asset at time 0 then it will worth $\beta_0(1+r_0)$ surely at time T. Similarly, r_i will denote the rate of return of the i-th asset which is, however, a random variable. To put it differently, r_i can be considered as a random interest rate and thus investing β_i in asset i will lead the investor to possess $\beta_i(1+r_i)$ at the terminal time. Let X_0 be the initial capital (to be invested) and β_i $(i = 1, \dots, n)$ denote the money invested in the i-th asset. We have a portfolio $\pi = (\beta_0, \beta_1, \dots, \beta_n)$ with value equal to $X_0 = \sum_{i=0}^n \beta_i$ at time 0 and

$$\begin{aligned} X_T^\pi = \sum_{i=0}^n \beta_i(1+r_i) &= \left(X_0 - \sum_{i=1}^n \beta_i\right)(1+r_0) + \sum_{i=1}^n \beta_i(1+r_i) \\ &= X_0(1+r_0) + \sum_{i=1}^n \beta_i(r_i - r_0) \end{aligned} \tag{25}$$

at time T.

4.8 Definition. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, $r_0 > 0$ and for $i = 1, \dots, n$ let $r_i : \Omega \mapsto (-1, \infty)$ be a random variable on the underlying probability space such that $\mathbb{E}\, r_i^2 < \infty$ and $\mathbb{P}(r_i = r_0) < 1$ and write $r = (r_0, r_1, \dots, r_n)$.*
Then the set $\{\Omega, \mathcal{F}, \mathbb{P}, r, n\}$ is said to be a securities market (with n risky assets).
An $n+1$ dimensional vector $\pi = (\beta_0, \beta_1, \dots, \beta_n)$, $\beta_i \in \mathbb{R}$, is called portfolio, where β_i indicates the amount of money invested in asset i.

The condition $\mathbb{P}(r_i = r_0) < 1$ does certainly not cause loss of generality since financial assets with rate of return satisfying $\mathbb{P}(r_i = r_0) = 1$ are indifferent to the riskless asset with interest rate r_0.

It is also realistic to assume (though it is not included in Definition 4.8) that the rate of return of a risky asset takes larger and smaller values than r_0, both with positive probability. Otherwise, $\mathbb{P}(r_i \geq r_0) = 1$ or $\mathbb{P}(r_i \leq r_0) = 1$ would both lead the investor to realize as large profit as he wants by getting a loan on asset 0 and putting that money in asset i in the first case or, in the second case, by acting exactly in the opposite way in the market as in the first.

Such an opportunity of gaining profit without any risk and with no need of initial capital is called arbitrage. Similarly, the condition $\mathbb{P}(r_i \leq r_j) < 1$, for $i \neq j$, $i,j \in \{1,\ldots,n\}$, can be also required in our problems.

4.9 Notation. *In the following T will denote the terminal date when the returns on the assets are realized and X_T^π denotes the value of portfolio π at T as we already used this notation in (25). Furthermore we define*

$$C_{X_0} = \left\{ \pi \;\middle|\; \pi = (\beta_0, \beta_1, \ldots, \beta_n) \in \mathbb{R}^n, \quad \sum_{i=0}^{n} \beta_i = X_0 \right\}$$

which is called the set of feasible portfolios corresponding to initial capital X_0 if the trading with the assets is not limited and hence the β_i's can take any real value.

4.10 Lemma. *Let $U \in \mathcal{C}^2(\mathbb{R})$ be a strictly concave utility function, $X_0 > 0$ and assume that $\{\Omega, \mathcal{F}, \mathbb{P}, r, n\}$ is a securities market.*
Then π^* is an optimal portfolio meaning that

$$\mathbb{E}\, U\left(X_T^{\pi^*}\right) = \max_{\pi \in C_{X_0}} \mathbb{E}\, U\left(X_T^\pi\right) \tag{26}$$

if and only if

$$\mathbb{E}\,\Big(U'\big(X_T^{\pi^*}\big)(r_i - r_0)\Big) = 0 \quad \textit{for} \quad i = 1, \ldots, n. \tag{27}$$

Furthermore, if $r_1, \ldots, r_n$ are independent then the optimal portfolio, if there is any, is unique.

Proof. Note that n of the entries of a portfolio π can be chosen freely as long as $\pi \in C_{X_0}$ and then the remaining entry is uniquely determined by them. Let $\pi = (\beta_0, \beta_1, \ldots, \beta_n) \in C_{X_0}$

and define

$$F(\beta_1, \ldots, \beta_n) := \mathbb{E}\, U\left(X_T^\pi\right) = \mathbb{E}\, U\left(X_0(1+r_0) + \sum_{i=1}^n \beta_i(r_i - r_0)\right). \tag{28}$$

Now we are looking for the maximum of F over $\mathbb{R}^n$. Since $U \in \mathcal{C}^2(\mathbb{R})$ the expectation and the differentiation can be interchanged in the first order condition for the maximum at $\beta = (\beta_1, \ldots, \beta_n)$ and we obtain

$$0 = \frac{\partial}{\partial \beta_i} F(\beta) = \mathbb{E}\, \frac{\partial}{\partial \beta_i} U\left(X_T^\pi\right) = \mathbb{E}\, \Big(U'\big(X_T^\pi\big)(r_i - r_0)\Big).$$

The function F is concave over $\mathbb{R}^n$ which can be verified as follows. Let $\beta \in \mathbb{R}^n$ and $z_i \in \mathbb{R}$ $(i = 1, \ldots, n)$ and define $\pi = (X_0 - \sum_{i=1}^n \beta_i, \beta_1, \ldots, \beta_n)$. Now we have

$$\begin{aligned}\sum_{i=1}^n \sum_{j=1}^n z_i z_j \frac{\partial^2 F(\beta)}{\partial \beta_i \partial \beta_j} &= \sum_{i=1}^n \sum_{j=1}^n z_i z_j \mathbb{E}\, U''\left(X_T^\pi\right)(r_i - r_0)(r_j - r_0) \\ &= \mathbb{E}\, U''\left(X_T^\pi\right)\left(\sum_{i=1}^n z_i(r_i - r_0)\right)^2 \leq 0\end{aligned} \tag{29}$$

since U'' is negative everywhere. The concavity of F follows from (29) which means that the first order necessary condition (27) for maxima derived above is sufficient as well.

Moreover, if there exists an optimum then (27) implies $\mathbb{P}(r_j - r_0 < 0) > 0$ and hence $\mathbb{P}(r_j - r_0 > 0) > 0$ is valid as well for $j = 1, \ldots, n$. Therefore it follows from the independence of the rates of return of the risky assets in the market that the left-hand side of inequality (29) is negative provided that $z_i \neq 0$ for some i in $\{1, \ldots, n\}$. This implies the strict concavity of F over $\mathbb{R}^n$ and thus the uniqueness of the optimum. □

4.4 Demand For Financial Asset

One could expect some relationship between the individual's risk aversion and the way he chooses an optimal portfolio since, as we have seen before, the risk aversion characterized the behaviour of the individual under uncertainty. There can be found such relationship, indeed. To see this we will mainly deal with the simple market where a riskless asset and only one risky financial asset are available.

In economics a good is said to be normal if the individual's demand for the good increases as his wealth (or income) increases whereas the demand in case of a so-called inferior good decreases as the income increases. The same properties can be defined for financial assets as follows.

4.11 Definition. *Suppose that U is a strictly concave, monotone increasing utility function and $\{\Omega, \mathcal{F}, \mathbb{P}, r, n\}$ is a securities market such that for each $X > 0$ there is an optimal portfolio $\pi(X) = (\beta_0(X), \ldots, \beta_n(X))$ in the sense of (26).*

Then asset i ($i \in \{0, \ldots, n\}$) is called normal if $\beta_i(X)$, the demand for the asset at wealth X, is a monotone increasing function of X whereas it is called inferior if β_i is monotone decreasing.
If β_i is differentiable at X and it does not vanish at X then

$$\varepsilon_i(X) = \frac{\frac{d\beta_i(X)}{dX}}{\frac{\beta_i(X)}{X}} = \frac{d\beta_i(X)}{dX} \frac{X}{\beta_i(X)}, \quad X > 0,$$

is said to be the wealth elasticity of the demand for asset i at wealth X.

Intuitively, the elasticity shows the change of the demand for the risky asset in terms of percentage if the wealth is changed by one percent. Therefore an elasticity value larger than 1 means that the relative proportion of the risky asset in the optimal portfolio will increase if the wealth increases whereas for $\varepsilon(X) \in [0, 1)$ this relative proportion decreases. The case $\varepsilon(X) = 1$ does not cause change in the proportion of the risky asset at issue. But negative elasticity means the decrease of the total demand for the risky asset as the wealth increases.

4.12 Theorem. *Let $U \in \mathcal{C}^2(\mathbb{R})$ be an increasing, strictly concave utility function of an individual possessing capital $X_0 > 0$ in a one-risky-asset securities market $\{\Omega, \mathcal{F}, \mathbb{P}, r, 1\}$. Suppose, furthermore, that $\pi^* = (\beta_0^*, \beta_1^*)$ is the optimal portfolio for X_0 in the sense of (26).*

Then

$$\beta_1^* > 0 \qquad \Longleftrightarrow \qquad \mathbb{E}\, r_1 > r_0$$

and similarly

$$\beta_1^* < 0 \qquad \Longleftrightarrow \qquad \mathbb{E}\, r_1 < r_0.$$

There exists an optimal portfolio for each $X > 0$ if either

$$\lim_{x\to\infty} U'(x) = 0 \qquad \textit{or} \qquad \lim_{x\to-\infty} U'(x) = \infty.$$

Proof. Recall the function F defined in (28) which is now a univariate function. Define the portfolio $\pi_0 = (X_0, 0)$. Then taking the derivative of F at 0 we have

$$F'(0) = \mathbb{E}\, U'\left(X_T^{\pi_0}\right)(r_1 - r_0) = U'\big(X_0(1 + r_0)\big)(\mathbb{E}\, r_1 - r_0). \tag{30}$$

Since F is strictly concave and $U' > 0$ everywhere, it is easy to see from (30) that the location of the maximum of F is larger than 0 if and only if $\mathbb{E}\, r_1 - r_0 > 0$ and smaller than zero if and only if $\mathbb{E}\, r_1 - r_0 < 0$.

In the proof of Lemma 4.10 we showed the strict concavity of F from which it is clear that F' is strictly monotone decreasing. By the monotone convergence theorem

$$\begin{aligned}
\lim_{\beta\to\infty} F'(\beta) &= \lim_{\beta\to\infty} \mathbb{E}\, U'\big(X_0(1 + r_0) + \beta(r_1 - r_0)\big)(r_1 - r_0) \\
&= \mathbb{E}\, \lim_{\beta\to\infty} \Big[U'\big(X_0(1 + r_0) + \beta(r_1 - r_0)\big)(r_1 - r_0)\Big]^+ \\
&\quad + \mathbb{E}\, \lim_{\beta\to\infty} \Big[U'\big(X_0(1 + r_0) + \beta(r_1 - r_0)\big)(r_1 - r_0)\Big]^- \qquad (31) \\
&= u_+ \mathbb{E}\, [r_1 - r_0]^+ + u_- \mathbb{E}\, [r_1 - r_0]^-,
\end{aligned}$$

where $u_+ = \lim_{x\to\infty} U'(x)$, $u_- = \lim_{x\to-\infty} U'(x)$ and $[\cdot]^+ = \max(0, \cdot)$, $[\cdot]^- = \min(0, \cdot)$. Note that $\mathbb{E}\, [r_1 - r_0]^+ > 0$ and $\mathbb{E}\, [r_1 - r_0]^- < 0$ and therefore it is clear that the right-hand side of (31) is negative or equal to $-\infty$ if either $u_+ = 0$ or $u_- = \infty$. With a similar argument,

$$\begin{aligned}
\lim_{\beta \to -\infty} F'(\beta) = \mathbb{E} \lim_{\beta \to -\infty} \Big[U'\big(X_0(1+r_0) + \beta(r_1 - r_0)\big)(r_1 - r_0)\Big]^+ \\
+ \mathbb{E} \lim_{\beta \to -\infty} \Big[U'\big(X_0(1+r_0) + \beta(r_1 - r_0)\big)(r_1 - r_0)\Big]^- \qquad (32) \\
= u_- \mathbb{E}\,[r_1 - r_0]^+ + u_+ \mathbb{E}\,[r_1 - r_0]^- > 0,
\end{aligned}$$

if $u_+ = 0$ or $u_- = \infty$ holds. Finally Darboux theorem provides a point β where F' vanishes and thus (27) is achieved. $\square$

4.13 Theorem. *Let us assume that $U \in \mathcal{C}^2(\mathbb{R})$ is a monotone increasing, strictly concave utility function and $\{\Omega, \mathcal{F}, \mathbb{P}, r, 1\}$ is a one-risky-asset securities market where the risky asset's risk premium ($\mathbb{E}\, r_1 - r_0$) is positive. Suppose that the optimal portfolio $\pi(X)$ corresponding to capital X exists for any $X > 0$.*

If R, the risk aversion function of an individual having utility function U, is strictly decreasing over $\mathbb{R}$ then the risky asset is normal (for the individual).
If R is strictly increasing over $\mathbb{R}$ then the risky asset is inferior.
A constant relative risk aversion on $\mathbb{R}$ implies constant demand for the risky asset.

Proof. Define

$$f(X,\beta) = \mathbb{E}\,\Big(U'\big(X(1+r_0) + \beta(r_1 - r_0)\big)(r_1 - r_0)\Big), \qquad (X,\beta) \in (0,\infty) \times \mathbb{R}$$

and note that f has continuous partial derivatives, in particular

$$\frac{\partial f(X,\beta)}{\partial \beta} = \mathbb{E}\,\Big(U''\big(X(1+r_0) + \beta(r_1 - r_0)\big)(r_1 - r_0)^2\Big) < 0 \qquad \forall (X,\beta) \in (0,\infty) \times \mathbb{R}.$$

Applying the implicit function theorem (see Appendix, Theorem 8.2), there exist a function $\beta \in \mathcal{C}^1\big((0,\infty)\big)$ such that

$$f\big(X, \beta(X)\big) = 0 \qquad \text{for} \quad X \in (0,\infty)$$

and hence we obtain the following. For $Y > 0$

$$\begin{aligned}\frac{df\big(Y,\beta(Y)\big)}{dY} &= \mathbb{E}\, U''\left(X^{\pi(Y)}\right)\big((1+r_0)+\beta'(Y)(r_1-r_0)\big)(r_1-r_0)\\ &= \mathbb{E}\, U''\left(X^{\pi(Y)}\right)(1+r_0)(r_1-r_0)+\beta'(Y)\mathbb{E}\, U''\left(X^{\pi(Y)}\right)(r_1-r_0)^2 = 0\end{aligned}$$

Recalling $U''(x) < 0$, $x \in \mathbb{R}$, and $\mathbb{P}(r_1 \neq r_0) > 0$ we have

$$\frac{d\beta}{dY} = \frac{\mathbb{E}\, U''\left(X^{\pi(Y)}\right)(1+r_0)(r_1-r_0)}{-\mathbb{E}\, U''\left(X^{\pi(Y)}\right)(r_1-r_0)^2}, \qquad Y > 0, \tag{33}$$

and hence

$$\begin{aligned} sgn\left(\frac{d\beta}{dY}\right) &= sgn\left(\frac{\mathbb{E}\, U''\left(X^{\pi(Y)}\right)(1+r_0)(r_1-r_0)}{-\mathbb{E}\, U''\left(X^{\pi(Y)}\right)(r_1-r_0)^2}\right)\\ &= sgn\Big(\mathbb{E}\, U''\left(X^{\pi(Y)}\right)(r_1-r_0)\Big).\end{aligned} \tag{34}$$

Now suppose that R is strictly monotone decreasing. Then

$$R\big(Y(1+r_0)\big) \geq R\left(X^{\pi(Y)}\right) = R\big(Y(1+r_0)+\beta(Y)(r_1(\omega)-r_0)\big) \tag{35}$$

for each $\omega \in \{\ r_1 \geq r_0\ \}$ (for this note that $\beta(Y) > 0$ holds due to Theorem (4.12)) and

$$R\big(Y(1+r_0)\big) < R\left(X^{\pi(Y)}\right) = R\big(Y(1+r_0)+\beta(Y)(r_1(\omega)-r_0)\big) \tag{36}$$

for each $\omega \in \{\ r_1 < r_0\ \}$. Multiplying (35) and (36) by $-U'\left(X^{\pi(Y)}\right)(r_1(\omega)-r_0)$ and taking the expectation we finally obtain

$$\mathbb{E}\, U''\left(X^{\pi(Y)}\right)(r_1-r_0) > -R\big(Y(1+r_0)\big)\mathbb{E}\, U'\left(X^{\pi(Y)}\right)(r_1-r_0) = 0$$

which together with (34) implies the first statement.

The second statement can be proved in the same way whereas the third is an immediate corollary of the two previous ones. □

4.14 Theorem. *Let us suppose that the conditions of Theorem 4.13 hold.*

If R_A, the absolute risk aversion of the individual, is strictly increasing over $\mathbb{R}$ then the wealth elasticity is less than 1.
If R_A is strictly decreasing over $\mathbb{R}$ then the wealth elasticity is larger than 1.
Finally, constant absolute risk aversion make the elasticity equal 1.

Proof. Using formula (33) one can write the elasticity in the following form. (For convenience ε will be simply used to denote the elasticity of demand for the only risky asset of the market in this proof.)

$$\begin{aligned}\varepsilon(Y) &= \frac{d\beta(Y)}{dY}\frac{Y}{\beta(Y)} = 1 + \frac{\frac{d\beta(Y)}{dY}Y - \beta(Y)}{\beta(Y)} \\ &= 1 + \frac{\mathbb{E}\, U''\left(X^{\pi(Y)}\right)(1+r_0)(r_1-r_0)Y + \beta(Y)\mathbb{E}\, U''\left(X^{\pi(Y)}\right)(r_1-r_0)^2}{-\beta(Y)\mathbb{E}\, U''\left(X^{\pi(Y)}\right)(r_1-r_0)^2} \\ &= 1 + \frac{\mathbb{E}\, U''\left(X^{\pi(Y)}\right)(r_1-r_0)\left(X^{\pi(Y)}\right)}{-\beta(Y)\mathbb{E}\, U''\left(X^{\pi(Y)}\right)(r_1-r_0)^2}\end{aligned}$$

Therefore

$$sgn(\varepsilon(Y)-1) = sgn\Big(\mathbb{E}\, U''\left(X^{\pi(Y)}\right)(r_1-r_0)\left(X^{\pi(Y)}\right)\Big) \tag{37}$$

Now take the case when R_A is strictly monotone decreasing. Then

$$\begin{aligned}U''\left(X^{\pi(Y)}(\omega)\right)\left(X^{\pi(Y)}(\omega)\right)(r_1(\omega)-r_0) &= -R_A\Big(\left(X^{\pi(Y)}(\omega)\right)\Big)U'\left(X^{\pi(Y)}(\omega)\right)(r_1(\omega)-r_0) \\ &\le -R_A\big(Y(1+r_0)\big)U'\left(X^{\pi(Y)}(\omega)\right)(r_1(\omega)-r_0)\end{aligned}$$

holds if either $r_1(\omega) \ge r_0$ or $r_1(\omega) < r_0$ and we know that $\mathbb{P}(r_1(\omega) < r_0) > 0$. Hence we conclude with

$$\mathbb{E}\, U''\left(X^{\pi(Y)}\right)\left(X^{\pi(Y)}\right)(r_1-r_0) < -R_A\big(Y(1+r_0)\big)\mathbb{E}\, U'\left(X^{\pi(Y)}\right)(r_1-r_0) = 0$$

which together with (37) completes the proof of the first statement.

Clearly, the proof of the other two statements are analogous. □

5 Stochastic Dominance

So far we have seen some useful notions to characterize the individual's preferences under uncertainty and, in particular, when he is acting in a financial market. There occurs naturally the question of finding some features of the financial assets in order to compare their riskiness. Such a feature of the assets can be especially useful if it has some consequence about the market behaviour of a (large) class of individuals regarding to these assets. The classes of our further studies are the class of risk averse individuals and the class of individuals who prefer more to less.

For this investigation, we will use the concept of first and second order stochastic dominance.

5.1 Definition. *Let $\{\Omega, \mathcal{F}, \mathbb{P}, r, n\}$ be a securities market and suppose that the individual has a continuous utility function U. Let $i, j \in \{1, 2, \ldots, n\}$.*

Then we say that the individual prefers asset i to asset j if

$$\mathbb{E}\, U(1 + r_i) \geq \mathbb{E}\, U(1 + r_j).$$

Asset i dominates asset j in the sense of first order stochastic dominance (FSD) if all individuals having monotone increasing and continuous utility function (i.e. who prefer more to less) prefer asset i to asset j. This relation will be denoted by $r_i \succcurlyeq_{FSD} r_j$ or $r_j \preccurlyeq_{FSD} r_i$.

Whereas, the second order stochastic dominance (SSD) of asset i over asset j is said to be satisfied, which will be denoted by $r_i \succcurlyeq_{SSD} r_j$ or $r_j \preccurlyeq_{SSD} r_i$, if all risk averse individuals having utility function (i.e. a concave one) in $\mathcal{C}(\mathbb{R})$ prefer asset i to asset j.

Note that an asset having rate of return r_i is defined to be preferred to another asset of rate of return r_j with respect to a certain utility function U if $\mathbb{E}\, U(1+r_i) \geq \mathbb{E}\, U(1+r_j)$. In other words, the investment of one unit of the currency provides larger expected utility with asset i that with asset j. In case of stochastic dominance, however, the size of the capital (to be invested) does not play any role since the definition requires a feature of the asset to be valid for a whole class of utility functions. Therefore it is easy to see that defining the individual's preference of asset i to asset j for instance by $\mathbb{E}\, U(r_i) \geq \mathbb{E}\, U(r_j)$ instead of $\mathbb{E}\, U(1+r_i) \geq \mathbb{E}\, U(1+r_j)$ would lead us the same notion of stochastic dominance (either of type first order or of type second order).

5.1 First Order Stochastic Dominance

Here we give a theorem to describe the notion of FSD.

5.2 Theorem. *Let us assume that* $\{\Omega, \mathcal{F}, \mathbb{P}, r, n\}$ *is a securities market,* $i, j \in \{1, 2, \ldots, n\}$, *such that* $\mathbb{P}(r_i < u) = \mathbb{P}(r_j < u) = 1$ *for some real number* u.

Then

$$r_i \succcurlyeq_{FSD} r_j \qquad \textit{iff} \qquad F_i(x) \leq F_j(x) \quad \textit{for} \quad x \in \mathbb{R},$$

where F_k *denotes the distribution function of* r_k, $k \in \{1, 2, \ldots, n\}$.

Proof. Define

$$G(z) = F_i(z) - F_j(z) \qquad \text{for} \quad z \in \mathbb{R},$$

and let $U \in \mathcal{C}(\mathbb{R})$ be increasing. Then by the continuity of U the formula of integration by parts (see Appendix, Theorem 8.3) yields

$$\begin{aligned} &\int_{[-1,u]} U(1+z)dG(z) + \int_{[-1,u]} G(z)dU(1+z) \\ &= U(1+u)G(u) - U(0)G(-1) = 0, \end{aligned} \tag{38}$$

since $G(u) = G(-1) = 0$. Hence $r_i \succcurlyeq_{FSD} r_j$, that is

$$\mathbb{E}\, U(1+r_i) - \mathbb{E}\, U(1+r_j) = \int_{[-1,u]} U(1+z)dG(z) \geq 0$$

holds for all monotone increasing $U \in \mathcal{C}(\mathbb{R})$, if and only if

$$\int_{[-1,u]} G(z) dU(1+z) \leq 0$$

holds for all monotone increasing $U \in \mathcal{C}(\mathbb{R})$ which is equivalent to $G(z) \geq 0$ a.s. for $z \in \mathbb{R}$ and thus the proof is complete. □

Now it is clear that an asset displays first order stochastic dominance over another asset if the probability for the rate of return to exceed any level is not less in case of the first asset than in case of the second one. Hence, it implies that an asset dominating another one in the sense of FSD has at least as large expected rate of return than the one being dominated.

But the reverse of this statement does not hold of course. Imagine for example asset 1 with a uniformly distributed rate of return over the interval $[-0.5, 0.5]$ and let the rate of return of asset 2 be uniformly distributed on $[-0.2, 0.4]$. By Theorem 5.2 it is trivial that none of the assets displays stochastic dominance over the other one in the sense of FSD, although their expected rates of return are not equal.

It can also be mentioned that reflexivity and transitivity are both satisfied by the relation $\preccurlyeq_{FSD}$ though this relation does not necessary provides a linear ordering on the set of financial assets of a certain market, since some assets might not be comparable in this way, as it has been shown in the last example for instance.

5.2 Second Order Stochastic Dominance

The second order stochastic dominance, similarly to the first order one, can be also described by some basic properties of the distributions of the rates of returns, which makes it easy to verify whether an arbitrary pair of financial assets are in the relation or not. The following theorem provides the precise statement.

5.3 Theorem. *Let us assume that the conditions of Theorem 5.2 are valid and recall the notations used in it.*

Then

$$r_i \succcurlyeq_{SSD} r_j \qquad \Longleftrightarrow \qquad \mathbb{E}\, r_i = \mathbb{E}\, r_j \quad \text{and} \quad S(x) \le 0 \quad \text{for all} \quad x \in [-1, u],$$

where

$$S(x) = \int_{[-1,x]} F_i(z) - F_j(z) dz, \qquad x \in [-1, u].$$

Proof. Let U belong to $\mathcal{C}^1(\mathbb{R})$. Note first that clearly $S(-1) = 0$ and

$$S(u) = \int_{[-1,u]} (F_i - F_j)(z) dz = -\int_{[-1,u]} z d(F_i - F_j)(z) = \mathbb{E}\, r_j - \mathbb{E}\, r_i.$$

Then recalling formula 38 and applying the formula of integration by parts again (see Theorem 8.3 in Appendix) we find (with $G(z) = F_i(z) - F_j(z)$, $z \in \mathbb{R}$)

$$\begin{aligned} \mathbb{E}\, U(1+r_i) - \mathbb{E}\, U(1+r_j) &= \int_{[-1,u]} U(1+z) dG(z) = -\int_{[-1,u]} G(z) dU(1+z) \\ &= -\int_{[-1,u]} G(z) U'(1+z) dz = -\int_{[-1,u]} U'(1+z) dS(z) \\ &= -U'(1+u)S(u) + U'(0)S(-1) + \int_{[-1,u]} S(z) dU'(1+z) \\ &= U'(1+u)\big(\mathbb{E}\, r_i - \mathbb{E}\, r_j\big) + \int_{[-1,u]} S(z) dU'(1+z). \end{aligned} \tag{39}$$

Necessity. If $r_i \succcurlyeq_{SSD} r_j$ then the left-hand side of (39) is non-negative for each concave utility function U in $\mathcal{C}^1(\mathbb{R})$. The integral in the last line of (39), however, vanishes for linear utility function. In particular, we have $\mathbb{E}\, r_i - \mathbb{E}\, r_j \ge 0$ with $U(x) = x$ and $\mathbb{E}\, r_i - \mathbb{E}\, r_j \le 0$ with $U(x) = -x$ which together imply the equality of the expected rates of return of the underlying financial assets.

To show that the function S does not exceed zero over $[-1, u]$ let us suppose that there is a point in $[-1, u]$ where S takes a positive value. The function S is defined to be continuous and hence there can be found an interval $[a, b] \subset [-1, u]$ which S is positive over. Now, taking for instance the function

$$U(x) = \mathbf{1}_{(-\infty,a)}\, 2|a|x - \mathbf{1}_{[a,b]}\, x^2 - \mathbf{1}_{(b,\infty)}\, 2|b|x, \qquad \forall x \in \mathbb{R},$$

which is clearly a concave function in $\mathcal{C}^1(\mathbb{R})$, leads us to

$$\int_{[-1,u]} S(z)dU'(1+z) = \int_{[a,b]} S(z)dU'(1+z) < 0.$$

This relation is a contradiction to (39) and therefore the necessity is proved.

Sufficiency. If $U \in \mathcal{C}^1(\mathbb{R})$ is concave then U' is monotone decreasing. Therefore

$$\int_{[-1,u]} S(z)dU'(1+z) \geq 0$$

for such a utility function which together with (39) completes the proof. □

We have seen earlier that the class of individuals who prefer more to less is uniform in the sense that all will prefer an asset, say 1, to another, say 2, if the first asset displays first order stochastic dominance over the second one.

Another important class of individuals (or the corresponding class of utility functions) has turned to be uniform in a similar sense by Theorem 5.3. Namely, if the first asset dominates the second one this time in the sense of SSD then we can unambiguously say that any risk averse individual prefer asset 1 to asset 2.

Note that reflexivity and transitivity hold also for the relation $\succcurlyeq_{SSD}$ but the linearity is not satisfied either in this case.

5.4 Remark. It easily follows from Definition 5.1 and from Theorem 5.3 that the properties

$$\mathbb{E}\, r_i = \mathbb{E}\, r_j \qquad \text{and} \qquad Var\ r_i \leq Var\ r_j \tag{40}$$

are involved in the relation $r_i \preccurlyeq_{SSD} r_j$, where r_i and r_j are the rates of return of two financial assets in a securities market. For this, choose the concave utility function $U(x) = (x-\mu-1)^2$,

$x \in \mathbb{R}$, with $\mu = \mathbb{E}\ r_i = \mathbb{E}\ r_j$ which immediately yields the above relation between the variances.

Properties (40) make it clear why an asset is sometimes said to be more risky than another one by some authors instead of saying that the latter one dominates the first asset in the sense of SSD.

The two properties in (40), however, do not provide sufficient condition for the second order stochastic dominance. Indeed, take for instance a securities market $\{\Omega, \mathcal{F}, \mathbb{P}, r, 2\}$ where the risky assets are defined as follows. Let r_1 be uniformly distributed on $[-a, a]$ with $a \in (0, 1)$ whereas $r_2(\omega) \in \{-a, 0, a\}$ for each $\omega \in \Omega$ such that

$$\mathbb{P}(r_2 = -a) = \mathbb{P}(r_2 = a) = \varepsilon$$

and

$$\mathbb{P}(r_2 = 0) = 1 - 2\varepsilon \qquad \text{with} \quad 0 < \varepsilon < \frac{1}{6}\ .$$

It is easy to see that $\mathbb{E}\ r_1 = \mathbb{E}\ r_2 = 0$ and $Var\ r_1 > Var\ r_2$. On the other hand, there can be found a right neighbourhood of $-a$ (e.g. $(-a, -a + 2a\varepsilon)$) where the distribution function of r_2, say F_2, exceeds the distribution function of r_1, say F_1, which means that

$$S(y) = \int_{[-1,y]} F_2(x) - F_1(x) dx > 0$$

if y belongs to this neighbourhood. Hence, it follows from Theorem 5.3 that an asset with rate of return r_2 cannot dominate an asset with rate of return r_1 in the sense of SSD.

5.3 Demand Versus Stochastic Dominance

Consider two risky financial assets such that asset 1 displays stochastic dominance over asset 2. If the dominance is first order type then all individuals who prefer more to less will prefer any amount of money invested in asset one to the same amount invested in the second asset if these investments are made at the same time (namely, it was time point 0 in Section 4.3). Whereas in case of second order stochastic dominance risk averse people will prefer an investment in the first asset to an investment in the second one if these investments are worth the same at time 0.

One would guess naturally that the above assertions should imply a clear relation between the demand of two different financial assets where one is stochastically dominating the other one. One could expect, for instance, the dominant asset to have a larger demand (for any level of wealth to be invested).

Such a relationship cannot be established in general. In the following we will present examples contradicting the relation suggested above. We will see that studying securities markets, where there is at least one more asset traded in the market besides the investigated risky asset (e.g. one riskless is available for sure), the situation might be more complicated than in the previous sections. However, one can find some additional necessary conditions under what the relation at issue becomes valid, as we will show in Remark 5.6.

5.5 Example. Consider two one-risky-asset securities markets such that the riskless asset, say bond, is the same in these markets. The bond's interest rate is $r_0 > 0$. The second (risky) assets of the markets, say stocks, are defined by their random rate of return, what will be denoted by r_1 and r_2 respectively. Define

$$\mathbb{P}(r_1 = a) = \mathbb{P}(r_1 = b_1) = \frac{1}{2}$$

and

$$\mathbb{P}(r_2 = a) = \mathbb{P}(r_2 = b_2) = \frac{1}{2}\ ,$$

where for ease of calculations we assume that

$$a - r_0 = -\frac{1}{10}, \qquad b_1 - r_0 = 1 \qquad \text{and} \qquad b_2 - r_0 = 1 - \varepsilon$$

with $0 < \varepsilon < 9/11$. It is trivial that first order stochastic dominance of stock 1 over stock 2 holds.

Now consider the portfolio problem of the individual acting in both markets with initial capital X_0 (see Section 4.3). Let $\beta > 0$ and suppose that the individual's utility function U belongs to $\mathcal{C}^2(\mathbb{R})$ and satisfies

$$U'\big(g(a)\big) = 10\, U'\big(g(b_1)\big) \quad \text{and} \quad U'\big(g(b_2)\big) = (10-\varepsilon)\, U'\big(g(b_1)\big)$$

such that $U'\big(g(b_1)\big) > 0$ where the function g is defined by

$$g(x) = X_0(1+r_0) + \beta x \qquad \text{for} \quad x \in \mathbb{R}.$$

Note that $g(a) < g(b_2) < g(b_1)$ and hence the above constraints made on U still allow us to suppose that U is strictly concave and and monotone increasing. Moreover, either $\lim_{x\to\infty} U'(x) = 0$ or $\lim_{x\to-\infty} U'(x) = \infty$ can be fulfilled, which will be assumed as well.

In the above set up Lemma 4.10 and Theorem 4.12 assures us the existence and uniqueness of the optimal portfolio in both markets. Furthermore, checking the first order conditions (27) we find

$$\mathbb{E}\, U'\big(X_0(1+r_0)+\beta(r_1-r_0)\big)(r_1-r_0) = \frac{1}{2}\Big[U'\big(g(a)\big)(a-r_0) + U'\big(g(b_1)\big)(b_1-r_0)\Big] = 0$$

which means that $\pi = (X_0-\beta, \beta)$ is the solution of the portfolio problem in market 1. In market 2 we have

$$\begin{aligned}
\mathbb{E}\, U'\big(X_0(1+r_0)+\beta(r_2-r_0)\big)(r_2-r_0) &= \frac{1}{2}\Big[U'\big(g(a)\big)(a-r_0) + U'\big(g(b_2)\big)(b_2-r_0)\Big] \\
&= \frac{1}{2}\Big[U'\big(g(a)\big)(a-r_0) + (10-\varepsilon)\, U'\big(g(b_1)\big)(b_1-r_0-\varepsilon)\Big] \\
&= \frac{1}{2}\Big[U'\big(g(a)\big)(a-r_0) + U'\big(g(b_1)\big)(b_1-r_0)\Big] + \frac{1}{2}U'\big(g(b_1)\big)(9+\varepsilon^2-11\varepsilon) \\
&> U'\big(g(b_1)\big)\frac{9-11\varepsilon}{2} > 0.
\end{aligned} \tag{41}$$

It follows now from (41) that the function

$$F(x) = \mathbb{E}\, U\big(X_0(1+r_0)+x(r_2-r_0)\big), \qquad x \in \mathbb{R},$$

will achieve its maximum at a point β^* larger than β (see the proof of Lemma 4.10).

We have shown that there can be found a risk averse individual who prefers more to less such that he will undertake more of stock 2 in case of acting in market 2 than of stock 1 in case of acting in market 1 in spite of the fact that stock 2 is dominated by stock 1 in the sense of FSD (which also implies that any amount of stock 1 would be preferred to stock 2 being worth the same since the utility function of the individual is monotone increasing).

5.6 Remark. Let us suppose again that we are given two one-risky-asset securities market with the same bond representing the riskless asset in both markets and denote the interest rate of this common bond by $r_0 > 0$. As in Example 5.5, we will use the notations r_1 and r_2 to indicate the rates of returns of stock 1 (in market 1) and stock 2 (in market 2) respectively. Let $\mathbb{E}\, r_1 > r_0$ and $\mathbb{E}\, r_2 > r_0$.

We assume that stock 1 displays second order stochastic dominance over stock 2, in other words, stock 2 is more risky than stock 1.

Now consider the portfolio choice problem in both markets with a strictly risk averse individual having utility function $U \in \mathcal{C}^2(\mathbb{R})$ and possessing initial capital X_0. Furthermore, we assume that U has a form which provides the existence and uniqueness of the optimal portfolio in both markets. (For this necessary conditions are given in Theorem 4.12.)

If $(X_0 - \beta, \beta)$ (where β must be positive) is the optimal portfolio in market 1 then from the first order condition (27) we have

$$\mathbb{E}\, U'\big(X_0(1+r_0) + \beta(r_1 - r_0)\big)(r_1 - r_0) = 0.$$

Define

$$f(x) = U'\big(X_0(1+r_0) + \beta(x - r_0)\big)(x - r_0), \qquad x \in \mathbb{R}.$$

By Theorem 5.3 we can state the following about the demands of the stocks.

If function f, which is determined by the individual's utility, is concave over the real line then the second order stochastic dominance of stock 1 over stock 2 implies the fact that *the individual will invest more money in the less risky risky asset (in the first market) than in the more risky risky asset (in the second market).*

So, we have found a relationship, indeed, between the stochastic dominance and the demand for the risky asset what we were discussing at the beginning of this section. This assertion, however, is hardly applicable since the concavity of the above f does not hold in case of many commonly used types of utility functions.

5.7 Example. In the above remark we investigated what circumstances would lead the individual to invest less in the more risky asset than in the less risky one. We also mentioned that the conditions had been found sufficient are not always the case in practice. Therefore now we demonstrate an example where the investigated relation of the demand and the riskiness is reversed.

For this we keep the notations of Remark 5.6 and define the rates of returns of the assets and give a particular utility function.

So, first suppose that

$$\mathbb{P}(r_1 = a_0) = \mathbb{P}(r_1 = b) = \frac{1}{2}$$

where $a_0 = -0.5 + r_0$ and $b = 1 + r_0$. The second risky asset is defined by

$$\mathbb{P}(r_2 = a_1) = \mathbb{P}(r_2 = a_2) = \frac{1}{4} \qquad \text{and} \qquad \mathbb{P}(r_2 = b) = \frac{1}{2}$$

with $a_1 = -0.6 + r_0$ and $a_2 = -0.4 + r_0$.

We claim that stock 1 displays second order stochastic dominance over stock 2. To see this one can find

$$\mathbb{E}\, r_1 = r_0 + \frac{1}{4} = \mathbb{E}\, r_2, \tag{42}$$

furthermore,

$$sgn\big(S(x)\big) = sgn\left(\int_{[-1,1+r_0]} F_1(x) - F_2(x)dx\right) = \mathbf{1}_{(a_1+r_0,a_2+r_0)}. \tag{43}$$

Then (42) and (43) together with Theorem 5.3 imply $r_1 \succcurlyeq_{SSD} r_2$.

Let $\beta > 0$. Given capital $X_0 > 0$ recall the definition of function g from Example 5.5 and suppose that

$$0 < U'\big(g(a_0)\big) = 2U'\big(g(b)\big)$$

and

$$U'\big(g(a_1)\big) = U'\big(g(a_0)\big) + \varepsilon, \qquad U'\big(g(a_2)\big) = U'\big(g(b)\big) + \varepsilon$$

where $0 < \varepsilon < U'\big(g(a_0)\big)/5$. We can suppose that this utility function is strictly concave and monotone increasing satisfying either $\lim_{x\to\infty} U'(x) = 0$ or $\lim_{x\to-\infty} U'(x) = \infty$. We have

$$\begin{aligned} \mathbb{E}\, U'\big(X_0(1+r_0) &+ \beta(r_1 - r_0)\big)(r_1 - r_0) \\ &= \frac{1}{2}\, U'\big(g(a_0)\big)(a_0 - r_0) + \frac{1}{2}\, U'\big(g(b)\big)(b - r_0) = 0 \end{aligned} \tag{44}$$

and

$$\begin{aligned} &\mathbb{E}\, U'\big(X_0(1+r_0) + \beta(r_2 - r_0)\big)(r_2 - r_0) \\ &\quad = \frac{1}{4}\, U'\big(g(a_1)\big)(a_1 - r_0) + \frac{1}{4}\, U'\big(g(a_2)\big)(a_2 - r_0) + \frac{1}{2}\, U'\big(g(b)\big)(b - r_0) \\ &\quad = \frac{1}{4}\, \Big(U'\big(g(a_0)\big) + \varepsilon\Big)(-0.6) + \frac{1}{4}\, \left(\frac{1}{2}\, U'\big(g(a_0)\big) + \varepsilon\right)(-0.4) + \frac{1}{2}\, U'\big(g(b)\big)(b - r_0) \\ &\quad = \frac{U'\big(g(a_0)\big)}{20} + \frac{\varepsilon}{4} > 0. \end{aligned} \tag{45}$$

Using the same argument as in Example 5.5 one can obtain from (44) and (45) that the optimal portfolios in market 1 and market 2 are $(X_0 - \beta, \beta)$ and $(X_0 - \beta^*, \beta^*)$ respectively, with $\beta < \beta^*$.

Thus we have shown that a risk averse individual who prefers more to less may invest more in an asset than in another one even if the asset is more risky (in the sense of SSD) than the other one.

In Examples 5.5 and 5.7 we considered two securities markets. Both markets are equipped with two assets, a riskless bond and a risky stock such that the bonds of the markets are identical. Then we compared the stocks and established stochastic dominance of one of the stocks over the other one. Inspite of this fact, however, we have seen that some individuals would buy more of the more risky stock (the one which is worse in terms of stochastic dominance) even if they are risk averse and they prefer more to less.

One can interpret this assertion in different ways. First, we could imagine two markets in reality with such features and observe individuals who invest in both securities markets (certainly choosing optimal portfolios in both). But in fact this is not very realistic. Or, one can consider market 1 and market 2 as two different states of the same securities market at different time points, say, at t_1 and t_2 with $t_1 < t_2$. Then our assertions show that the stock of the market got more risky by time t_2 and then the individual restructured his portfolio according to the changes. But these do not necessarily imply that he took capital out of stock (by selling some of the stocks held by him) and bought bond for that capital. Moreover, he may increase his investment portion of stocks.

Although we should like to note that this setup of the two markets is not equivalent with a two-risky-asset securities market where all the three assets of the two markets are available, that is: the bond is the same as before and also stock 1 and stock 2 are both traded (on the same market). In that case the optimal portfolio would not contain more of the more risky asset even if considering an individual with utility function like in Examples 5.5, 5.7.

6 Risk measures

One of the main issues of modern finance is to develop tools and methods which enable us to compare financial assets and especially portfolios, and to describe their riskiness. We have seen in the previous chapters that several tools are provided by the different notions of stochastic dominance for the comparison, and it is known that the classical capital market models can also give a certain type of comparison of portfolios and a certain valuation of their riskiness.

Beyond the above mentioned tools it is also a natural requirement that simple financial indicators should describe the financial assets/portfolios and especially their riskiness. In what follows we will call the indicators that describes the riskiness of the financial assets and portfolios risk measures. Recall for instance the celebrated P/E indicator (price/earning), which gives the investor some information on the asset. However, the classical financial indicators (such as P/E) do not really give information on the riskiness of the financial assets.

Several risk measures have been introduced in the literature. Among these risk measures it is undoubtable that the Value at Risk has become the most widespreadly used both in practice and in theoretical works. According to many financial (and capital market) regulations (law) the financial institutions and possibly other market actors are required to calculate VaR and to fulfill other related operational requirements. We mention to the interested reader that the Basel Committee on Banking Supervision has deeply considered several problems related with risk measures and issued some standards that they recommend (see e.g. International Convergence of Capital Measurement and Capital Standards). These standards are commonly known as Basel I and Basel II, in which among others VaR and related issues play a crucial role, they are recommended to the countries.

At this point we should also notice that we cannot state that the VaR is the most appropriate measure for the purpose we have discussed above, furthermore, in order to find the most appropriate risk measure one first has to precisely fix his or her requirements that such an indicator should fulfill. In Section 6.1 we study such possible features one would require, whereas after this we consider the VaR and discuss its properties in Section 6.2. Finally in Section 6.3 we study another risk measure, the so-called expected shortfall, which has been suggested by many authors as an alternative of VaR because of the non-satisfactory features of the VaR.

6.1 Coherent risk measures

As we have already mentioned one can propose several requirements that a risk measure should satisfy. It is of course subjective, that is we cannot say that everyone would find the same properties equally important. In the literature many different properties have been proposed. Maybe the most commonly accepted properties in the literature are collected in the notion of coherency, which we study in this section. We also study some possible alternative properties as well.

The risk measures are defined on a set of random variables. Since given a financial asset or portfolio the future value of its profit is described by a random variable. Similarly, the value of the portfolio itself can also be described by the random variable of which the risk is to be measured.

6.1 Definition. *Let V be a set of random variables (describing the profit/value of the set of the corresponding portfolios or financial assets) over a probability space $(\Omega, \mathcal{F}, \mathbb{P})$.*

A function $\varrho : V \to \mathbb{R}$ is called risk measure.

A risk measure V is said to be

(1) *monotone, if for all $X, Y \in V$ and $\mathbb{P}(X \leq Y) = 1$ we have $\varrho(X) \geq \varrho(Y)$;*

(2) *subadditive, if for all* $X, Y, X+Y \in V$ *we have* $\varrho(X+Y) \leq \varrho(X) + \varrho(Y)$*;*

(3) *positive homogeneous, if for all* $h > 0$, $X, hX \in V$ *we have* $\varrho(hX) = h\varrho(X)$*;*

(4) *translation invariant, if for all* $a \in \mathbb{R}$, $X, X+a \in V$ *we have* $\varrho(X+a) = \varrho(X) - a$*.*

A risk measure is said to be coherent if it satisfies properties **(1)-(4)**.

Assumption. *For the sake of simplicity and convenience in this chapter we will assume that* V *is closed under addition, under multiplication by a positive scalar and under translation, that is* $V + V, h \cdot V, a + V \subset V$ *for all* $h > 0$, $a \in \mathbb{R}$*. We will assume, furthermore, that* V *contains the element* $X \equiv 0$ *(* $\mathbb{P}(X = 0) = 1$*).*

The properties of the risk measures (monotonicity, subadditivity, etc) will sometimes be also referred to as axioms.

Considering the motivation behind the properties mentioned in the definition we can interpret them as follows. If a portfolio promises more then another one in all situations then it should not have larger risk then that of the other one (monotonicity). Merging two portfolios should not increase the riskiness, i.e. we should not exceed the sum of the risks of the separate portfolios (subadditivity). If we multiply the portfolio value but we keep the relative proportions of the assets in it then the risk of the portfolio should change proportionally to the value of it (positive homogeneity). If we will realise an additional fixed (non-random) cash flow then the risk should be reduced by exactly the sum of the cash flow (translation invariance).

It is important to mention that one could consider several other properties that could also be reasonable from a financial point of view, they could possibly be similar to the ones mentioned, on the other hand, one could doubt the necessity of any of the properties contained by coherency. Next we mention two of the possible alternative properties.

6.2 Definition. *A risk measure* $\varrho : V \to \mathbb{R}$ *is said to be*

(5) positive, *if for all* $X \in V$, $\mathbb{P}(X \geq 0) = 1$ *we have* $\varrho(X) \leq 0$*;*

(6) convex, *if for all* $\lambda \in [0,1]$ *and* $X, Y, \lambda X + (1-\lambda)Y \in V$ *we have* $\varrho(\lambda X + (1-\lambda)Y) \leq \lambda\varrho(X) + (1-\lambda)\varrho(Y)$.

6.3 Theorem. *Let* ϱ *be a risk measure.*

(a) *If* ϱ *is positive homogeneous and* $X \equiv 0$ *then* $\varrho(X) = 0$.

(b) *If* ϱ *is positive homogeneous and translation invariant then* $\varrho(a) = -a$ *for all* $a \in \mathbb{R}$.

(c) *If* ϱ *is subadditive, positive homogeneous and translational invariant then* $-\varrho(X) = \varrho(-X)$ *for all* $X, -X \in V$.

(d) *A risk measure is coherent if and only if it satisfies properties* **(2)-(5)**.

(e) *If* ϱ *is subadditive and positive homogeneous then it is convex.*

Proof. **(a)** Let $X \equiv 0$. We have $X \in V$, thus the first statement is immediate due to positive homogeneity, since $\varrho(X) = \varrho(2X) = 2\varrho(X)$.

(b) If $a \in \mathbb{R}$ then the translation invariance and statement **(a)** together imply $\varrho(a) = \varrho(0 + a) = \varrho(0) - a = -a$.

(c) Applying subadditivity we obtain $0 = \varrho(0) \leq \varrho(-X) + \varrho(X)$. Let us suppose now indirectly that $\varrho(-X) + \varrho(X) = b > 0$. Then statement **(b)** and subadditivity together imply $\varrho(b) = \varrho(X + b - X) \leq \varrho(X + b) + \varrho(-X) = \varrho(X) - b + \varrho(-X) = 0$, from which we obtain that $b = 0$ which is clearly a contradiction.

(d) **(1)-(4)** $\Rightarrow$ **(2)-(5)**. If ϱ is coherent then for a nonnegative random variable X ($\mathbb{P}(X \geq 0) = 1$) the first statement and the monotonicity imply that $\varrho(X) \leq \varrho(0) = 0$, and hence the positivity is shown.

(2)-(5) $\Rightarrow$ **(1)-(4)**. Let $X, Y \in V$ such that $\mathbb{P}(X \leq Y) = 1$. Let us assume indirectly that $\varrho(Y) > \varrho(X)$. Notice that by statement **(c)** and the positivity we have $\varrho(X - Y) =$

$-\varrho(Y - X) \geq 0$. Furthermore, due to subadditivity $0 \leq \varrho(X - Y) \leq \varrho(X) + \varrho(-Y) < \varrho(Y)+\varrho(-Y)$, which contradicts to the fact that by statement **(c)** we have $\varrho(Y)+\varrho(-Y) = 0$.

(e) First applying subadditivity, next applying positive homogeneity one gets immediately the statement. □

Statement **(b)** means that the risk of a cash flow that comes with probability one is exactly the 'opposite' of its value. This means, that an almost sure fixed loss has positive risk and it is equal to the sum of the loss, whereas the risk of an investment that gives an almost sure positive fixed profit has negative risk such that its absolute value is equal to the profit. According to statement **(d)** we can see that monotonicity can be replaced in the definition of coherency by positivity (such that we obtain an equivalent system of 'axioms'). Finally we note that some authors propose a weaker set of axioms instead of coherency, namely, they suggest that in the four axioms of coherency the subadditivity and positive homogeneity should be replaced by the weaker axiom of convexity (see **(e)**).

In the past the riskiness of portfolios and financial assets were often described by the aid of the standard deviation or the variance of its future value. It is important, however, to note that the standard deviation or the variance are not appropriate for our purposes. It is trivial that none of them is monotone, or for instance positivity is not satisfied by any of them either.

6.2 Value at Risk

Notation. Given a random variable X we will denote its (left continuous) distribution function by F_X, i.e. $F_X(z) = \mathbb{P}(X < z)$. The right continuous version of the distribution function will be denoted by $\tilde{F}_X$, i.e. $\tilde{F}_X(z) = \mathbb{P}(X \leq z)$. We find it important to introduce both versions here because several authors use F_X, many others $\tilde{F}_X$ as the distribution function of X. Therefore in the followings we will always emphasise the differences caused by the two different definitions of the distribution function.

6.4 Definition. *Let X be a random variable and $\alpha \in (0,1)$. Then the lower α-quantile of X is defined by*

$$q_\alpha(X) = \sup\{y \mid F_X(y) < \alpha\},$$

where F_X is the distribution function of X, whereas

$$q^\alpha(X) = \inf\{y \mid F_X(y) > \alpha\}$$

will be called the upper α-quantile of X.

Let X be a random variable which represents the (random) profit of an asset or a portfolio on a probability space. Let $\alpha \in (0,1)$. The lower α-Value at Risk of X is

$$VaR_\alpha(X) = -q_\alpha(X).$$

Similarly, the upper α-Value at Risk of X is defined as

$$VaR^\alpha(X) = -q^\alpha(X).$$

The α-Value at Risk of X can be interpreted briefly as follows. It divides the possible future values of the portfolio profit into two sets: in α*100 % of the possible cases the profit will be less than the VaR, and with probability $(1-\alpha)$ we can say that the profit will be at least as much as the VaR. In other words, the α-VaR gives the value of loss that will be exceeded by the realised loss with probability $(1-\alpha)$. That is, the α-VaR shows the best of the worst α*100 % of the possible loss values, or to put in a different way, it shows the worst case of the best $(1-\alpha)$*100 % of the possible scenarios. However, it is important to emphasise that these decriptions are not rigorous statements. Furthermore, the upper VaR and lower VaR do not necessarily coincide. We will discuss these issues later in details.

Next we summarise some useful properties of quantiles, and afterwards we consider the consequent properties of VaR.

6.5 Remark. It is easy to see that the quantiles can be written in other forms as well, since

$$q_\alpha(X) = \inf\{y \mid F_X(y) \geq \alpha\}, \quad \text{whereas} \quad q^\alpha(X) = \sup\{y \mid F_X(y) \leq \alpha\}.$$

Furthermore, we note that in Definition 6.4 or in the equivalent forms given above one could replace the left continuous distribution function $y \mapsto F_X(y) = \mathbb{P}(X < y)$ by its right continuous counterpart $y \mapsto \tilde{F}_X(y) = \mathbb{P}(X \leq y)$, the value of $q_\alpha(X)$ and $q^\alpha(X)$ would not change.

We have $q^\alpha(X) = -q_{1-\alpha}(-X)$ and $q_\alpha(X) = -q^{1-\alpha}(-X)$. These statements are immediate taking into account the previous remark of ours, since considering for instance the first case one obtains that

$$\begin{aligned} q^\alpha(X) &= \inf\{y \mid F_X(y) > \alpha\} = \inf\{y \mid \mathbb{P}(-X \leq -y) < 1-\alpha\} \\ &= -\sup\{-y \mid \mathbb{P}(-X \leq -y) < 1-\alpha\} = -\sup\left\{y \mid \tilde{F}_{-X}(y) < 1-\alpha\right\} \\ &= -q_{1-\alpha}(-X). \end{aligned}$$

Since we clearly have that $\{y \mid F_X(y) > \alpha\} \subset \{y \mid F_X(y) \geq \alpha\}$, therefore taking the infimum of these sets we have

$$q_\alpha(X) \leq q^\alpha(X).$$

Based on this one can also observe that the lower and upper quantiles are not necessarily equal, namely, by the monotonicity of the distribution function we get that

$$q_\alpha(X) = q^\alpha(X) \iff \text{if the set } \{y \in \mathbb{R} \mid F_X(y) = \alpha\} \text{ has at most one element}$$

(since by the monotonicity of the distribution function we can see that the sets $\{y \mid F_X(y) > \alpha\}$ and $\{y \mid F_X(y) \geq \alpha\}$ forms intervals such that their right end-points are ∞). It is also trivial that the set $\{y \in \mathbb{R} \mid F_X(y) = \alpha\}$ has at most one element if and only if the set $\{y \in \mathbb{R} \mid \tilde{F}_X(y) = \alpha\}$ has at most one element.

This means that if the distribution function is constant over an interval and this constant value just equals α then the lower and upper quantile corresponding to this α are not the same. We cannot have such a situation in case of absolutely continuous distributions, however, taking for instance a discrete distribution the lower and upper quantiles are different for any $\alpha \in (0,1)$ which is an element of the image $F_X(\mathbb{R})$ of the distribution function. One can also say that the two quantiles determine (as end points) the interval where the distribution function takes the value α, more precisely: if $q_\alpha(X) < q^\alpha(X)$ then

$$\{x \in \mathbb{R} \mid F_X(x) = \alpha\} = \begin{cases} (q_\alpha(X), q^\alpha(X)], & \text{if } \mathbb{P}(X = q_\alpha(X)) > 0 \\ [q_\alpha(X), q^\alpha(X)], & \text{if } \mathbb{P}(X = q_\alpha(X)) = 0, \end{cases}$$

or to put it in a different way

$$\{x \in \mathbb{R} \mid \tilde{F}_X(x) = \alpha\} = \begin{cases} [q_\alpha(X), q^\alpha(X)), & \text{if } \mathbb{P}(X = q^\alpha(X)) > 0 \\ [q_\alpha(X), q^\alpha(X)], & \text{if } \mathbb{P}(X = q^\alpha(X)) = 0. \end{cases}$$

6.6 Theorem. *Let U be a uniformly distributed random variable over the interval $(0,1)$ and X be an arbitrary random variable. Then $\eta_1 = q_U(X)$, $\eta_2 = q^U(X)$ and X are equal in distribution (they have the same distribution).*

Given a random variable X it is usual to define the generalised inverse of its distribution function F_X by F_X^{-1} where

$$F_X^{-1}(y) := q_y(X), \qquad y \in (0,1),$$

which is clearly the inverse of the distribution function as long as the random variable is absolutely continuous. Thus according to the theorem above the distribution function of $F_X^{-1}(U)$ is F_X if U is uniformly distributed over the interval $(0,1)$.

The proof of Theorem 6.6. Since the U is absolutely continuous, $\eta_1 = q_U(X)$ and $\eta_2 = q^U(X)$ have the same distribution (because they only differ over a set of probability zero). We will show that

$$A_y := \{\omega \in \Omega \mid U(\omega) < F_X(y)\} \subset B_y := \left\{\omega \in \Omega \mid F_X^{-1}(U(\omega)) < y\right\} \tag{46}$$

and

$$B_y \subset C_y := \{\omega \in \Omega \mid U(\omega) \leq F_X(y)\} \tag{47}$$

for all $y \in \mathbb{R}$. These imply the required statement, for this notice that U has absolutely continuous distribution and $F_X(y) = F_U(F_X(y)) = \mathbb{P}(A_y) = \mathbb{P}(C_y)$, furthermore we clearly have $F_{\eta_1}(y) = \mathbb{P}(B_y)$.

Let us turn to the proof of (46). Let $\omega \in A_y$, that is $U(\omega) < F_X(y)$ and let $H := \{z \mid F_X(z) < U(\omega)\}$. Then $F_X^{-1}(U(\omega)) = \sup H$ and due to the left continuity of F_X we have $F_X\left(F_X^{-1}(U(\omega))\right) \leq U(\omega) < F_X(y)$. If $y = F_X^{-1}(U(\omega))$ were valid then it would imply $F_X(y) \leq U(\omega)$. If $y < F_X^{-1}(U(\omega))$ were valid then $y \in H$ would be satisfied, which would lead us to $F_X(y) < U(\omega)$. In other words, in case of $y \leq F_X^{-1}(U(\omega))$ by the monotonicity of F_X we would have $F_X(y) \leq F_X\left(F_X^{-1}(U(\omega))\right) \leq U(\omega)$, therefore $F_X^{-1}(U(\omega)) < y$. Thus we have shown (46).

Finally we turn to the proof of (47). Let us take now $\omega \in B_y$, that is $F_X^{-1}(U(\omega)) < y$ and let H denote the same set as above. Then $y \notin H$, therefore $F_X(y) \geq U(\omega)$, hence we obtain (47). $\square$

6.7 Remark. We can, of course, use $\tilde{F}_X$ instead of F_X in order to prove Theorem 6.6. In this case one can easily show the following implications:

$$\tilde{A}_y := \left\{\omega \in \Omega \mid U(\omega) < \tilde{F}_X(y)\right\} \subset \tilde{B}_y := \left\{\omega \in \Omega \mid F_X^{-1}(U(\omega)) \leq y\right\} \tag{48}$$

and

$$\tilde{B}_y \subset \tilde{C}_y := \left\{\omega \in \Omega \mid U(\omega) \leq \tilde{F}_X(y)\right\} \tag{49}$$

for all $y \in \mathbb{R}$.

Now one can show (48) in a similar way as (47) has been shown. Consider an appropriate formulation of F_X^{-1}: let $\bar{H} := \left\{z \mid \tilde{F}_X(z) \geq U(\omega)\right\}$, since in this case $F_X^{-1}(U(\omega)) = \inf \bar{H}$. Furthermore, due to the right continuity of $\tilde{F}_X$ we obtain $\tilde{F}\left(F_X^{-1}(U(\omega))\right) \geq U(\omega)$. Let us suppose now that $y < F_X^{-1}(U(\omega))$. Then by $y < \inf \bar{H}$ we have $y \notin \bar{H}$, thus we obtain $\tilde{F}(y) < U(\omega)$, which is a contradiction.

The implication (49) can be derived as (46) has been shown. To see this recall that $\tilde{F}_X$ is monotone and hence if $y \geq F_X^{-1}(U(\omega))$ then $\tilde{F}_X(y) \geq \tilde{F}_X(F_X^{-1}(U(\omega))) \geq U(\omega)$.

6.8 Theorem. *The lower VaR and the upper VaR are both monotone, positive homogeneous and translation invariant risk measures (over the set of all random variables of a probability space).*

The Value at Risk can be written in the forms $VaR_\alpha(X) = q^{1-\alpha}(-X)$ *and* $VaR^\alpha(X) = q_{1-\alpha}(-X)$ *as well.*

Moreover, given a random variable X *the functions* $\alpha \mapsto VaR_\alpha(X)$ *and* $\alpha \mapsto VaR_\alpha(X)$ $(\alpha \in (0,1))$ *are monotone non-increasing.*

Proof. We give the proof for the case of lower VaR, the statements can be similarly proved for upper VaR.

Monotonicity. If $\mathbb{P}(X \leq Y) = 1$ then $F_X(y) \geq F_Y(y)$ for all $y \in \mathbb{R}$. Therefore $\{y \mid F_X(y) < \alpha\} \subset \{y \mid F_Y(y) < \alpha\}$ that is

$$-VaR_\alpha(X) = \sup\{y \mid F_X(y) < \alpha\} \leq \sup\{y \mid F_Y(y) < \alpha\} = -VaR_\alpha(Y).$$

Positive homogeneity. If $h > 0$ then $F_{hX}(y) = F_X(y/h)$, $y \in \mathbb{R}$, hence $-VaR_\alpha(hX) = \sup\{y \mid F_{hX}(y) < \alpha\} = \sup\{y \mid F_X(y/h) < \alpha\} = h \sup\{z \mid F_X(z) < \alpha\} = -hVaR_\alpha(X)$.

Translation invariance. Given $a \in \mathbb{R}$ we have $F_{X+a}(y) = F_X(y-a)$, $y \in \mathbb{R}$, hence $-VaR_\alpha(X+a) = \sup\{y \mid F_{X+a}(y) < \alpha\} = \sup\{y \mid F_X(y-a) < \alpha\} = a + \sup\{z \mid F_X(z) < \alpha\} = a - VaR_\alpha(X)$.

The equivalent forms of VaR follows from the properties discussed in Remark 6.5.

Finally, if $\alpha < \beta$ then $\{z \mid F_X(z) < \alpha\} \subset \{z \mid F_X(z) < \beta\}$, therefore $q_\alpha(X) < q_\beta(X)$, which implies the monotonicity of VaR.

This proof would be almost literally the same if we had used the right continuous version of the distribution function ($\tilde{F}$) instead of the left continuous one (F). □

6.9 Remark. In the proof of Theorem 6.8 we could see that the lower and upper quantiles are both positive homogeneous, furthermore if $\mathbb{P}(X \leq Y) = 1$ then $q_\alpha(X) \leq q_\alpha(Y)$, $q^\alpha(X) \leq q^\alpha(Y)$, finally, for all $a \in \mathbb{R}$ we have $q_\alpha(X + a) = q_\alpha(X) + a$, $q^\alpha(X + a) = q^\alpha(X) + a$.

So far we have shown that the VaR fulfills three requirements needed for the coherency of a risk measure, it is only subadditivity that has not yet discussed. However, the VaR is not subadditive in general and hence it is not a coherent risk measure either. One can construct easy counterexamples to subaddtivity of VaR, we show such examples next.

6.10 Example. Consider a simple probability space, say, $(\Omega, \mathcal{F}, \mathbb{P})$, where $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}$, and let $\mathbb{P}(\omega_1) = 0,01$, $\mathbb{P}(\omega_2) = \mathbb{P}(\omega_3) = 0,03$. Let X and Y show the profit of two portfolios with

$$X(\omega_1) = -30, \quad X(\omega_2) = -20, \quad X(\omega_3) = -5, \quad X(\omega_4) = 20, \tag{50}$$

$$Y(\omega_1) = -30, \quad Y(\omega_2) = -5, \quad Y(\omega_3) = -20, \quad Y(\omega_4) = 20. \tag{51}$$

Then $VaR_{0,05}(X) = VaR^{0,05}(X) = VaR_{0,05}(Y) = VaR^{0,05}(X) = 5$, on the other hand

$$\mathbb{P}(X + Y = -60) = 0,01, \quad \mathbb{P}(X + Y = -25) = 0,06, \quad \mathbb{P}(X + Y = 40) = 0,93,$$

that is $VaR_{0,05}(X + Y) = VaR^{0,05}(X + Y) = 25$.

The next example is given by Paul Embrechts.

6.11 Example. Let us suppose that the random variables Y_i $(i = 1, \ldots, 100)$ are independent, identically distributed such that:

$$\mathbb{P}(Y_i = 2) = 0,99 \quad \text{and} \quad \mathbb{P}(Y_i = -100) = 0,01.$$

Consider a financial institution which grants credits, for instance the credit is of 100 thousand dollars, granted for one year with (borrowing) rate 2%. Then Y_i denotes the profit realised by the institution from such a contract: with high probability the credit institution earns 2 thousand dollars, however there is a 1 % chance that the credit will not be payed back, for instance due to the fact that the borrower goes bankrupt. (For the sake of simplicity we do not consider present value calculations, which should have been done in case of cash flows realised possibly at different time points.)

It is clear that $VaR_{0,05}(Y_i) = -2$. Consider now the VaR of the portfolio L_1 that consists of the 100 credit contracts mentioned above, i.e. let $L_1 = \sum_{i=1}^{100} Y_i$. We can put the portfolio in a different form, namely,

$$L_1 = \sum_{i=1}^{100} Y_i = \sum_{i=1}^{100} (102\xi_i - 100) = -100^2 + 102 \sum_{i=1}^{100} \xi_i,$$

where the random variables ξ are appropriate independent, Bernoulli distributed variables with parameter 0,99. In other words, $\sum_{i=1}^{100} \xi_i = \eta$, where η has binomial distribution with parameters $(100;\ 0,99)$. Applying the translation invariance and the positive homogeneity we obtain

$$VaR_{0,05}(L_1) = 102 VaR_{0,05}(\eta) + 100^2.$$

Furthermore notice that

$$VaR_{0,05}(\eta) = -q_{0,05}(\eta) > -100$$

since the largest possible value of η is 100.

Thus we conclude that

$$VaR_{0,05}\left(\sum_{i=1}^{100} Y_i\right) > \sum_{i=1}^{100} VaR_{0,05}(Y_i),$$

which means that subadditivity is not satisfied in our case.

This example is fairly surprising from several aspects, moreover, we can say that it is rather frightening for the financial profession and financial institutions. One of the reasons is that in contrast to the previous counterexample now the subadditivity is disproved in a case of independent and identically distributed portfolios/assets.

Another reason can be seen when we have a look at the result from a different perspective. Let L_2 be a portfolio that contains only one credit contract with the same rate and borrowing time as before (see Y_1) but written on 10000 thousand dollars, i.e. $L_2 = 100Y_1$. Since due to the positive homogeneity $VaR_{0,05}(L_2) = 100VaR_{0,05}(Y_1)$, thus we can write our results as

$$VaR_{0,05}(L_1) > VaR_{0,05}(L_2).$$

That is, this example just states the opposite of our thoughts and beliefs we had about portfolio diversification. The risk expressed in terms of VaR of the one-credit-contract portfolio is smaller than that of the diversified portfolio where the same amount of credit is diversified among independent clients given the same credit amount. And to top it all we note that $VaR_\alpha(Y_1) = -2$ whenever $0,01 < \alpha < 1$. Whereas in the same range the value of $VaR_\alpha(L_2)$ is changing in a very sensitive way as the value of α is changing.

As we have seen above the VaR is not a coherent risk measure. Moreover, the property which is not satisfied by VaR is the subadditivity, which would probably be the most important property according to many experts. This means that calculating first the risks of separate portfolios and then summing the risks one might get less than the risk of the portfolio in which the separate portfolios have been merged. Though one would expect that due to merging the portfolios a part of the risk would be eliminated or at least reduced. To solve these problems other risk measures have been introduced and studied in the literature. Among the alternative risk measures, the so-called expected shortfall has become the most widely used and appreciated risk measure due to some of its nice features.

6.3 The expected shortfall – An 'average' of great losses

The expected shortfall is based on a simple idea: consider the worst α*100 % of the possible future outcomes of the portfolio, like in case of VaR. Now take the average of these outcomes instead of just taking the upper bound of them (which is the best of them). That is, the expected shortfall is to show the expectation of the profit (loss) of the worst α*100 % cases. The rigorous definition can be formulated as follows.

6.12 Definition. *Let X be a random variable such that $\mathbb{E}(X)^- < \infty$ and let $\alpha \in (0,1)$, where $(X)^-$ denotes the negative part of X[4]. The α-expected shortfall (value) of X is*

$$ES_\alpha(X) = -\frac{1}{\alpha}\left(\mathbb{E}\left[X\mathbf{1}_{\{X \le q_\alpha(X)\}}\right] + q_\alpha(X)\left[\alpha - \mathbb{P}(X \le q_\alpha(X))\right]\right).$$

The following result helps us to derive and understand the features of the expected shortfall, in fact it is an equivalent form of the notion at issue.

6.13 Theorem. *Let X be a random variable with $\mathbb{E}(X)^- < \infty$ and let $\alpha \in (0,1)$. Then*

$$ES_\alpha(X) = -\frac{1}{\alpha}\int_0^\alpha q_u(X)du = -\frac{1}{\alpha}\int_0^\alpha F_X^{-1}(u)du.$$

Proof. Let U be a uniformly distributed random variable over the interval $(0,1)$ and define $\eta := F_X^{-1}(U)$. We have shown earlier that the distributions of X and U are the same. Having in mind that the function F_X^{-1} is monotone increasing we can obtain the following simple statements:

$$\{U \le \alpha\} \subset \{\eta \le q_\alpha(X)\},$$

[4]If $x \in \mathbb{R}$ then the positive part of x is:

$$(x)^- = \begin{cases} x & \text{if } x > 0 \\ 0 & \text{otherwise.} \end{cases}$$

The negative part of x is defined as $(x)^- = (-x)^+$.

furthermore, $U(\omega) > \alpha$, $\omega \in \Omega$, (and hence $\eta(\omega) \geq q_\alpha(X)$) and $\eta(\omega) \leq q_\alpha(X)$ can both be fulfilled only if $\eta(\omega) = q_\alpha(X)$. Based on these we obtain

$$\begin{aligned}
\int_0^\alpha q_u(X)du &= \mathbb{E}\left(\eta \mathbf{1}_{\{U\leq\alpha\}}\right) \\
&= \mathbb{E}\left(\eta \mathbf{1}_{\{\eta\leq q_\alpha(X)\}}\right) - \mathbb{E}\left(\eta \mathbf{1}_{\{U>\alpha\}\cap\{\eta\leq q_\alpha(X)\}}\right) \\
&= \mathbb{E}\left(X \mathbf{1}_{\{X\leq q_\alpha(X)\}}\right) - q_\alpha(X)\mathbb{P}\left(\{U>\alpha\}\cap\{\eta\leq q_\alpha(X)\}\right) \\
&= \mathbb{E}\left(X \mathbf{1}_{\{X\leq q_\alpha(X)\}}\right) - q_\alpha(X)\left(\mathbb{P}(X\leq q_\alpha(X)) - \alpha\right).
\end{aligned}$$

One gets the desired statement by dividing by $-\alpha$. □

The results of Theorem 6.13 can be formulated with the use of upper quantiles. It is easy to see that $\int_0^\alpha q_u(X)du = \int_0^\alpha q^u(X)du$, thus $ES_\alpha(X) = -\frac{1}{\alpha}\int_0^\alpha q^u(X)du$. Similarly, one could replace the lower quantiles by the corresponding upper quantiles in Definition 6.12, the replacement would not change the value of the expected shortfall. To see this first note that the statement is trivial if $q^\alpha(X) = q_\alpha(X)$. If $q^\alpha(X) > q^\alpha(X)$ then $\mathbb{P}(X \leq q_\alpha(X)) = \alpha$ and we have

$$\begin{aligned}
&\mathbb{E}\left[X\mathbf{1}_{\{X\leq q^\alpha(X)\}}\right] + q^\alpha(X)\left[\alpha - \mathbb{P}(X\leq q^\alpha(X))\right] \\
&\quad= \mathbb{E}\left[X\mathbf{1}_{\{X\leq q_\alpha(X)\}}\right] + q^\alpha(X)\mathbb{P}(X = q^\alpha(X)) + q^\alpha(X)\left[\alpha - \mathbb{P}(X\leq q_\alpha(X)) - \mathbb{P}(X = q^\alpha(X))\right] \\
&\quad= \mathbb{E}\left[X\mathbf{1}_{\{X\leq q_\alpha(X)\}}\right] = ES_\alpha(X).
\end{aligned}$$

Thus in contrast to the VaR we do not differentiate between a lower version of expected shortfall and an upper version of it. Moreover, the expected shortfall can be written in some other alternative forms as well, namely:

$$ES_\alpha(X) = -\frac{1}{\alpha}\left(\mathbb{E}\left[X\mathbf{1}_{\{X<q_\alpha(X)\}}\right] + q_\alpha(X)\left[\alpha - \mathbb{P}(X < q_\alpha(X))\right]\right),$$

or

$$ES_\alpha(X) = -\frac{1}{\alpha}\left(\mathbb{E}\left[X\mathbf{1}_{\{X\leq s\}}\right] + s\left[\alpha - \mathbb{P}(X\leq s)\right]\right), \quad \forall s \in [q_\alpha(X), q^\alpha(X)].$$

6.14 Remark. Before turning to the most important properties of expected shortfall we make some preparatory technical remarks. Let X be a random variable with $\mathbb{E}(X)^- < \infty$ and let $\alpha \in (0,1)$. Let us introduce the following function for all $y \in \mathbb{R}$ and $\omega \in \Omega$:

$$\mathbf{1}^{(\alpha)}_{\{X(\omega)\le y\}} := \begin{cases} \mathbf{1}_{\{X(\omega)\le y\}}, & \text{if } \mathbb{P}(X=y)=0 \\ \mathbf{1}_{\{X(\omega)\le y\}} + \frac{\alpha-\mathbb{P}(X\le y)}{\mathbb{P}(X=y)}\mathbf{1}_{\{X(\omega)=y\}}, & \text{if } \mathbb{P}(X=y)>0. \end{cases}$$

Next we summarise some simple properties of this function. If $X(\omega) > y$ then $\mathbf{1}^{(\alpha)}_{\{X(\omega)\le y\}} = 0$, if $X(\omega) < y$ then $\mathbf{1}^{(\alpha)}_{\{X(\omega)\le y\}} = 1$. Furthermore, for $X(\omega) = q_\alpha(X)$ if $\mathbb{P}(X = y) > 0$ then $0 \le \mathbb{P}(X \le q_\alpha(X)) - \alpha < \mathbb{P}(X = q_\alpha(X))$. Due to the assertions just mentioned we have

$$\mathbf{1}^{(\alpha)}_{\{X\le q_\alpha(X)\}} \in [0,1]. \tag{52}$$

One can directly see from the definition the following:

$$\mathbb{E}\mathbf{1}^{(\alpha)}_{\{X\le q_\alpha(X)\}} = \alpha, \tag{53}$$

$$ES_\alpha(X) = -\alpha^{-1}\mathbb{E}X\mathbf{1}^{(\alpha)}_{\{X\le q_\alpha(X)\}}. \tag{54}$$

6.15 Theorem. *Consider $V = \{X \text{random variable } \mid \mathbb{E}(X)^- < \infty\}$ on a probability space. Then the expected shortfall is coherent on the set V.*

Proof. The properties of the quantiles given in Remark 6.9 together with Theorem 6.13 imply that the expected shortfall is monotone, positive homogeneous and translation invariant.

Thus it remains to prove the subadditivity. Let X and Y be random variables with $\mathbb{E}(X)^- < \infty$, $\mathbb{E}(Y)^- < \infty$, and write $Z = X + Y$. Then by (52) we have

$$(X - q_\alpha(X))\left(\mathbf{1}^{(\alpha)}_{\{Z\le q_\alpha(Z)\}} - \mathbf{1}^{(\alpha)}_{\{X\le q_\alpha(X)\}}\right) \ge 0. \tag{55}$$

Since for $X(\omega) > q_\alpha(X)$ one gets $\mathbf{1}^{(\alpha)}_{\{X(\omega)\leq q_\alpha(X)\}} = 0$ and for $X(\omega) < q_\alpha(X)$ one gets $\mathbf{1}^{(\alpha)}_{\{X(\omega)\leq q_\alpha(X)\}} = 1$. Furthermore

$$\begin{aligned}
&\alpha[ES_\alpha(X) + ES_\alpha(Y) - ES_\alpha(Z)] \\
&\quad = \mathbb{E}\left(Z\mathbf{1}^{(\alpha)}_{\{Z\leq q_\alpha(Z)\}} - X\mathbf{1}^{(\alpha)}_{\{X\leq q_\alpha(X)\}} - Y\mathbf{1}^{(\alpha)}_{\{Y\leq q_\alpha(Y)\}}\right) \\
&\quad = \mathbb{E}\left(X\left[\mathbf{1}^{(\alpha)}_{\{Z\leq q_\alpha(Z)\}} - \mathbf{1}^{(\alpha)}_{\{X\leq q_\alpha(X)\}}\right] + Y\left[\mathbf{1}^{(\alpha)}_{\{Z\leq q_\alpha(Z)\}} - \mathbf{1}^{(\alpha)}_{\{Y\leq q_\alpha(Y)\}}\right]\right) \\
&\quad = \mathbb{E}\left([X - q_\alpha(X)]\left[\mathbf{1}^{(\alpha)}_{\{Z\leq q_\alpha(Z)\}} - \mathbf{1}^{(\alpha)}_{\{X\leq q_\alpha(X)\}}\right] + [Y - q_\alpha(Y)]\left[\mathbf{1}^{(\alpha)}_{\{Z\leq q_\alpha(Z)\}} - \mathbf{1}^{(\alpha)}_{\{Y\leq q_\alpha(Y)\}}\right]\right) \\
&\qquad + \mathbb{E}\left(q_\alpha(X)\left[\mathbf{1}^{(\alpha)}_{\{Z\leq q_\alpha(Z)\}} - \mathbf{1}^{(\alpha)}_{\{X\leq q_\alpha(X)\}}\right] + q_\alpha(Y)\left[\mathbf{1}^{(\alpha)}_{\{Z\leq q_\alpha(Z)\}} - \mathbf{1}^{(\alpha)}_{\{Y\leq q_\alpha(Y)\}}\right]\right) \\
&\quad \geq q_\alpha(X)\mathbb{E}\left[\mathbf{1}^{(\alpha)}_{\{Z\leq q_\alpha(Z)\}} - \mathbf{1}^{(\alpha)}_{\{X\leq q_\alpha(X)\}}\right] + q_\alpha(Y)\mathbb{E}\left[\mathbf{1}^{(\alpha)}_{\{Z\leq q_\alpha(Z)\}} - \mathbf{1}^{(\alpha)}_{\{Y\leq q_\alpha(Y)\}}\right] \\
&\quad = q_\alpha(X)(\alpha - \alpha) + q_\alpha(Y)(\alpha - \alpha) = 0,
\end{aligned}$$

where the first equality is due to (54), the inequality is due to (55), whereas (53) has been applied to obtain the last line. □

6.16 Theorem. *If X is a random variable with $\mathbb{E}(X)^- < \infty$ then the function*

$$\alpha \mapsto ES_\alpha(X), \qquad \alpha \in (0,1)$$

is continuous and monotone decreasing.

Proof. Recalling Theorem 6.13 we can see that the expected shortfall is a continuous function of the confidence level α. Since $-q_\alpha(X)$ is non-increasing in α, hence

$$\frac{\int_0^\alpha -q_u(X)du}{\alpha}$$

is also non-increasing, which implies that the expected shortfall is monotone decreasing in α. To see this in more details let $0 < \alpha_1 < \alpha_2 < 1$. Due to the monotonicity of the quantile there exists an $m \in \mathbb{R}$ such that $\int_{\alpha_1}^{\alpha_2} -q_u(X)\, du = m\,(\alpha_2 - \alpha_1)$ and $m \leq -q_{\alpha_1}(X)$. Hence we also have $m \leq ES_{\alpha_1}(X)$. Thus

$$\begin{aligned} ES_{\alpha_2}(X) &= \frac{1}{\alpha_2}\int_0^{\alpha_2} -q_u(X)\, du + \frac{1}{\alpha_2}\int_{\alpha_1}^{\alpha_2} -q_u(X)\, du \\ &= \frac{\alpha_1}{\alpha_2} ES_{\alpha_1}(X) + \left(1 - \frac{\alpha_1}{\alpha_2}\right) m \leq ES_{\alpha_1}(X). \end{aligned}$$

□

7 Bibliographic Notes

As we have seen utility theory forms the base of our set up and we were especially focused on the Neumann-Morgenstern type of utility functions and related topics. For this we have found the papers [Berde & Petró] and [Eső & Lóránd] fairly useful. They provide a very good overview on the issue. Introduction to utility theory can be found in many books on microeconomics and related areas, like in [Nordhaus & Samuelson] or [Kreps]. To find further references and more about recent research we refer again to [Berde & Petró] and [Eső & Lóránd] who also supply a large list of references.

There are many publications and mainly textbooks available on portfolio management which discuss the problem of portfolio selection. For our purposes we have used [Ingersoll] and [Huang & Litzenberger] the most though we also mention that the mathematics in these books is not found sufficiently precise for our aims. For readers with main interest in economics [Elton & Gruber] can help the better understanding on portfolio problems whereas [Hull] provides a great overview on (derivative) securities and related practical problems. We also mention the excellent work of [Duffie] where not only portfolio problems but mainly pricing problems and related topics in securities market are studied.

During the development of Chapter 6 the following works were used fruitfully: [Acerbi], [Acerbi2], where we found a detailed discussion of the VaR and expected shortfall; [Delbaen], where coherency of measures is studied in details; [Dowd], where a great overview of the VaR, its usage in economics and the discussion of its estimation can be found. Concerning quantiles and its properties [Acerbi2] and [Major] provided excellent help for us. Finally we mention the useful works (papers, slides, etc) of Paul Embrechts where one can find the discussion of many practical and theoretical issues concerning risk measures (see http://www.math.ethz.ch/~embrechts/).

For better understanding on the means of classical mathematical calculus used in this work we refer to [Lang] which provides a good exposition like many others. The work [Rockafellar] gives a detailed discussion on convex analysis. We also mention the excellent book of [Cohn] to those who are interested in measure theory.

8 Appendix

The following two theorems are fundamental and well-known in classical calculus. They together with their proofs can certainly be found in many (introductory) books on mathematical analysis.

8.1 Theorem. (The method of Lagrangian multipliers) *Let U be an open set in $\mathbb{R}^p$ ($p \in \mathbb{N}^+$) and suppose that $f, g_1, \ldots, g_k \in \mathcal{C}^1(U)$. Define*

$$S = \{x \in U \mid g_i(x) = 0, \quad i = 1, \ldots, k\}.$$

If $c \in S$ is a point where f achieves a local maximum then the gradient vectors

$$\triangledown g_1(c) = \begin{pmatrix} \frac{\partial}{\partial x_1} g_1(c) \\ \cdot \\ \cdot \\ \cdot \\ \frac{\partial}{\partial x_p} g_1(c) \end{pmatrix}, \ldots, \triangledown g_k(c) = \begin{pmatrix} \frac{\partial}{\partial x_1} g_k(c) \\ \cdot \\ \cdot \\ \cdot \\ \frac{\partial}{\partial x_p} g_k(c) \end{pmatrix}, \triangledown f(c) = \begin{pmatrix} \frac{\partial}{\partial x_1} f(c) \\ \cdot \\ \cdot \\ \cdot \\ \frac{\partial}{\partial x_p} f(c) \end{pmatrix}$$

are linearly dependent.

Note furthermore that using Theorem 8.1 one can show the existence of constants $\lambda_1, \ldots, \lambda_k$ such that

$$\triangledown f(c) = \lambda_1 \triangledown g_1(c) + \ldots + \lambda_k \triangledown g_k(c).$$

Hence, the locations of the local maximums of f can be found as the solutions of the following system of equations:

$$\frac{\partial}{\partial x_i} \left(f(x) + \sum_{j=1}^{k} \lambda_j g_j(x) \right) = 0, \qquad i = 1, \ldots, p,$$

$$\frac{\partial}{\partial \lambda_t} \left(f(x) + \sum_{j=1}^{k} \lambda_j g_j(x) \right) = 0, \qquad t = 1, \ldots, k.$$

8.2 Theorem. (The implicit function theorem) *Let $U \subset \mathbb{R}^{p+q}$ be an open set, $p, q \in \mathbb{N}^+ \cup \{\infty\}$, and assume that $f \in \mathcal{C}^n(U \to \mathbb{R}^q)$. Define $S = \{x \in U \mid f(x) = 0\}$. Let us suppose that $a \in S$ such that*

$$det \begin{pmatrix} \frac{\partial}{\partial x_{p+1}} f_1(a) & \frac{\partial}{\partial x_{p+2}} f_1(a) & \dots & \frac{\partial}{\partial x_{p+q}} f_1(a) \\ . & & & . \\ . & & & . \\ . & & & . \\ \frac{\partial}{\partial x_{p+1}} f_q(a) & \frac{\partial}{\partial x_{p+2}} f_q(a) & \dots & \frac{\partial}{\partial x_{p+q}} f_q(a) \end{pmatrix} \neq 0.$$

Then there is a neighbourhood $U_0 \subset \mathbb{R}^{p+q}$ of point a and there is a function g defined on an open set $W \subset \mathbb{R}^p$ such that $g \in \mathcal{C}^n(W \to \mathbb{R}^q)$ and

$$S \cap U_0 = \Big\{ \big(w, g(w)\big) \;\Big|\; w \in W \Big\}.$$

Furthermore, $a_I \in W$, $g(a_I) = a_{II}$, where $a = (a_1, \dots, a_{p+q})$, $a_I = (a_1, \dots, a_p)$, $a_{II} = (a_{p+1}, \dots, a_{p+q})$ and $f(w, g(w)) = 0$ for all $w \in W$.

Next we give some useful versions of the formula of integration by parts. The proof can be found in several books on measure theory. We refer to Chapter 5 of [Cohn].

8.3 Theorem. (Integration by parts) *Let F and G be monotone increasing, left continuous, bounded, real valued functions on the real line with $\lim_{x\to-\infty} F(x) = \lim_{x\to-\infty} G(x) = 0$. Let $a, b \in \mathbb{R}$, $a < b$.*

Then

$$\int_{[a,b]} G(x) dF(x) + \int_{[a,b]} F(x+) dG(x) = F(b+)G(b+) - F(a)G(a), \tag{56}$$

where $F(x+)$ and $G(x+)$ denotes the right limit of F and G at x respectively.

From (56) one can obtain another form, namely

$$\int_{[a,b]} \frac{G(x)+G(x+)}{2} dF(x) + \int_{[a,b]} \frac{F(x)+F(x+)}{2} dG(x) = F(b+)G(b+) - F(a)G(a).$$

Note also that if F and G have no common point of discontinuity then we can simplify the above formula and write

$$\int_{[a,b]} G(x)dF(x) + \int_{[a,b]} F(x)dG(x) = F(b+)G(b+) - F(a)G(a).$$

References

[Acerbi] Acerbi, C.: "Coherent Representations of Subjective Risk Aversion", in Giorgio Szeg (Ed): *Risk Measures for the 21st Century*, Wiley, New York, 2004.

[Acerbi2] Acerbi, C.: *"On the coherence of Expected Shortfall"*, http://www.gloriamundi.org/detailpopup.asp?ID=453054765, 2002.

[Arató] Arató, M.: *"Általános biztosításmatematika (General insurance mathematics)"* (in Hungarian), ELTE Eötvös Kiadó, Budapest, 1997.

[Berde & Petró] Berde, É. and Petró, K.: "A különféle hasznosságfogalmak szerepe a közgazdaságtanban (The role of the different notions of utility in economics)" (in Hungarian), *Közgazdasági Szemle*, 1995, 5, pp. 511-529.

[Cohn] Cohn, D. L.: *"Measure Theory"*, Birkhäuser Boston, 1980.

[Delbaen] Delbaen, F.: *"Coherent risk measures on general probability spaces"*, Eidgenssische Technische Hochschule, Zrich, 2000.

[Dowd] Dowd, K.: *"Beyond Value at Risk"*, The New Science of Risk Management, John Wiley & Sons, 1999.

[Duffie] Duffie, D.: *"Dynamic Asset Pricing Theory"*, Princeton University Press, New Jersey, 1992.

[Elton & Gruber] Elton, E. J. and Gruber M. J.: *"Modern portfolio theory and investment analysis"*, John Wiley & Sons, Inc., 1995.

[Eső & Lóránd] Eső, P. and Lóránd, G.: "A racionalitás közgazdasági értelmezéséről (On the economic interpretation of rationality)" (in Hungarian), *Közgazdasági Szemle*, 1993, 4, pp. 311-324.

[Huang & Litzenberger] Huang, Chi-Fu and Litzenberger, R. H.: *"Foundations for financial economics"*, Prentice Hall, 1988.

[Hull] Hull, J. C.: *"Options, futures, and other derivative securities"*, Prentice-Hall International, Inc., 1993.

[Ingersoll] Ingersoll, J. E.: *"Theory of Financial Decision Making"*, Rowman & Littlefield, 1987.

[Korn] Korn, R.: *"Optimal portfolios"*, World Scientific, 1998.

[Kreps] Kreps, D. M.: *"A course in microeconomic theory"*, Princeton University Press, Princeton, 1990.

[Lang] Lang, S.: *"Analysis"* I., II., Addison-Wesley Publishing Company, 1968,1969.

[Major] Major, P.: *" The relation between the closeness of random variables and their distributions"*, http://www.renyi.hu/~major/probability/couple.html, 2005.

[Nordhaus & Samuelson] Nordhaus, W. D. and Samuelson, P. A.: *"Economics"*, McGraw-Hill, Inc., 1992.

[Potters] Potters, J.A.M.: *"Mathematische economie"*, Universiteit Nijmegen, 1996.

[Potters2] Potters, J.A.M.: *"Preferences on stochastic outcomes"*, manuscript, Universiteit Nijmegen, 1998.

[Rockafellar] Rockafellar, R. T.: *"Convex Analysis"*, Princeton, New Jersey, 1970.

[Schmidt] Schmidt, U.: *"Axiomatic Utility Theory under Risk"*, Springer, Berlin, 1998.