Univerzita Karlova v Praze
Matematicko-fyzikální fakulta

DISERTAČNÍ PRÁCE

RNDr. Ing. Miloš Kopa

Užitkové funkce v úlohách optimalizace portfolia

Katedra pravděpodobnosti a matematické statistiky

Vedoucí disertační práce:	Prof. RNDr. Jitka Dupačová, DrSc.
Studijní obor:	M5 – Ekonometrie a operační výzkum

Praha, 2006

Charles University in Prague
Faculty of Mathematics and Physics

DOCTORAL THESIS

RNDr. Ing. Miloš Kopa

Utility functions in portfolio optimization

Department of Probability and Mathematical Statistics

Supervisor: Prof. RNDr. Jitka Dupačová, DrSc.
Branch: M5 – Econometrics and Operations Research

Prague, 2006

Acknowledgements

I am very much indebted to Prof. Jitka Dupačová for her excellent guidance and huge support during my research at Charles University. I would like to express deep thanks to Prof. Thierry Post for his patient leadership during my study stay at Erasmus University in Rotterdam.

I thank to my parents for financial and moral support.

This work was partly supported by the Grant Agency of the Czech Republic (grants 402/02/1015, 201/05/H007, 402/05/0115 and 201/05/2340)

Prague, May 31, 2006 RNDr. Ing. Miloš Kopa

Contents

1 Introduction **6**

2 Stability in portfolio optimization **15**
- 2.1 Preliminaries . . . 15
- 2.2 Portfolio selection problem . . . 16
- 2.3 Stability of optimal portfolio . . . 17
  - 2.3.1 Stability of optimal expected utility . . . 17
  - 2.3.2 Stability of optimal investment strategy . . . 20
- 2.4 Variational analysis approach . . . 28

3 Risk premiums **33**
- 3.1 Preliminaries . . . 33
- 3.2 Multivariate risk premium . . . 33
- 3.3 Multivariate risk premium with random initial wealth . . . 34
- 3.4 Multiperiod risk premium . . . 37
- 3.5 Generalizations of multiperiod risk premium . . . 40

4 SSD portfolio efficiency **46**
- 4.1 Preliminaries . . . 46
- 4.2 Stochastic dominance . . . 47
- 4.3 VaR and CVaR . . . 49
  - 4.3.1 CVaR for scenario approach . . . 51
  - 4.3.2 Relationship between risk premium, VaR and CVaR . . . 53
- 4.4 SSD portfolio efficiency criteria . . . 54
- 4.5 Algorithm for testing SSD portfolio efficiency . . . 61
- 4.6 SSD portfolio inefficiency measure . . . 62
- 4.7 Numerical application: SSD efficiency of mean-VaR optimal portfolios . . . 69

5 FSD portfolio efficiency **74**
5.1 Preliminaries 74
5.2 FSD optimality versus FSD admissibility 75
5.3 Representative utility functions 77
5.4 Mathematical programming formulation 80
5.5 Mixed-integer Programming Algorithm for Testing FSD Optimality 83
5.6 Numerical example 85

6 Summary and open problems **92**

Bibliography **94**

Chapter 1

Introduction

Dealing with decision problems one has to choose an action from a given set of alternatives with uncertain consequences. For example, consider a decision maker who wishes to allocate his resources to different investment opportunities in an "optimal way". There are several approaches how to construct the decision criterion under risk to choose the optimal alternative. Almost all of these models are based on some measure of yield and risk. Typically a measure of yield of an alternative (investment opportunity) is maximized and a measure of risk is minimized. However, the behavior of decision maker depends on his risk attitude. One of the classical ways of involving risk factor in portfolio selection problem is considering a utility function when the optimal alternatives maximize expected utility.

If the yield and risk of an asset are measured separately, the yield is usually measured by expected value. On the other hand, there is no generally accepted measure of risk. Therefore there are several different mean-risk models for various types of risk measures: variance, semi-variance, upper semi-deviation, Value at Risk, conditional Value at Risk, etc. When the concept of maximizing expected utility is applied, from the type of utility function one can derive another risk measures: Arrow-Pratt absolute (relative) risk aversion measure and risk premiums.

In all portfolio optimizing models some kind of a risk parameter is included and some distribution of yields is assumed. Since the risk parameter and the distribution of yields are usually not exactly known, one can analyze the dependence of optimal solutions on these inputs.

When no information about risk attitude of the decision maker is known one can apply a stochastic dominance approach. In the context of the stochastic dominance for portfolio selection problems the efficiency of a given portfolio is analyzed. In this thesis we will examine a utility theory, risk measures and stochastic dominance approach with application to portfolio selection problems.

The basics of utility theory are connected with von Neumann & Morgenstern [54] where the existence of von Neumann - Morgenstern utility function is analyzed. This function u has such a property that a rational decision maker prefers alternative X to alternative Y if and only if $Eu(X) \geq Eu(Y)$. More general axiomatic theory of utility was presented in Černý *et al.* [6], Ziemba & Vickson [57] and references therein, especially Herstein & Milnor [19] and Fishburn [12].

There are two approaches to construction of utility functions: direct (cardinal utility function) and indirect (ordinal utility funciton). An ordinal utility function for an individual consists of a rank ordering of possible states of affairs for that individual. An ordinal function tells us that decision maker prefers X to Y, but it doesn't tell us whether X is much better than Y or only a little better. A cardinal utility function assigns a real-number value for each possible state of affairs. The assumptions for existence of the cardinal or ordinal utility function are derived in e.g. Černý *et al.* [6] and Glűckaufová & Černý [16]. In this thesis we focus on the cardinal utility function where the utility is assigned to the total wealth of a decision maker.

There are several characterization of utility functions. In Kopa [25], three ways of utility function classification are presented. They are based on: Arrow-Pratt absolute (relative) risk aversion measure, "preference switching" and "star shape". The classical characterization of Arrow [1] and Pratt [45] deals with twice differentiable and increasing utility functions. There is a close relationship between risk aversion, risk seeking or risk neutrality of an investor and the sign of the Arrow-Pratt absolute risk aversion measure. A concave (convex, linear) utility functions represent risk averse (risk seeking, risk neutral) decision maker. Another way how to express the risk attitude of decision maker is represented by "risk premiums". The preference switching characterization explores the number of switching preferences between any two gambles, as initial wealth increases, see e.g. Pedersen & Satchel [41], Kopa [25]. Especially zero-switch utility functions are of interest. One-switch utility functions, where at most one preference switching between any two gambles occurs due to changes in wealth, were analyzed in Bell [4]. Similarly to concave utility functions, star-shape utility functions also exhibit risk aversion at some wealth position, see Landsberger & Meilijson [35]. The comparison of concave and star-shaped utility functions shows that concave functions have decreasing marginal slope whereas star-shaped functions have decreasing average slope from the point at which they are star-shaped.

There is a host of areas where utility theory can be applied. For example, the utility function can be used in medical survival analysis. In insurance theory one can exploit utility function to estimate fair insurance premium level. In this thesis we will apply the utility theory to a portfolio selection problem in order to analyze the optimal investment strategy of the decision maker. In optimization models with utility functions the expected utility of the final wealth is maximized. Therefore the portfolio selection problem is a problem of stochastic programming.

According to Rőmisch [48] and references therein, one may derive a stability result for set of optimal solutions in the case when an underlying probability distribution is perturbed or approximated. As a consequence of this theory, we can provide a scenario-based approximation of distribution of yields in the portfolio selection problem and estimate the maximal error caused by using approximate distribution.

In classical approach, utility functions for one-period investment possibility are considered. When a multiperiod investment possibilities are analyzed the decision problem is dynamic and it leads to dynamic portfolio selection problem. In this case, one can search for investment strategy as a sequence of decisions. In this thesis we assume discrete time multiperiod problems defined in e.g. Dupačová *et al.* [10]. In these problems, a multidimensional utility function is maximized. These functions are shown and analyzed in e.g. Ambarish & Kallberg [2], Duncan [9], Dupačová *et al.* [10], Kihlstrom & Mirman [23] and Richard [46].

In spite of a large number of papers dealing with utility functions, the theory of utility functions with application to portfolio selection problem is still actual and of interest due to three reasons. Firstly, the computational aspect of solving one-period portfolio selection problems is no more limiting. Secondly, less conventional classes of utility functions become more important. For example, according to Kopa & Post [32], the representative set of utility functions in the case of first-order stochastic dominance consists of discontinuous utility functions. It opens an area for research concerning suitable assumptions for utility functions in context of portfolio selection problem. Finally, portfolio selection problem with multiperiod investment possibilities can be formulated using multiperiod utility functions.

An alternative formulation of the portfolio selection problem is represented by mean-risk models. If risk is measured by variance, the Markowitz model is considered, see Markowitz [39]. In Ogryczak & Ruszczyński [40], some of the other risk measures such as: absolute deviation, absolute semideviation, standard semideviation, Value at Risk (VaR), conditional Value at Risk (CVaR) and Gini mean difference are analyzed with respect to relationship to stochastic dominance. All of these measures are based on some risk parameter and on certain distribution of yields. Some of the corresponding mean-risk models can be derived as a special case of maximizing expected utility problem. For example, if quadratic utility function is assumed, variance is the appropriate measure of risk. If a decision maker has not a utility function consistent with any of these mean-risk models, he needs to quantify his risk by another, more general measure of risk, so-called the risk premium.

Risk premiums can be derived from any type of utility function and for any investment opportunity. The basic ideas of the risk premium approach come from Pratt [45] when the risk premium for one-period and univariate gamble is constructed. A generalization of this approach in order to define multidimensional premium for one-period gamble was suggested in e.g. Duncan [9], Kihlstrom & Mirman [23] or Richard [46]. To derive risk premium for multiperiod risks, one can apply the modification of multidimensional premium in Ambarish & Kallberg [2]. The construction of the multiperiod risk premium based on the preference indifference between accepting a multiperiod gamble and rejecting the gamble with possibility of accepting the gamble only in some time periods was presented in Kopa [29]. This approach is a generalization of Duncan [9], Kihlstrom & Mirman [23], Richard [46] or Ambarish & Kallberg [2]. Another way how to construct multiperiod risk measures was shown in Eichhorn & Rőmisch [11] using polyhedral risk measures. This measures are defined as optimal values of certain linear stochastic programs where the arguments of the risk measure appear on the right-hand side of the dynamic constraints. Multiperiod extensions of CVaR are an example for polyhedral risk measure.

The portfolio selection problem may be regarded as a two-step procedure. Firstly, an efficient set among all available portfolios is chosen and then the risk preferences of a decision maker to this set are applied. When no information about risk preferences is known, an efficiency of a given portfolio can be tested with respect to stochastic dominance rules. First-order stochastic dominance (FSD) is one of the fundamental concepts of decision making under uncertainty, relying only on the assumption of nonsatiation, or decision makers preferring more to less. Assuming a concavity of utility functions, a second-order stochastic dominance (SSD) approach can be employed.

There are well-known, simple tests for establishing FSD and SSD relationships between a pair of choice alternatives; see, e.g. Hanoch & Levy [18], Levy [36], Levy [37]. The third or higher degree of stochastic dominance was analyzed in e.g. Levy [36], Whitmore [55] and Whitmore [56]. Unfortunately, these tests have a limited use in applications with more than two choice alternatives. At present, the analysis of investment portfolios is a case of interest; investors generally can form a large number of portfolios by diversifying across individual assets. For such applications, there were developed special tests that analyzed whether a given portfolio is FSD efficient or SSD efficient relative to all possible portfolios. In this thesis, one of SSD efficiency tests is introduced and the FSD efficiency test based on FSD optimality is derived.

Assuming scenario approach for distribution of outcomes, Kuosmanen [34], Post [43], Post [44] presented linear programming SSD efficiency tests of a given portfolio. There was a historical development of SSD efficiency property. The first ideas come from Post [43]. The Post test exploits a structure of the set of representative utility functions when the diversification is allowed. For pairwise comparisons, Russel & Seo [50] showed that the set of two-piece linear utility functions is representative for all concave utility functions. In context of portfolio optimization, Post [43] proved that the set of piece-wise linear utility functions is representative. He presented very fast linear programing test. The Kousmanen SSD efficiency test is based on so-called dominating sets of portfolio return profile employing empirical distribution functions and pairwise SSD criteria. Under the assumption of scenario approach, in this thesis, a linear programming SSD efficiency test based on the relationship between CVaR and a dual second-order stochastic dominance properties is derived. In contrast to the Post approach, we follow Kuosmanen [34] and Ruszczyńsky & Vanderbei [51] in considering less stringent definition of SSD efficiency. Therefore, the Post criterion is only a necessary condition. From empirical point of view, this necessary condition is very powerful. However, this criterion fail in detecting SSD dominating portfolio with the same mean as a tested portfolio. It means, that the Post criterion classifies portfolio as SSD efficient even if there exists a SSD dominating portfolio in the sense of Hanoch & Levy [18] or Levy [36]. Comparing our test with the Kuosmanen test, our test leads to smaller linear problem than the Kuosmanen test. Moreover, contrary to both the Post and the Kuosmanen tests, if a given portfolio is SSD inefficient, our test detects a dominating portfolio which is SSD efficient. More general stochastic problem with stochastic dominance constraints was solved in Dentcheva & Ruszczyński [8]. However, there is no application to SSD efficiency test in this reference. In Ruszczyński & Vanderbei [51] a SSD efficiency in a mean-risk space was analyzed. A specialized parametric method for the entire mean-risk efficient frontiers was developed.

A complication in testing FSD portfolio efficiency is that we must distinguish between efficiency criteria based on "admissibility" and "optimality". There is a subtle difference between these two concepts. An alternative is FSD admissible if and only if no second alternative is preferred by all non-satiable decision-makers. This concept is relevant for expected utility theory with non-decreasing utility functions, as well as other theories of risky choice that are consistent with FSD, such as cumulative prospect theory. However, when using expected utility theory, admissibility is generally weaker than optimality. An alternative is FSD optimal if and only if it is the optimal choice for at least some non-decreasing and non-constant utility function. For pairwise comparison, the two concepts are identical. However, more generally, when multiple alternatives are available, FSD admissibility is a necessary but not sufficient condition for FSD optimality. In other words, an alternative may be admissible even if it is not optimal for any non-decreasing and non-constant utility function.

Bawa *et al.* [3] and Kuosmanen [34] propose FSD tests that apply under more general conditions than a pairwise test does. The two tests differ in a subtle way. While Bawa *et al.* [3] consider all convex combinations of the distribution functions of a given set of alternatives, Kuosmanen [34] considers the distribution function for all convex combinations of a given set of alternatives. Each of these two tests captures an important aspect of portfolio choice that is not captured by a pairwise FSD test. Still, both tests miss some key aspects of a proper FSD portfolio optimality test and both tests generally give a necessary but not sufficient condition. The linear programming test of Bawa *et al.* is based on optimality, but it does not account for diversification across the choice alternatives. Even though the mixed-integer linear programming test of Kuosmanen does account for diversification, it relies on admissibility rather than optimality.

In Kopa & Post [32], a proper test for FSD optimality of a given portfolio relative to all portfolios formed from a set of alternatives is derived. The reformulation of the FSD optimality criterion in terms of a set of elementary representative utility functions is presented. For pairwise FSD comparisons, Russell & Seo [50] showed that the set of three-piece linear functions - where the first and the last piece is constant - is representative for all admissible utility functions. In portfolio context, with diversification allowed, a more general class of piecewise constant functions is relevant. Kopa & Post [32] developed a linear programming test for searching over all representative utility functions in order to test a portfolio optimality. To identify the input for the linear programming problem, they suggest to use mixed-integer linear programming or subsampling techniques. In contrast to Bawa *et al.* [3], they consider diversified portfolios in addition to the individual, undiversified alternatives, and in contrast to Kuosmanen [34], they rely on optimality rather than admissibility.

Due to concavity of utility functions, the analysis of SSD efficiency is simpler than FSD efficiency. First, SSD admissibility and SSD optimality are equivalent and the definition of SSD efficiency is less ambiguous than FSD efficiency. Second, SSD efficiency can be tested directly using linear program while FSD optimality linear programming test requires mixed-integer linear programming algorithm or subsampling techniques as an initial phase. Third, FSD representative set of utility functions consist of discontinuous utility functions. This discontinuity causes a presence of the mixed-integer element.

The SSD efficiency tests in Kuosmanen [34] and Post [43] are applied in analysis of the Fama and French market portfolio relative to benchmark portfolios formed on market capitalization and book-to-market equity ratio using US stock market data. They showed that tested market portfolios were SSD inefficient. Kuosmanen [34], using a mixed-integer linear program, and Kopa & Post [32], using a linear program with subsampling initial phase, demonstrated FSD inadmissibility hence FSD non-optimality of the market portfolio. It implies the fact that no nonsatiable investor would hold the Fama and French market portfolio in the face of the considered benchmark portfolios i.e. small cap premium and the value stock premium.

The dissertation thesis is structured as follows. Chapter 2 is inspired by Kopa [25], Kopa [26], Kopa [27], Kopa [28] and deals with utility functions and their application in a portfolio selection problem. We will restrict our attention to classification of utility functions based on the Arrow - Pratt absolute risk aversion measure. It is assumed that the distribution of returns has a bounded support. The stability of expected utility of optimal portfolio in dependence on the choice of utility function is analyzed. Under the same assumptions, the stability of optimal investment strategy due to changes in Arrow - Pratt absolute risk aversion measure is discussed. The related result was proved in Kallberg & Ziemba [22] for normally distributed yields of assets using Rubinstein's measure of global risk aversion instead of absolute risk aversion measure. Applying the theory of variational analysis, see Rockafellar & Wets [47], under assumption of hypoconvergence of utility functions, the limit set of optimal portfolios is analyzed. In comparison with general stability results in stochastic programming, see Rőmisch [48], we analyze the stability with respect to perturbations of utility functions instead of changes in probability measures.

Chapter 3 is based on Kopa [29]. It develops characterizations of multiperiod risk premium. In general, risk premiums represent a way how the risk of investment possibilities can be evaluated when utility function of decision maker is known. The construction of multiperiod risk premium is based on the preference indifference between accepting a multiperiod game and rejecting this game. The possibility of accepting the game only in some time periods is included. The results in Ambarish & Kallberg [2] and Chalfant & Finkelshtain [5] are generalized for multiperiod problem. Considering directional, partial and conditional multiperiod risk premiums, the connection between multiperiod risk aversion and multiperiod risk premiums is proved.

In comparison with maximizing utility criterion the concept of stochastic dominance offers a different approach to classification of considered portfolios. The differences are also in notation for investment strategy and scenarios of yields. Following the seminal works about stochastic dominance in context of the portfolio selection problem, see Post [43] and Kuosmanen [34], we hold the usual notation for stochastic dominance. Therefore the notation in chapter 4 and chapter 5 is not the same as in previous chapters.

Chapter 4, inspired by Kopa [30] and Kopa [31], describes SSD rules concerning the portfolio selection problem. As it was shown in Ogryczak & Ruszczyński [40], CVaR corresponds to second-order stochastic dominance. Using this property for discrete probability distributions of returns, necessary and sufficient conditions for efficient and inefficient portfolios relative to all possible portfolios created from a set of assets are derived. We suggest an algorithm based on these conditions for stochastic dominance and special properties of CVaR for discrete probability distributions of returns. We derive a SSD portfolio efficiency measure which is consistent with second-order stochastic dominance. Moreover, we explore the convexity of this measure. We adopt these results for testing second-order stochastic efficiency of mean-

VaR optimal portfolios.

Finally, chapter 5, based on Kopa & Post [32] and Kopa & Post [33], develops a test for FSD efficiency of a given portfolio of choice alternatives relative to all possible portfolios. To simplify the search over all utility functions, we reformulate the problem in terms of piecewise-constant utility functions, a generalization of the Russell & Seo [50] representative utility functions for pairwise FSD tests. We provide a linear programming criterion for implementing the test. To identify the input for the linear programming problem, we may use mixed-integer linear programming or subsampling techniques. In contrast to the test by Bawa *et al.* [3], our test considers diversified portfolios in addition to the individual, undiversified alternatives, and furthermore contrary to Kuosmanen [34], our analysis is based on optimality rather than admissibility. Both features lead to a more powerful FSD efficiency test than is currently available. In Kopa & Post [32], this test is applied in analysis of Fama and French market portfolio. The differences between the Kuosmanen FSD efficiency test, the Bawa FSD efficiency test and our approach are demonstrated on numerical example.

Chapter 2

Stability of optimal portfolio in portfolio selection problem

2.1 Preliminaries

In this chapter we use utility functions, so that when solving portfolio selection problem, the optimal portfolio has the maximal expected utility. Utility functions are very useful for modeling the investor's behavior, e.g. risk aversion (or seeking). On the other hand it can be difficult to solve the portfolio selection problem for some types of utility functions. In Section 2.2, we recall an additive and multiplicative formulation of maximizing expected utility problem. The stability of optimal portfolio due to changes in Arrow-Pratt risk aversion measure in Section 2.3 will be analyzed and supplemented with application of basic results of variational analysis in Section 2.4.

Definition 2.1:

A function $u : I \rightarrow R$ is called utility function if u is finite and nondecreasing in the interval $I \subseteq R$.

The basic analysis of utility functions of Arrow [1] and Pratt [45] offers an intuitive way of looking at absolute and relative risk aversion coefficients. The Arrow-Pratt coefficient of *absolute risk aversion*, also called absolute risk averse (ARA) function, is defined as

$$r(x) = -\frac{u''(x)}{u'(x)} \tag{2.1}$$

for $x \in I$ and for an increasing, twice differentiable utility function u in I.

We assume that investor (decision maker) has utility function u and initial wealth x. Let ε be a gamble with distribution P. The investor is called *risk averse* at wealth level x if:

$$Eu(x + \varepsilon) < u(x + E\varepsilon).$$

It is easily seen that $r(x) > 0$ for every risk averse investor at wealth level x (see Ingersoll [20] for more details). According to Pratt [45], a value $\pi(x, P)$ satisfying

$$u(x + E\varepsilon - \pi(x, P)) = Eu(x + \varepsilon), \tag{2.2}$$

is called a *risk premium.* We consider only the situations where $Eu(x+\varepsilon)$ exists and is finite. The risk averse decision maker would be indifferent between accepting a risk ε and receiving the non-random amount $E\varepsilon-\pi(x, P)$. Let us consider $\pi(x, P)$ for a risk ε with a small variance σ_ε^2. Then an approximation can be proved (see Pratt [45]):

$$\pi(x, P) \approx \frac{1}{2}\sigma_\varepsilon^2 r(x + E\varepsilon). \tag{2.3}$$

According to (2.3) it is clear that ARA function is a measure of investor's *local risk aversion.*
To examine the stability of optimal portfolio due to changes in absolute risk aversion measure the following assumption will be needed:

(2.i) Utility function $u : I \to R$ is increasing and twice differentiable in the interval $I \subseteq R$.

2.2 Portfolio selection problem

Suppose that the investor wishes to allocate his wealth among assets $i = 1, ..., n$ and he chooses $\mathbf{x} = (x_1, ..., x_n)'$ to maximize the expected utility of final wealth. This model will be formulated as:

$$\begin{aligned} \max Eu(x_0 &+ \boldsymbol{\varrho}'\mathbf{x}) \\ \text{subject to}: \mathbf{1}'\mathbf{x} &= x_0 \\ x_i &\geq 0, \end{aligned} \tag{2.4}$$

x_0 ... the initial wealth
$\boldsymbol{\varrho}$ ... the random vector of returns per unit of wealth
$\mathbf{x}$ ... the investment strategy
u ... the utility function.

Assuming a multiplicative approach, we could also formulate the problem as:

$$\begin{aligned} \max Eu(\boldsymbol{\varrho}'\mathbf{x}x_0) \\ \text{subject to}: \mathbf{1}'\mathbf{x} = 1 \\ x_i \geq 0. \end{aligned} \tag{2.5}$$

Of course, it is assumed that expected values in (2.4) and (2.5) exist.

2.3 Stability of optimal portfolio

Kallberg & Ziemba [22] proved that investors with the same Rubinstein measure of global risk aversion, defined as:

$$r_g(x_0) = -\frac{x_0 E\left[u''(w)\right]}{E\left[u'(w)\right]}$$

where $w = x_0\boldsymbol{\varrho}'\mathbf{x}$, have the same optimal investment strategies, i.e. the same optimal solutions of (2.5), under the additional assumption that $\boldsymbol{\varrho}'\mathbf{x}$ is normally distributed. The Rubinstein's risk aversion measure is an example of measure of *global risk aversion*. For a deeper discussion of differences between local and global risk aversion we refer to Pratt [45].

2.3.1 Stability of optimal expected utility

Kallberg & Ziemba [22] also empirically examined the extent to which investors with "similar" ARA measures have "similar" optimal portfolios. We will formulate this result precisely for the class of probability distributions described by the following assumption:

(2.ii) There exists an interval $\langle a, b\rangle \subseteq I$ such that $P(x_0 + \boldsymbol{\varrho}'\mathbf{x} \in \langle a, b\rangle) = 1$. for any choice of $x_i \geq 0$, $i = 1, ..., n$, satisfying: $\mathbf{1}'\mathbf{x} = x_0$.

Proposition 2.2:

Let $\boldsymbol{\varrho}= (\varrho_1, \varrho_2, \dots, \varrho_n)'$ be the returns on investments satisfying (2.ii). Let $u_1(x)$, $u_2(x)$ satisfy assumption (2.i) on $\langle a, b\rangle$ and let $r_1(x)$, $r_2(x)$ be their ARA measures. Let δ be positive. If

$$|r_1(x) - r_2(x)| < \delta \tag{2.6}$$

for all $x \in \langle a, b\rangle$ then

$$Eu_1(x_0 + \boldsymbol{\varrho}'\mathbf{x}^1) - Eu_1(x_0 + \boldsymbol{\varrho}'\mathbf{x}^2) \leq [u_1(b) - u_1(a)](e^{2\delta(b-a)} - 1),$$

where $\mathbf{x}^1$, $\mathbf{x}^2$ are the optimal solutions of (2.4) for the utility functions $u_1(x)$, $u_2(x)$, respectively.

Proof: According to (2.6) we have

$$-\delta < \frac{u_2''(x)}{u_2'(x)} - \frac{u_1''(x)}{u_1'(x)} < \delta$$

for all $x \in \langle a, b\rangle$. Integrating it from a to any $y \in \langle a, b\rangle$ we obtain

$$-\delta(y - a) < \log u_2'(y) - \log u_2'(a) - \log u_1'(y) + \log u_1'(a) < \delta(y - a).$$

Set $v_1(x) = \frac{u_1(x)}{u_1'(a)}$; $v_2(x) = \frac{u_2(x)}{u_2'(a)}$ and combining it with $y \leq b$ we get

$$-\delta(b - a) < \log \frac{v_2'(y)}{v_1'(y)} < \delta(b - a)$$

or in an equivalent form

$$e^{-\delta(b-a)}v_1'(y) < v_2'(y) < e^{\delta(b-a)}v_1'(y).$$

After one more integration from a to any $x \in \langle a, b\rangle$ we have

$$e^{-\delta(b-a)}\,[v_1(x) - v_1(a)] < v_2(x) - v_2(a) < e^{\delta(b-a)}\,[v_1(x) - v_1(a)]$$

and by substitution $w_1(x) = v_1(x) - v_1(a)\,; w_2(x) = v_2(x) - v_2(a)$ we obtain

$$e^{-\delta(b-a)}w_1(x) < w_2(x) < e^{\delta(b-a)}w_1(x). \tag{2.7}$$

By the substitutions $w_1(x) = \frac{u_1(x)-u_1(a)}{u_1'(a)}$; $w_2(x) = \frac{u_2(x)-u_2(a)}{u_2'(a)}$, it is easy to check that $\mathbf{x^1}$, $\mathbf{x^2}$ are optimal solutions of (2.4) also for utility functions $w_1(x)$, $w_2(x)$. Combining (2.7) and optimality of $\mathbf{x^1}$, $\mathbf{x^2}$ we can estimate the difference of expected utilities between these optimal portfolios

$$\begin{aligned} 0 &\le E\left[w_1(x_0+\boldsymbol{\varrho}'\mathbf{x}^1) - w_1(x_0+\boldsymbol{\varrho}'\mathbf{x}^2)\right] \\ &< E\left[w_2(x_0+\boldsymbol{\varrho}'\mathbf{x}^1)e^{\delta(b-a)} - w_1(x_0+\boldsymbol{\varrho}'\mathbf{x}^2)\right] \\ &< E\left[w_2(x_0+\boldsymbol{\varrho}'\mathbf{x}^2)e^{\delta(b-a)} - w_1(x_0+\boldsymbol{\varrho}'\mathbf{x}^2)\right] \\ &< (e^{2\delta(b-a)}-1)Ew_1(x_0+\boldsymbol{\varrho}'\mathbf{x}^2) \end{aligned}$$

Since $w_1(x)$ is increasing and $x_0 + \boldsymbol{\varrho}'\mathbf{x}^2 \le b$ a.s., we can conclude

$$E\left[w_1(x_0+\boldsymbol{\varrho}'\mathbf{x}^1) - w_1(x_0+\boldsymbol{\varrho}'\mathbf{x}^2)\right] \le (e^{2\delta(b-a)}-1)w_1(b).$$

It follows immediately that

$$\begin{aligned} E\left[w_1(x_0+\boldsymbol{\varrho}'\mathbf{x}^1) - w_1(x_0+\boldsymbol{\varrho}'\mathbf{x}^2)\right] &= E\left[\frac{u_1(x_0+\boldsymbol{\varrho}'\mathbf{x}^1) - u_1(x_0+\boldsymbol{\varrho}'\mathbf{x}^2)}{u_1'(a)}\right] \\ w_1(b) &= \frac{u_1(b)-u_1(a)}{u_1'(a)}. \end{aligned}$$

Substituting it into last inequality we obtain

$$E\frac{[u_1(x_0+\boldsymbol{\varrho}'\mathbf{x}^1) - u_1(x_0+\boldsymbol{\varrho}'\mathbf{x}^2)]}{u_1'(a)} \le \frac{u_1(b)-u_1(a)}{u_1'(a)}(e^{2\delta(b-a)}-1),$$

which completes the proof. Q.E.D

More details about application of this stability result can be found in Kopa [25]. The above proposition gives information about the stability of optimal expected utility. However, Proposition 2.2 yields no information about the stability of optimal investment strategy. We will look more closely at this problem.

2.3.2 Stability of optimal investment strategy

By Lagrange's method, we obtain the necessary conditions for the optimal solution of (2.4):

$$\frac{\partial Eu(x_0 + \boldsymbol{\varrho}'\mathbf{x})}{\partial x_i} - \lambda + \eta_i = 0, \quad i = 1, 2, .., n \tag{2.8}$$

$$\eta_i x_i = 0, \quad \eta_i \geq 0, \quad i = 1, 2, .., n \tag{2.9}$$

$$\mathbf{1}'\mathbf{x} = x_0, \quad x_i \geq 0, \quad i = 1, 2, .., n.$$

From now on we make the assumptions:

(2.iii) $\boldsymbol{\varrho} = (\varrho_1, \varrho_2, \ldots, \varrho_n)'$ are the returns on investments satisfying (2.ii),

(2.iv) $u(x)$, $u_1(x), u_2(x), \ldots$ satisfy (2.i) and $r(x)$, $r_1(x), r_2(x), \ldots$ are their ARA measures,

(2.v) $\lim_{k\to\infty} r_k(x) = r(x) \quad \forall x \in \langle a, b\rangle$,

(2.vi) $u''(x), u''_k(x)$, $k = 1, 2, ...$ are continuous and negative in interval $\langle a, b\rangle$.

Set

$$\begin{aligned} X &= \{\mathbf{x} = (x_1, x_2, ..., x_n) : \mathbf{1}'\mathbf{x} = x_0, \quad x_i \geq 0, \quad i = 1, 2, .., n\} \\ X^k &= \arg\max_{\mathbf{x}\in X} Eu_k(x_0 + \boldsymbol{\varrho}'\mathbf{x}) \\ X^* &= \arg\max_{\mathbf{x}\in X} Eu(x_0 + \boldsymbol{\varrho}'\mathbf{x}). \end{aligned}$$

In this notation, X^k denote the set of optimal solutions of (2.4) using $u_k(x)$ and let us denote by $\mathbf{x}^k$ the element of X^k. Similarly, we will denote by $\mathbf{x}^*$ the element of the set of optimal solutions of (2.4) using $u(x)$.

Corollary 2.3:

Let assumptions (2.iii) - (2.v) hold. Then

$$\begin{aligned} \lim_{k\to\infty} Eu(x_0 + \boldsymbol{\varrho}'\mathbf{x}^k) - Eu(x_0 + \boldsymbol{\varrho}'\mathbf{x}^*) &= 0, \\ \lim_{k\to\infty} Eu_l(x_0 + \boldsymbol{\varrho}'\mathbf{x}^k) - Eu_l(x_0 + \boldsymbol{\varrho}'\mathbf{x}^*) &= 0, \quad l = 1, 2, ..., \end{aligned}$$

where $\mathbf{x}^k \in X^k$ and $\mathbf{x}^* \in X^*$.

Proof:

Use (2.v) and Proposition 2.2 with $\delta \to 0$. □

Proposition 2.4:

Let assumptions (2.iii) - (2.vi) hold. Then from any sequence $\mathbf{x}^1, \mathbf{x}^2, \ldots$, where $\mathbf{x}^k \in X^k$, $k = 1, 2\ldots$, a subsequence $\mathbf{x}^{k_1}, \mathbf{x}^{k_2}, \ldots$ can be extracted such that

$$\boldsymbol{\varrho}'\mathbf{x}^{k_n} \stackrel{k_n \to \infty}{\longrightarrow} \boldsymbol{\varrho}'\mathbf{x}^* \quad \text{a.s.} \quad \text{and} \;\; \mathbf{x}^* \in X^*.$$

Proof:

To simplify notation, set

$$\frac{\partial f(\mathbf{x})}{\partial \mathbf{x}} = \left(\frac{\partial f(\mathbf{x})}{\partial x_1}, \frac{\partial f(\mathbf{x})}{\partial x_2}, \ldots, \frac{\partial f(\mathbf{x})}{\partial x_n} \right). \tag{2.10}$$

By Taylor's formula, we have:

$$-Eu(x_0 + \boldsymbol{\varrho}'\mathbf{x}^k) = -Eu(x_0 + \boldsymbol{\varrho}'\mathbf{x}^*) - A + B \tag{2.11}$$

where

$$A = \frac{\partial Eu(x_0 + \boldsymbol{\varrho}'\mathbf{x}^*)}{\partial \mathbf{x}}(\mathbf{x}^k - \mathbf{x}^*) \tag{2.12}$$

$$B = \frac{1}{2}E(\mathbf{x}^k - \mathbf{x}^*)' \left(-\frac{\partial^2 u(x_0 + \boldsymbol{\varrho}'\mathbf{x}^*)}{\partial^2 \mathbf{x}} \right)_{\mathbf{x}=\overline{\mathbf{x}}} (\mathbf{x}^k - \mathbf{x}^*) \tag{2.13}$$

and $\overline{\mathbf{x}} = \alpha\mathbf{x}^* + (1-\alpha)\mathbf{x}^k$, $\alpha \in (0\,,\,1)$.
Since $\mathbf{x}^*$ is an optimal solution of (2.4), applying (2.8)-(2.9) we obtain

$$A = (\lambda.\mathbf{1} - \boldsymbol{\eta})(\mathbf{x}^k - \mathbf{x}^*) = -\boldsymbol{\eta}\mathbf{x}^k \geq 0. \tag{2.14}$$

By assumption (2.vi), $\xi > 0$ exists such that

$$B = \frac{1}{2}E(\mathbf{x}^k - \mathbf{x}^*)'\boldsymbol{\varrho}(-u''(x_0 + \boldsymbol{\varrho}'\overline{\mathbf{x}}))\boldsymbol{\varrho}'(\mathbf{x}^k - \mathbf{x}^*) \geq \frac{\xi}{2}E\left[\boldsymbol{\varrho}'(\mathbf{x}^k - \mathbf{x}^*)\right]^2 \tag{2.15}$$

Combining Corollary 2.3 with (2.11),(2.14) and (2.15) we obtain

$$Eu(x_0 + \boldsymbol{\varrho}'\mathbf{x}^*) - Eu(x_0 + \boldsymbol{\varrho}'\mathbf{x}^k) \geq \boldsymbol{\eta}\mathbf{x}^k + \frac{\xi}{2}E\left[\boldsymbol{\varrho}'(\mathbf{x}^k - \mathbf{x}^*)\right]^2 \stackrel{k \to \infty}{\longrightarrow} 0,$$

Thus

$$E\left[\boldsymbol{\varrho}'(\mathbf{x}^k - \mathbf{x}^*)\right]^2 \stackrel{k\to\infty}{\longrightarrow} 0,$$

which completes the proof. □

Since the limit of any Cauchy sequence is equal to the limit of any its convergent subsequence the following Corollary follows from Proposition 2.4.

Corollary 2.5:

Let assumptions (2.iii) - (2.vi) hold. Assume that $\mathbf{x}^1, \mathbf{x}^2, ...$ where $\mathbf{x}^k \in X^k$, $k = 1, 2...$, is a Cauchy sequence. Then

$$\boldsymbol{\varrho}'\mathbf{x}^k \stackrel{k\to\infty}{\longrightarrow} \boldsymbol{\varrho}'\mathbf{x}^* \quad \text{a.s.} \quad \text{and} \quad \mathbf{x}^* \in X^*.$$

Set

$$\begin{aligned} Y &= \{\boldsymbol{y} \in R^n : \quad \mathbf{1}'\boldsymbol{y} = 0, \quad \boldsymbol{y} \neq \mathbf{0}\}, \\ \mathcal{P} &= \{\boldsymbol{\varrho} : \exists \delta > 0 : P(\boldsymbol{\varrho} = \mathbf{0}) \leq 1 - \delta, \quad P(\boldsymbol{\varrho}'\boldsymbol{y} = 0) \leq 1 - \delta, \ \forall \boldsymbol{y} \in Y\}. \end{aligned}$$

Proposition 2.4 and Corollary 2.5 present the qualitative stability of total yields ($\boldsymbol{\varrho}'\mathbf{x}$) of optimal portfolio. To examine the stability of investment strategies of optimal portfolios, we assume that:

(2.vii) $P(\boldsymbol{\varrho} = \mathbf{0}) < 1$.

Let

$$\begin{aligned} \overline{Y}_{\boldsymbol{\varrho}} &= \{\boldsymbol{y} \in Y : \quad P(\boldsymbol{\varrho}'\boldsymbol{y} = 0) = 1\} \quad && \text{for } \boldsymbol{\varrho} \notin \mathcal{P} \\ &= \emptyset && \text{for } \boldsymbol{\varrho} \in \mathcal{P}. \end{aligned}$$

Proposition 2.6:

Let assumptions (2.iii) - (2.vii) hold. Let $\boldsymbol{\varrho} \in \mathcal{P}$. Then

(i) portfolio selection problem (2.4) has a unique solution when using $u(x), u_k(x)$, $k = 1, 2, ...$

(ii) from the sequence $\mathbf{x}^1, \mathbf{x}^2, ...,$ where $\mathbf{x}^k \in X^k$, $k = 1, 2...$, a Cauchy subsequence $\mathbf{x}^{l_1}, \mathbf{x}^{l_2}, ...$ can be extracted such that

$$\mathbf{x}^{l_n} \stackrel{l_n\to\infty}{\longrightarrow} \mathbf{x}^* \quad \text{and} \quad \mathbf{x}^* \in X^*.$$

Proof:

(i) Assume that $\mathbf{x}^k \in X^k$ and $\overline{\mathbf{x}}^k \in X^k$. Then

$$Eu_k(x_0 + \boldsymbol{\varrho}'\mathbf{x}^k) - Eu_k(x_0 + \boldsymbol{\varrho}'\overline{\mathbf{x}}^k) = 0. \tag{2.16}$$

By assumption (2.vi), $\xi > 0$ exists such that

$$-u_k''(x_0 + \boldsymbol{\varrho}'\mathbf{x}) \geq \xi, \; \forall \mathbf{x} \in X.$$

As in the proof of Proposition 2.4, by Taylor's formula, we obtain

$$0 = Eu_k(x_0 + \boldsymbol{\varrho}'\mathbf{x}^k) - Eu_k(x_0 + \boldsymbol{\varrho}'\overline{\mathbf{x}}^k) \geq \boldsymbol{\eta}\overline{\mathbf{x}}^k + \frac{\xi}{2}E\left[\boldsymbol{\varrho}'(\overline{\mathbf{x}}^k - \mathbf{x}^k)\right]^2.$$

Since $\boldsymbol{\eta}\overline{\mathbf{x}}^k \geq 0$ and $\xi > 0$ we have

$$E\left[\boldsymbol{\varrho}'(\overline{\mathbf{x}}^k - \mathbf{x}^k)\right]^2 = 0.$$

Hence

$$\boldsymbol{\varrho}'(\overline{\mathbf{x}}^k - \mathbf{x}^k) = 0 \quad \text{a.s.}$$

Since $\boldsymbol{\varrho} \in \mathcal{P}$, we obtain: $\overline{\mathbf{x}}^k = \mathbf{x}^k$.
In the same manner we can see that portfolio selection problem (2.4) has a unique solution using $u(x)$.

(ii) Proposition 2.4 shows that from any sequence $\mathbf{x}^1, \mathbf{x}^2, \ldots$, where $\mathbf{x}^k \in X^k$, $k = 1, 2 \ldots$, a subsequence $\mathbf{x}^{k_1}, \mathbf{x}^{k_2}, \ldots$ can be extracted such that

$$\boldsymbol{\varrho}'\mathbf{x}^{k_n} \overset{k_n \to \infty}{\longrightarrow} \boldsymbol{\varrho}'\mathbf{x}^* \quad \text{a.s.} \quad \text{and} \;\; \mathbf{x}^* \in X^*.$$

Let $\mathbf{x}^{l_1}, \mathbf{x}^{l_2}, \ldots$ be a Cauchy subsequence of the sequence $\mathbf{x}^{k_1}, \mathbf{x}^{k_2}, \ldots$. Then

$$\boldsymbol{\varrho}'\mathbf{x}^{l_n} \overset{l_n \to \infty}{\longrightarrow} \boldsymbol{\varrho}'\mathbf{x}^* \quad \text{a.s.} \tag{2.17}$$

Let

$$\overline{\mathbf{x}} = \lim_{l_n \to \infty} \mathbf{x}^{l_n}$$

then

$$\boldsymbol{\varrho}'\mathbf{x}^{l_n} \overset{l_n \to \infty}{\longrightarrow} \boldsymbol{\varrho}'\overline{\mathbf{x}} \quad \text{a.s.}$$

Combining it with (2.17) we have $\boldsymbol{\varrho}'(\overline{\mathbf{x}} - \mathbf{x}^*) = 0$ a.s. Since $\boldsymbol{\varrho} \in \mathcal{P}$, we obtain: $\overline{\mathbf{x}} = \mathbf{x}^*$, and the proof is complete. □

Proposition 2.7:

Let assumptions (2.iii) - (2.vii) hold. Let X^* be a singleton. Then from the sequence $\mathbf{x}^1, \mathbf{x}^2, ...$, where $\mathbf{x}^k \in X^k$, $k = 1, 2...$, a Cauchy subsequence $\mathbf{x}^{l_1}, \mathbf{x}^{l_2}, ...$ can be extracted such that

$$\mathbf{x}^{l_n} \stackrel{l_n \to \infty}{\longrightarrow} \mathbf{x}^* \quad \text{and} \quad \mathbf{x}^* \in X^*.$$

Proof:

Proposition 2.4 shows that from any sequence $\mathbf{x}^1, \mathbf{x}^2, ...$, where $\mathbf{x}^k \in X^k$, $k = 1, 2...$, a subsequence $\mathbf{x}^{k_1}, \mathbf{x}^{k_2}, ...$ can be extracted such that

$$\boldsymbol{\varrho}'\mathbf{x}^{k_n} \stackrel{k_n \to \infty}{\longrightarrow} \boldsymbol{\varrho}'\mathbf{x}^* \quad \text{a.s.} \quad \text{and} \quad \mathbf{x}^* \in X^*.$$

Let $\mathbf{x}^{l_1}, \mathbf{x}^{l_2}, ...$ be a Cauchy subsequence of the sequence $\mathbf{x}^{k_1}, \mathbf{x}^{k_2}, ...$ Then

$$\boldsymbol{\varrho}'\mathbf{x}^{l_n} \stackrel{l_n \to \infty}{\longrightarrow} \boldsymbol{\varrho}'\mathbf{x}^* \quad \text{a.s.} \tag{2.18}$$

Let

$$\overline{\mathbf{x}} = \lim_{l_n \to \infty} \mathbf{x}^{l_n}$$

then

$$\boldsymbol{\varrho}'\mathbf{x}^{l_n} \stackrel{l_n \to \infty}{\longrightarrow} \boldsymbol{\varrho}'\overline{\mathbf{x}} \quad \text{a.s.}$$

Combining it with (2.18) we have $\boldsymbol{\varrho}'\overline{\mathbf{x}} = \boldsymbol{\varrho}'\mathbf{x}^*$ a.s. Hence $\overline{\mathbf{x}} \in X^*$. Since X^* is a singleton, we obtain: $\overline{\mathbf{x}} = \mathbf{x}^*$, and the proof is complete. □

We recall the definition of the *Hausdorf distance* between two sets, A and B:

$$d_h(A, B) = \max\{\max_{a \in A} d(a, B), \max_{b \in B} d(b, A)\} \quad \text{where} \quad d(p, Q) = \min_{q \in Q} d(p, q)$$

and $d(p, q)$ is the Euclidean distance from p to q. To prove the main stability result the following lemma describing the structure of sets of optimal solutions will be needed.

Lemma 2.8:

Assume that $\mathbf{x}^* \in X^*$, $\mathbf{x}^k \in X^k$, $k = 1, 2...$, are fixed.
Let $Z^k = \{\boldsymbol{z} \in R^n : \boldsymbol{z} = \mathbf{x}^k + \boldsymbol{y};\ \boldsymbol{y} \in \overline{Y}_{\boldsymbol{\varrho}}\}$, $\quad k = 1, 2, ...,$
$Z^* = \{\boldsymbol{z} \in R^n : \boldsymbol{z} = \mathbf{x}^* + \boldsymbol{y};\ \boldsymbol{y} \in \overline{Y}_{\boldsymbol{\varrho}}\}$.
Then $X^k = Z^k \bigcap X$, $\quad k = 1, 2, ...$ and $X^* = Z^* \bigcap X$.

Proof:

Let $\mathbf{x} \in Z^k \bigcap X$. If $\overline{Y}_{\boldsymbol{\varrho}} \neq \emptyset$ then there exists $\boldsymbol{y} \in \overline{Y}_{\boldsymbol{\varrho}}$ such that $\mathbf{x} = \mathbf{x}^k + \boldsymbol{y}$. Since $\boldsymbol{\varrho}'\boldsymbol{y} = 0$ a.s., we obtain $\boldsymbol{\varrho}'\mathbf{x} = \boldsymbol{\varrho}'\mathbf{x}^k$ a.s. Thus $Eu_k(x_0 + \boldsymbol{\varrho}'\mathbf{x}) = Eu_k(x_0 + \boldsymbol{\varrho}'\mathbf{x}^k)$. Since $\mathbf{x} \in X$, the last equality yields $\mathbf{x} \in X^k$. Therefore $X^k \supseteq Z^k \bigcap X$.
Let $\mathbf{x} \in X^k$. Then according to the proof of Proposition 2.6 (ii), we have

$$\boldsymbol{\varrho}'(\mathbf{x} - \mathbf{x}^k) = 0 \quad \text{a.s.}$$

Since $\mathbf{x} = \mathbf{x}^k + (\mathbf{x} - \mathbf{x}^k)$ and $(\mathbf{x} - \mathbf{x}^k) \in \overline{Y}_{\boldsymbol{\varrho}}$, we obtain $\mathbf{x} \in Z^k$. By assumption, $\mathbf{x} \in X$ hence $\mathbf{x} \in Z^k \bigcap X$. Therefore $X^k \subseteq Z^k \bigcap X$. In the same manner it is easy to check that $X^* = Z^* \bigcap X$, and the proof is complete. □

Theorem 2.9:

Let assumptions (2.iii) - (2.vii) hold. Then

$$\limsup_{k \to \infty} d_h(X^k, X^*) = 0.$$

Proof:

Proposition 2.4 shows that from any sequence $\mathbf{x}^1, \mathbf{x}^2, ...$, where $\mathbf{x}^k \in X^k$, $k = 1, 2...$, a subsequence $\mathbf{x}^{k_1}, \mathbf{x}^{k_2}, ...$ can be extracted such that

$$\mathbf{x}^{k_n} \overset{k_n \to \infty}{\longrightarrow} \boldsymbol{\varrho}'\mathbf{x}^* \quad \text{a.s.} \quad \text{and} \quad \mathbf{x}^* \in X^*.$$

Since X is compact set, there exists a Cauchy subsequence $\mathbf{x}^{l_1}, \mathbf{x}^{l_2}, ...$ of the sequence $\mathbf{x}^{k_1}, \mathbf{x}^{k_2}, ...$ Let

$$\overline{\mathbf{x}} = \lim_{l_n \to \infty} \mathbf{x}^{l_n} \quad \text{then} \quad \boldsymbol{\varrho}'\mathbf{x}^{l_n} \overset{l_n \to \infty}{\longrightarrow} \boldsymbol{\varrho}'\overline{\mathbf{x}} \quad \text{a.s.}$$

Proposition 2.4 now implies $\boldsymbol{\varrho}'\mathbf{x}^{l_n} \overset{l_n \to \infty}{\longrightarrow} \boldsymbol{\varrho}'\mathbf{x}^*$ a.s. Combining these limits we obtain: $\boldsymbol{\varrho}'\overline{\mathbf{x}} = \boldsymbol{\varrho}'\mathbf{x}^*$ a.s. Therefore $\overline{\mathbf{x}} \in X^*$. We have just proved that

$$\limsup_{k \to \infty} \max_{\mathbf{x}^k \in X^k} d(\mathbf{x}^k, X^*) = 0.$$

Applying Proposition 2.4 for any subsequence of $\mathbf{x}^1, \mathbf{x}^2, ...$, where $\mathbf{x}^k \in X^k$, it remains to prove that for any $\mathbf{x}^* \in X^*$ a sequence $\mathbf{x}^1, \mathbf{x}^2, ...$, where $\mathbf{x}^k \in X^k$,

$k = 1, 2...$, exists such that at least one Cauchy subsequence of this sequence converges to $\mathbf{x}^*$. We will construct such sequence.
Choose $\overline{\mathbf{x}}^* \in X^*$. Consider a sequence $\mathbf{x}^1, \mathbf{x}^2, ...$ where $\mathbf{x}^k \in X^k$, $k = 1, 2...$, and a Cauchy subsequence $\mathbf{x}^{l_1}, \mathbf{x}^{l_2}, ...$ Set

$$\mathbf{x}^* = \lim_{l_n \to \infty} \mathbf{x}^{l_n}$$

If $\overline{\mathbf{x}}^* = \mathbf{x}^*$ then the construction follows immediately. In the opposite case, by the lemma above, there exists $\boldsymbol{y^*} \in \overline{Y}_{\boldsymbol{\varrho}}$ such that $\overline{\mathbf{x}}^* = \mathbf{x}^* + \boldsymbol{y^*}$. Let us analyze two cases:

(i) If $\overline{x}_i^* > 0, \quad \forall i \in \{1, 2, .., n\}$ then define $\overline{\mathbf{x}}^k = \mathbf{x}^k + \boldsymbol{y^*}, k = 1, 2, ...$
Since $\mathbf{x}^{l_n} \overset{l_n \to \infty}{\longrightarrow} \mathbf{x}^*$, we obtain $\overline{\mathbf{x}}^{l_n} \overset{l_n \to \infty}{\longrightarrow} \overline{\mathbf{x}}^*$. Since $\overline{\mathbf{x}}^*$ is a positive vector, there exists n_0 such that: $\overline{\mathbf{x}}^{l_n} \in X, \forall n \geq n_0$. Finally, Lemma 2.8 implies $\overline{\mathbf{x}}^k \in Z^k$. Hence $\overline{\mathbf{x}}^1, \overline{\mathbf{x}}^2, ...$ is the sequence we wanted to find.

(ii) Let $I = \{i \in \{1, 2, .., n\} : \overline{x}_i^* = 0, y_i < 0 \}$. Define $\overline{\mathbf{x}}^k = \mathbf{x}^k + \boldsymbol{y^k}$, $k = 1, 2, ...$ where $\boldsymbol{y^k} = \boldsymbol{y^*}(1 - \alpha_k)$. It is clear that $\boldsymbol{y^k} \in \overline{Y}_{\boldsymbol{\varrho}}$ and $\overline{\mathbf{x}}^k \in Z^k$, $\forall k \in N$. Let

$$\alpha_k = \max_{i \in I} \frac{x_i^k + y_i^*}{y_i^*}$$

then $\alpha_{l_n} \to 0$. Thus $\overline{\mathbf{x}}^{l_n} \overset{l_n \to \infty}{\longrightarrow} \overline{\mathbf{x}}^*$. Since $y_i^* < 0, \forall i \in I$ it is easy to check that this choice of α_k guarantees that $\overline{x}_i^k \geq 0$, $\forall i \in I$. If $\overline{x}_i^* = 0$ and $y_i^* = 0$ then it follows immediately that $\overline{x}_i^k \geq 0$. If $\overline{x}_i^* > 0$ we apply the similar arguments to the case (i). Hence there exists n_0 such that: $\overline{\mathbf{x}}^{l_n} \in X^{l_n}, \forall n \geq n_0$. Thus $\overline{\mathbf{x}}^1, \overline{\mathbf{x}}^2, ...$ is the sequence we wanted to find, and the proof is complete. □

Since Hausdorf distance is always non-negative

$$0 = \limsup_{k \to \infty} d_h(X^k, X^*) = \liminf_{k \to \infty} d_h(X^k, X^*)$$

which together with Theorem 2.9 implies the following result.

Corollary 2.10:

Let assumptions (2.iii) - (2.vii) hold. Then

$$\lim_{k \to \infty} d_h(X^k, X^*) = 0.$$

To derive these stability results, we assumed twice differentiability of utility functions (2.iv) and convergence of ARA measures (2.v). If we drop the assumption of differentiability, i.e. the assumption of existence of ARA measures, we can follow Rockafellar & Wets [47] and apply assumption of hypoconvergence of expected utility functions.

Comparing these two approaches, when assuming convergence of ARA measures, the full information about utility functions of the decision maker is not needed. This advantage can be used in the situation when we have full information about ARA measure of decision maker, but the portfolio selection problem can not be solved, because it is impossible to express analytically the exact form of utility function. In this case we can use approximation by another suitable utility function. The stability results in Proposition 2.7, Theorem 2.9 or Corollary 2.10 can be useful for examination of quality of the approximation. The following example will demonstrate this situation where ARA measure of decision maker can be estimated in various ways, for example, from risk premium using (2.3).

Example 2.11:

Consider a decision maker with unknown utility function. Let $x_0 = 1$, $\boldsymbol{\varrho} = (1,3)'$ and $\boldsymbol{\varrho} = (1,0)'$ with equal probabilities. Assume K time instants where $K \in \mathcal{N}$ is large enough. In each moment k we estimate his ARA measure from the available data till this moment. We obtain the sequence of ARA measures:

$$r_k(x) = e^{-\frac{1}{k}x^2}.$$

Since $r_k(x) \overset{k\to\infty}{\longrightarrow} 1$ the limit utility function is: $u(x) = -e^{-x}$. Since the exact form of utility functions corresponding to estimated ARA measures can not be derived, we can use the limit utility function. Thus we can solve the problem:

$$\begin{aligned} \max -\frac{1}{2}e^{-(1+x_1+3x_2)} \quad &- \quad \frac{1}{2}e^{-(1+x_1)} \\ \text{subject to: } x_1 + x_2 \quad &= \quad 1 \\ x_i \quad &\geq \quad 0, \qquad i = 1,2 \end{aligned}$$

and the optimal solution of this problem is $x_1^* = 1 - \frac{\log(2)}{3}, x_2^* = \frac{\log(2)}{3}$. In spite of the fact that the optimal solutions of portfolio selection problems corresponding to estimated ARA measures are not known, applying Proposition

2.7, every Cauchy sequence of these optimal solutions converges to (x_1^*, x_2^*). Thus (x_1^*, x_2^*) can be regarded as an approximation of optimal solution of the original portfolio selection problem.

Assuming hypoconvergence of expected utility functions, we can obtain a stability result for larger class of utility functions than the class given by (2.iv). On the other hand, to verify this assumption, the full information about utility functions is needed which can be unreachable as demonstrated in Example 2.11. Typically, a verification of assumptions (2.iv) and (2.v) is less demanding than a verification of the assumption of hypoconvergence.

2.4 Variational analysis approach

Firstly, we recall the basic terms of variational analysis. In this approach, we consider expected utility as a function of investment strategy i.e.

$$f(\mathbf{x}) = -Eu(x_0 + \boldsymbol{\varrho}'\mathbf{x}).$$

Definition 2.12:

(i) The function $f : R^n \to R$ is *lower semicontinuous* (lsc) at $\overline{\mathbf{x}}$ if

$$\liminf_{\mathbf{x}\to\overline{\mathbf{x}}} f(\mathbf{x}) \geq f(\overline{\mathbf{x}})$$

and lower semicontinuous on R^n if this holds for every $\overline{\mathbf{x}} \in R^n$. The function $f : R^n \to R$ is *upper semicontinuous* (usc) at $\overline{\mathbf{x}}$ if $-f$ is lsc at $\overline{\mathbf{x}}$ and upper semicontinuous on R^n if $-f$ is lower semicontinuous on R^n.

(ii) For $f : R^n \to R$, the *epigraph* of f is the set

$$\text{epi} f = \{(x, a) \in R^n \times R | a \geq f(\mathbf{x})\}.$$

(iii) For $f : R^n \to R$, the *level set* of f is the set

$$\text{lev}_\alpha f = \{\mathbf{x} \in R^n | f(\mathbf{x}) \leq \alpha\}.$$

The epigraph consists of all the points of R^{n+1} lying on or above the graph of f. For α finite, the level sets correspond to the "horizontal cross section" of the epigraph. According to Rockafellar & Wets [47], Th. 1.6 the following properties of a function f are equivalent:

(a) f is lower semicontinuous on R^n;

(b) epi$f$ is closed in R^{n+1};

(c) $\text{lev}_\alpha f$ is a closed set in R^n for all α.

The basic tool for epiconvergence approach is definition of a limit of a sequence of sets $\{C^k\}_{k\in\mathcal{N}}$ and eventually level-bounded sequence using the following notation of index sets:

$$\begin{aligned} N_\infty &= \{N \subset \mathcal{N} | \mathcal{N}\backslash N \text{ is finite}\} \\ N_\infty^\sharp &= \{N \subset \mathcal{N} | N \text{ is infinite}\} \end{aligned}$$

where $\mathcal{N}$ represents the set of natural numbers. Since $N_\infty^\sharp$ consists of all subsequences of $\mathcal{N}$ it is easily seen that $N_\infty \subset N_\infty^\sharp$.

Definition 2.13:

(i) For a sequence $\{C^k\}_{k\in\mathcal{N}}$ of subsets of R^n, the *outer limit* is the set :

$$\limsup_{k\to\infty} C^k = \{x \,|\, \exists N \in N_\infty^\sharp,\, \exists x^k \in C^k, k \in N \text{ with } x^k \xrightarrow{N} x\}.$$

while the *inner limit* of $\{C^k\}_{k\in\mathcal{N}}$ is the set:

$$\liminf_{k\to\infty} C^k = \{x \,|\, \exists N \in N_\infty,\, \exists x^k \in C^k, \quad k \in N \text{ with } x^k \xrightarrow{N} x\}.$$

The *limit* of the sequence $\{C^k\}_{k\in\mathcal{N}}$ exists, if the outer and inner limit sets are equal:

$$\lim_{k\to\infty} C^k := \limsup_{k\to\infty} C^k = \liminf_{k\to\infty} C^k.$$

(ii) For any sequence $\{f_k\}_{k\in\mathcal{N}}$ of functions on R^n, the *lower epi-limit* $(e - \liminf_k f_k)$ is the function having as its epigraph the outer limit of the sequence of sets epi f_k:

$$\text{epi}(e - \liminf_k f_k) = \limsup_k(\text{epi}(f_k)).$$

The *upper epi-limit* $(e - \limsup_k f_k)$ is the function having as its epigraph the inner limit of the sequence of sets epi f_k:

$$\text{epi}(e - \limsup_k f_k) = \liminf_k(\text{epi}(f_k)).$$

When upper and inner limit coincide, the *epi-limit* ($e - \lim_k f_k$) is said to exist: $e - \lim_k f_k = e - \liminf_k f_k = e - \limsup_k f_k$. In this event the functions f_k are said to *epi-converge* to f ($f_k \overset{e}{\longrightarrow} f$).

(iii) A sequence $\{f_k\}_{k \in \mathcal{N}}$ of functions on R^n is *eventually level-bounded* if for each $\alpha \in R$ the sequence of level sets ($\text{lev}_\alpha f_k$) is eventually bounded, i.e. for some index set $N \in N_\infty$ the set $\bigcup_{k \in N} \text{lev}_\alpha f_k$ is bounded.

Directly from the definition of epi-limit and from the definition of the limit of sets (epigraphs) we can see that: $e - \liminf_k f_k \leq e - \limsup_k f_k$ and $f_k \overset{e}{\longrightarrow} f \Leftrightarrow \text{epi} f_k \longrightarrow \text{epi} f$. Applying Rockafellar & Wets [47], Th. 7.33 in the context of the portfolio selection problem we can conclude the following stability result.

Theorem 2.14:

Let $f_k(\mathbf{x}) = -Eu_k(x_0 + \boldsymbol{\varrho}'\mathbf{x})$ and $f(\mathbf{x}) = -Eu(x_0 + \boldsymbol{\varrho}'\mathbf{x})$. Suppose that the sequence $\{f_k\}_{k \in \mathcal{N}}$ is eventually level-bounded, and $f_k \overset{e}{\longrightarrow} f$ with f_k lsc. Then

(i) $\limsup_k X^k \subset X^*$

(ii) $Eu_k(x_0 + \boldsymbol{\varrho}'\mathbf{x}^k) \longrightarrow Eu(x_0 + \boldsymbol{\varrho}'\mathbf{x}^*)$ for any $\mathbf{x}^k \in X^k$ and $\mathbf{x}^* \in X^*$.

Reformulating the assumptions of Theorem 2.14 in terms of utility functions we obtain the following result.

Corollary 2.15:

Suppose the interval I is bounded. Let $u : I \longrightarrow R$ and $u_k : I \longrightarrow R$, $k = 1, 2, \ldots$, be usc utility functions with $-u_k \overset{e}{\longrightarrow} -u$. Let $\boldsymbol{\varrho}$ satisfies assumption (2.ii). Then

(i) $\limsup_k X^k \subset X^*$

(ii) $Eu_k(x_0 + \boldsymbol{\varrho}'\mathbf{x}^k) \longrightarrow Eu(x_0 + \boldsymbol{\varrho}'\mathbf{x}^*)$ for any $\mathbf{x}^k \in X^k$ and $\mathbf{x}^* \in X^*$.

Proof:

Since the union of domains of u, u_k, $k = 1, 2, \ldots$ is bounded and the support of $\boldsymbol{\varrho}$ is bounded the union of all level sets of expected utility functions $(\bigcup_{k\in\mathcal{N}} \text{lev}_\alpha[-Eu_k(x_0 + \boldsymbol{\varrho}'\mathbf{x})])$ is bounded for any choice of $\alpha \in R$, i.e. the sequence $\{-Eu_k(x_0 + \boldsymbol{\varrho}'\mathbf{x})\}_{k\in\mathcal{N}}$ is eventually level-bounded.

To show that $-u_k \xrightarrow{e} -u$ implies $-Eu_k(x_0+\boldsymbol{\varrho}'\mathbf{x}) \xrightarrow{e} -Eu(x_0+\boldsymbol{\varrho}'\mathbf{x})$ we apply Rockafellar & Wets [47], Th. 7.2. dealing with sufficient and necessary condition of epiconvergence: $f_k \xrightarrow{e} f$ if and only if at each point $\mathbf{x}$ both following statements hold true:

(a) $\liminf_k f_k(\mathbf{x}^k) \geq f(\mathbf{x})$ for every sequence $\mathbf{x}^k \longrightarrow \mathbf{x}$

(b) $\limsup_k f_k(\mathbf{x}^k) \leq f(\mathbf{x})$ for some sequence $\mathbf{x}^k \longrightarrow \mathbf{x}$.

Using Fatou's lemma and assumption $-u_k \xrightarrow{e} -u$, especially (a), we obtain:

$$\begin{aligned}\liminf_k \int_{R^n} -u_k(x_0 + \boldsymbol{\varrho}'\mathbf{x}^k)dP(\boldsymbol{\varrho}) &\geq \int_{R^n} \liminf_k - u_k(x_0 + \boldsymbol{\varrho}'\mathbf{x}^k)dP(\boldsymbol{\varrho}) \\ &\geq \int_{R^n} -u(x_0 + \boldsymbol{\varrho}'\mathbf{x})dP(\boldsymbol{\varrho})\end{aligned}$$

for every sequence $\mathbf{x}^k \longrightarrow \mathbf{x}$ which proves (a) with $f_k(\mathbf{x}^k) = -Eu_k(x_0+\boldsymbol{\varrho}'\mathbf{x}^k)$ and $f(\mathbf{x}) = -Eu(x_0+\boldsymbol{\varrho}'\mathbf{x})$. In the same manner, for some sequence $\mathbf{x}^k \longrightarrow \mathbf{x}$ we have:

$$\begin{aligned}\limsup_k \int_{R^n} -u_k(x_0 + \boldsymbol{\varrho}'\mathbf{x}^k)dP(\boldsymbol{\varrho}) &\leq \int_{R^n} \limsup_k - u_k(x_0 + \boldsymbol{\varrho}'\mathbf{x}^k)dP(\boldsymbol{\varrho}) \\ &\leq \int_{R^n} -u(x_0 + \boldsymbol{\varrho}'\mathbf{x})dP(\boldsymbol{\varrho}),\end{aligned}$$

i.e. (b) holds true and the proof of epiconvergence of sequence $\{-Eu_k(x_0 + \boldsymbol{\varrho}'\mathbf{x})\}_{k\in\mathcal{N}}$ is complete.

Finally, lower semicontinuity of $-Eu_k(x_0+\boldsymbol{\varrho}'\mathbf{x})$, $k = 1, 2, \ldots$, and $-Eu(x_0+\boldsymbol{\varrho}'\mathbf{x})$ will be derived. From the assumption of upper semicontinuity of u and u_k, $k = 1, 2, \ldots$ and Fatou's lemma we conclude:

$$\begin{aligned}\liminf_l \int_{R^n} -u_k(x_0 + \boldsymbol{\varrho}'\mathbf{x}^l)dP(\boldsymbol{\varrho}) &\geq \int_{R^n} \liminf_l - u_k(x_0 + \boldsymbol{\varrho}'\mathbf{x}^l)dP(\boldsymbol{\varrho}) \\ &\geq \int_{R^n} -u_k(x_0 + \boldsymbol{\varrho}'\mathbf{x})dP(\boldsymbol{\varrho}), \qquad k = 1, 2 \ldots\end{aligned}$$

$$\begin{aligned}\liminf_l \int_{R^n} -u(x_0 + \boldsymbol{\varrho}'\mathbf{x}^l)dP(\boldsymbol{\varrho}) &\geq \int_{R^n} \liminf_l - u(x_0 + \boldsymbol{\varrho}'\mathbf{x}^l)dP(\boldsymbol{\varrho}) \\ &\geq \int_{R^n} -u(x_0 + \boldsymbol{\varrho}'\mathbf{x})dP(\boldsymbol{\varrho})\end{aligned}$$

for every sequence $\mathbf{x}^l \longrightarrow \mathbf{x}$ which completes the proof. $\square$

Since x_0 is a given parameter, $\boldsymbol{\varrho}$ has a bounded support and the feasible set of investment strategies is compact, assumption of boundedness of interval I represents no addition restriction.

Chapter 3

Multivariate and multiperiod risk premiums

3.1 Preliminaries

In Chapter 2, the univariate risk premium was considered as an amount which is a risk averse investor willing to pay to eliminate the risk in a fair gamble. The classical Arrow-Pratt approach assumes certain (non-random) level of initial wealth. The generalization of this notion to random initial wealth was introduced in Ross [49] where a stronger measure of risk aversion was presented. Another extension of the Arrow-Pratt results for the case of random initial wealth was suggested in Kihlstrom & Romer & Williams [24].

In Section 3.2 and 3.3 of this chapter, we summarize the results of Ambarish & Kallberg [2], Chalfant & Finkelshtain [5], Duncan [9] and Kihlstrom & Mirman [23] with respect to the characterization of risk premiums for multivariate (multiattribute) risk. In Section 3.4, we develop a multiperiod risk premium. For this construction of risk premium in multiperiod problem, the basic relationship to multivariate risk aversion is proved. Finally, in section 3.5, several generalizations of multiperiod risk premium notion are suggested when some of considered assumptions are relaxed.

3.2 Multivariate risk premium

Suppose a decision maker with utility function $u(\mathbf{w})$ and with initial wealth $\mathbf{w} = (w_1, w_2, ..., w_n)'$. We can interpret $\mathbf{w}$ as a vector of n commodities. Assume that $u(\mathbf{w})$ is continuous and increasing in all variables. In this section, we follow Duncan [9] in assuming that $\mathbf{w}$ is non-random. By analogy to one-dimensional case, the multivariate risk premium $\boldsymbol{\pi}$ is given by the equation:

$$u(\mathbf{w} + E_{\mathbf{x}}\mathbf{x} - \boldsymbol{\pi}) = E_{\mathbf{x}}u(\mathbf{w} + \mathbf{x})$$

for a given multidimensional risk $\mathbf{x}$. The vector $\boldsymbol{\pi}$ is a function of initial wealth and probability distribution of multidimensional risk. The uniqueness of risk premium in univariate case was proved in Pratt [45]. It is clear that if $n > 1$ than $\boldsymbol{\pi}$ is not unique and using asymptotic characterization, we can conclude that $\boldsymbol{\pi}$ lies in an n-dimensional hyperplane. We refer to Duncan [9] for more details. As in the univariate case, we define the risk aversion at level $\mathbf{w}$ such that:

$$u(\mathbf{w} + E_{\mathbf{x}}\mathbf{x}) > E_{\mathbf{x}}u(\mathbf{w} + \mathbf{x})$$

for any given gamble $\mathbf{x}$. The interpretation is that the utility of having certain quantities $\mathbf{w} + E_{\mathbf{x}}\mathbf{x}$ is preferred to the expected utility of having uncertain quantities $\mathbf{w} + \mathbf{x}$. It is easy to show that if u is concave than there exists a nonnegative risk premium for any gamble and consequently u fulfills the condition of risk aversion (see Duncan [9]).

3.3 Multivariate risk premium with random initial wealth

The generalization of Duncan [9] for random initial wealth was introduced in Ambarish & Kallberg [2]. Similarly to the case of non-random initial wealth we are interested in determining a multivariate risk premium $\boldsymbol{\pi}$ such that the decision maker is indifferent between two random variables: $(\mathbf{w} - \boldsymbol{\pi})$ and $(\mathbf{w} + \mathbf{x})$. Observe that, while the uncertainty in $\mathbf{x}$ can be eliminated (insured) by $\boldsymbol{\pi}$, there is no insurance against the risk in $\mathbf{w}$, because the final wealth will be a random vector in both cases. We follow Ambarish & Kallberg [2] in defining the *multivariate risk premium* by

$$E_{\mathbf{w}}u(\mathbf{w} + E_{\mathbf{x}}\mathbf{x} - \boldsymbol{\pi}) = E_{\mathbf{w},\mathbf{x}}u(\mathbf{w} + \mathbf{x}). \tag{3.1}$$

In this notion, multivariate risk premium is a function of probability distribution of a gamble $\mathbf{x}$ and probability distribution of initial wealth $\mathbf{w}$. However it does not depend on the realization of $\mathbf{w}$. This is a disadvantage of this approach. It was demonstrated that also in the case of a random initial wealth $\boldsymbol{\pi}$ lies in an n-dimensional hyperplane in asymptotic characterization. See Ambarish & Kallberg [2] for more details. By analogy to univariate and multivariate case with non-random initial wealth, the condition of multivariate risk aversion can be given by the formula:

$$E_{\mathbf{w}}u(\mathbf{w} + E_{\mathbf{x}}\mathbf{x}) > E_{\mathbf{w},\mathbf{x}}u(\mathbf{w} + \mathbf{x}). \qquad (3.2)$$

However, contrary to the univariate case, concavity of utility function does not guarantee a risk aversion. The multivariate risk aversion given by (3.2) depends on the gamble, as we can see in the following example. Therefore the risk aversion defined by (3.2) has to be called a *multivariate risk aversion at wealth level* $\mathbf{w}$ *with respect to gamble* $\mathbf{x}$.

Example 3.1:

Let $u(\mathbf{w}) = \log(w_1+w_2)$ and $(w_1, w_2, x_1, x_2) = (\frac{1}{2}, 0, 1, -\frac{1}{2})$ or $(1, \frac{1}{2}, -1, \frac{1}{2})$ with equal probabilities. Consider $u(\mathbf{w}) = \log(w_1 + w_2)$. It is clear that $E_{\mathbf{x}}x_1 = E_{\mathbf{x}}x_2 = 0$ and

$$\begin{aligned} E_{\mathbf{w},\mathbf{x}}u(\mathbf{w} + \mathbf{x}) &= E_{\mathbf{w},\mathbf{x}} \log(w_1 + w_2 + x_1 + x_2) \\ &= \frac{1}{2}\log\left(\frac{1}{2} + 0 + 1 - \frac{1}{2}\right) + \frac{1}{2}\log\left(1 + \frac{1}{2} - 1 + \frac{1}{2}\right) \\ &= 0 \\ E_{\mathbf{w}}u\left(\mathbf{w} + E_{\mathbf{x}}\mathbf{x}\right) &= E_{\mathbf{w}} \log(w_1 + w_2) = \frac{1}{2}\log\left(\frac{1}{2} + 0\right) + \frac{1}{2}\log\left(1 + \frac{1}{2}\right) \\ &= \frac{1}{2}\log\left(\frac{3}{4}\right) < 0. \end{aligned}$$

Thus

$$E_{\mathbf{w},\mathbf{x}}u(\mathbf{w} + \mathbf{x}) > E_{\mathbf{w}}u\left(\mathbf{w} + E_{\mathbf{x}}\mathbf{x}\right).$$

It is easy to check that u is concave and increasing in w_1 and also in w_2. We can see that the correlation between $\mathbf{w}$ and $\mathbf{x}$ can cause the fact that the condition of risk aversion (3.2) does not hold even if u is concave and increasing in each variable. Moreover, we will see that considering the same utility function and initial wealth, the condition (3.2) is fulfilled for another gamble $\mathbf{x}$. If

$u(\mathbf{w}) = \log(w_1 + w_2)$ and $(w_1, w_2, x_1, x_2) = (\frac{1}{2}, 0, \frac{1}{2}, \frac{3}{4})$ or $(1, \frac{1}{2}, -\frac{1}{2}, -\frac{3}{4})$ with the same probabilities then

$$\begin{aligned}
E_{\mathbf{w},\mathbf{x}} u(\mathbf{w}+\mathbf{x}) &= E_{\mathbf{w},\mathbf{x}} \log(w_1 + w_2 + x_1 + x_2) \\
&= \frac{1}{2}\log\left(\frac{1}{2} + 0 + \frac{1}{2} + \frac{3}{4}\right) + \frac{1}{2}\log\left(1 + \frac{1}{2} - \frac{1}{2} - \frac{3}{4}\right) \\
&= \frac{1}{2}\log\left(\frac{7}{16}\right) \\
E_{\mathbf{w}} u\left(\mathbf{w} + E_{\mathbf{x}}\mathbf{x}\right) &= E_{\mathbf{w}} \log(w_1 + w_2) = \frac{1}{2}\log\left(\frac{1}{2} + 0\right) + \frac{1}{2}\log\left(1 + \frac{1}{2}\right) \\
&= \frac{1}{2}\log\left(\frac{3}{4}\right) > \frac{1}{2}\log\left(\frac{7}{16}\right)
\end{aligned}$$

hence

$$E_{\mathbf{w},\mathbf{x}} u(\mathbf{w}+\mathbf{x}) < E_{\mathbf{w}} u\left(\mathbf{w} + E_{\mathbf{x}}\mathbf{x}\right).$$

One particular set of risk premiums was considered in Ambarish & Kallberg [2]: let $\widehat{\pi}_i$ be defined to be the *risk premium in the i-th direction* i.e., a solution of (3.1) with the property that components of $\boldsymbol{\pi}$ satisfy

$$\begin{aligned}
\pi_j &= 0 \qquad j \neq i \\
&= \widehat{\pi}_i \qquad j = i.
\end{aligned}$$

Let us compute directional risk premiums $\widehat{\pi}_1$, $\widehat{\pi}_2$ for the first setting in Example 1 where $E_{\mathbf{x}}\mathbf{x} = 0$:

$$\begin{aligned}
E_{\mathbf{w},\mathbf{x}} u(\mathbf{w}+\mathbf{x}) &= E_{\mathbf{w}} u(\mathbf{w} - \boldsymbol{\pi}); \qquad \boldsymbol{\pi} = (\widehat{\pi}_1, 0)' \\
0 &= \frac{1}{2}\log\left(\frac{1}{2} - \widehat{\pi}_1\right) + \frac{1}{2}\log\left(\frac{3}{2}\right) \\
\widehat{\pi}_1 &= -\frac{1}{6}
\end{aligned}$$

$$\begin{aligned}
E_{\mathbf{w},\mathbf{x}} u(\mathbf{w}+\mathbf{x}) &= E_{\mathbf{w}} u(\mathbf{w} - \boldsymbol{\pi}); \qquad \boldsymbol{\pi} = (0, \widehat{\pi}_2)' \\
0 &= \frac{1}{2}\log\left(\frac{1}{2}\right) + \frac{1}{2}\log\left(\frac{3}{2} - \widehat{\pi}_2\right) \\
\widehat{\pi}_2 &= -\frac{1}{2}
\end{aligned}$$

We can see that concavity of u does not guarantee the risk averse attitude (with respect to all gambles) and nonnegativity of directional risk premiums as it did in the univariate case. However, it is easy to see that if u is increasing in each variable then there is a relationship corresponding to the univariate case: the condition (3.2) of risk aversion at wealth level $\mathbf{w}$ with respect to gamble $\mathbf{x}$ is equivalent to nonnegativity of all directional risk premiums.

The main disadvantage of this approach is the fact that $\mathbf{w}$ is not allowed to be a function of $\boldsymbol{\pi}$. Therefore this notion is not very useful in multiperiod models and we suggest another way, how to define multiperiod risk premium.

3.4 Multiperiod risk premium

Let $u(\mathbf{w})$ be an increasing utility function. In this section, we interpret the arguments of $\mathbf{w}$ as the random amounts of cash (single commodity) measured at times $1, \ldots, n$. It is the vector of initial wealth in each period. We will denote by $\mathbf{x}$ the random vector of all changes in wealth vector $\mathbf{w}$ at times $1, \ldots, n$, i.e., x_i is a random investment possibility (gamble) at time i. We would like to define i-th element of multiperiod risk premium $\mathbf{\Pi}$ such that a decision maker is indifferent between accepting the gamble x_i and paying $\Pi_i - E_{\mathbf{x}} x_i$ in i-th time period. If the probability distribution of $\mathbf{w}$ is known and we do not want Π_i to depend on realization of $\mathbf{w}$ then we can apply the approach mentioned in Section 3.3. However, it is not very realistic assumption. In our framework, the i-th element of multiperiod risk premium depends on the initial wealth at time i and on the probabilistic distribution of $\mathbf{x}$. The initial wealth w_i depends on w_{i-1} and on the decision of investor at time $i-1$, whether he accepted gamble x_{i-1} or paid $\Pi_{i-1} - E_{\mathbf{x}} x_{i-1}$. This decision is not known usually, because the investor is indifferent between these two possibilities. Thus, we assume that $\mathbf{w}$ is a function of $\mathbf{x}$ and $\mathbf{\Pi}$.

Without loss of generality from now on, we will follow Ambarish & Kallberg [2], Duncan [9] and Pratt [45] in assuming that $E_{\mathbf{x}}\mathbf{x} = 0$.

Finally, we assume that a history of decisions does not depend on $\mathbf{x}$ and all possible histories of decisions are described by the following decision scenarios where an investor has only two possibilities in each time period: to accept the gamble or to pay risk premium.

$$
\Big\langle \begin{array}{l}
\text{accept } x_1 \ \Big\langle \begin{array}{l} \text{accept } x_2 \ \langle \ \dots \ \begin{array}{l}\text{accept } x_{n-1} \\ \text{pay } \Pi_{n-1}\end{array} \\ \text{pay } \Pi_2 \ \langle \ \dots \ \begin{array}{l}\text{accept } x_{n-1} \\ \text{pay } \Pi_{n-1}\end{array} \end{array} \\
\qquad \vdots \\
\text{pay } \Pi_1 \ \Big\langle \begin{array}{l} \text{accept } x_2 \ \langle \ \dots \\ \text{pay } \Pi_2 \ \langle \ \dots \ \begin{array}{l}\text{accept } x_{n-1} \\ \text{pay } \Pi_{n-1}\end{array} \end{array}
\end{array}
$$

Note that an information about the decision in the last period is not relevant because it can not influence the initial wealth vector $\mathbf{w}$. Thus there are $m = 2^{n-1}$ scenarios. Let S denotes the set of all scenarios. Let $s \in S$. If the decision maker accepts a gamble in i-th time period then let $k_i^s = 1$, otherwise $k_i^s = 0$. The scenario s is represented by vector $K^s = (k_1^s, k_2^s, \dots, k_{n-1}^s)$ consisting of binary elements. Each scenario uniquely describes the decisions of investor in all time periods e.g. the scenario with $k_i^s = 1$, $i = 1, 2, ..., n-1$ corresponds to the investor who accepts a fair gamble in each time period. With this notation, the initial wealth in j-th time period along scenario s can be written in the form:

$$w_j^s(\Pi_1, \dots, \Pi_{j-1}) = w_1 + \sum_{i=1}^{j-1} [k_i^s x_i - (1 - k_i^s)\Pi_i]. \tag{3.3}$$

Therefore $\mathbf{w} = \mathbf{w}^s$ with unknown probability p^s for $s \in S$ where $\mathbf{w}^s$ depends on $\mathbf{x}$, hence $\mathbf{w}^s$ is a random vector. Observe that we consider multiperiod risk premium as a price of insurance against all risks. It is not allowed to separate risks in one time period and to compute the amount of multiperiod risk premium (insurance) only for some of them. For example if we receive \$1 in the second period from external resources and we can lose \$2 with probability 0.5 in the second period gamble then the considered investment possibility is to receive \$1 or to pay \$1, i.e. $x_2 = 1$ or $x_2 = -1$ with equal probabilities.

In a formal way, we would like to define multiperiod risk premium by the system of equations:

$$E_{\mathbf{x}} u(\mathbf{w}^s + \mathbf{x}) = E_{\mathbf{x}} u\left(\mathbf{w}^s - \mathbf{\Pi}\right) \qquad \forall s \in S. \tag{3.4}$$

We assume that all expected values exist for all scenarios. Recall that $\mathbf{w}^s$ is a function of $\mathbf{\Pi}$ and $\mathbf{x}$ (see (3.3)). However, this system of 2^{n-1} equations and n variables does not usually have a solution unless $n \leq 2$. Therefore we suggest another approach. Given $\mathbf{x}$, let

$$f^s(\mathbf{\Pi}) = |E_{\mathbf{x}}u(\mathbf{w}^s + \mathbf{x}) - E_{\mathbf{x}}u\,(\mathbf{w}^s - \mathbf{\Pi})\,|$$

for non-random w_1. It is clear that $\mathbf{\Pi}$ minimizes $f^s(\mathbf{\Pi})$ if and only if $\mathbf{\Pi}$ is a solution of the corresponding equation in (3.4) for scenario s. Hence, we are interested to find $\mathbf{\Pi}$ which minimizes $f^s(\mathbf{\Pi})$ jointly for all $s \in S$ as much as possible. This is a multi-criteria programming problem and we apply the goal programming approach. We are looking for a vector $(\mathbf{\Pi})$ which minimizes the maximal value of $f^s(\mathbf{\Pi})$ over all scenarios, i.e. is a solution of the problem:

$$\min_{\mathbf{\Pi}} \max_{s \in S} f^s(\mathbf{\Pi}),$$

which can be written in the equivalent form:

$$\begin{aligned} &\min_{\mathbf{\Pi}} \quad d \\ \text{s.t.} \quad & f^s(\mathbf{\Pi}) \leq \quad d \qquad \forall s \in S. \end{aligned} \tag{3.5}$$

Summarizing, the *multiperiod risk premium* is defined as a solution of the problem:

$$\begin{aligned} &\min_{\mathbf{\Pi}} \quad d \\ \text{s.t.} \quad & -d \leq E_{\mathbf{x}}u(\mathbf{w}^s + \mathbf{x}) \;-\; E_{\mathbf{x}}u\,(\mathbf{w}^s - \mathbf{\Pi}) \leq d \quad \forall s \in S, \end{aligned} \tag{3.6}$$

where the elements of $\mathbf{w}^s$ are given by (3.3).

It is easily seen that if an optimal solution $d^* = 0$ then the multiperiod risk premium is a solution of (3.4), else this system of equations has no solution.

We define the multiperiod risk aversion in the similar way as it was in the univariate and multivariate case using the scenario approach, i.e. the decision maker is *multiperiod risk averse at wealth level* $\mathbf{w}$ *with respect to gamble* $\mathbf{x}$ if

$$E_{\mathbf{x}}u(\mathbf{w}^s + \mathbf{x}) < E_{\mathbf{x}}u\,(\mathbf{w}^s) \qquad \forall s \in S. \tag{3.7}$$

We follow Ambarish & Kallberg [2] in applying the idea of directional risk premiums. They represent an amount that an investor can pay only in one time period to insure against all risks. We define *i-th directional multiperiod risk premium* $\widehat{\Pi}_i$ as a solution of the following problem:

$$\begin{array}{rrcl} & \min_{\mathbf{\Pi}} & d & \\ \text{s.t.} & -d \leq E_{\mathbf{x}} u(\mathbf{w}^s + \mathbf{x}) & - & E_{\mathbf{x}} u\left(\mathbf{w}^s - \mathbf{\Pi}\right) \leq d \quad \forall s \in S \\ & \Pi_j & = & 0 \quad j \neq i \end{array} \tag{3.8}$$

where the elements of $\mathbf{w}^s$ are given by (3.3).

Finally, we will prove a relationship between directional multiperiod risk premiums and multiperiod risk aversion. The corresponding property holds both for the multivariate case and the univariate case.

Theorem 3.2:

If the decision maker is multiperiod risk averse at wealth level $\mathbf{w}$ *with respect to gamble* $\mathbf{x}$ *then all directional multiperiod risk premiums are positive.*

Proof:

Choose $i \in \{1, 2, \ldots, n\}$. Let w_j^s be defined by (3.3) and $\mathbf{\Pi}^{s,i}$ be a solution of equation: $E_{\mathbf{x}} u(\mathbf{w}^s + \mathbf{x}) = E_{\mathbf{x}} u\left(\mathbf{w}^s - \mathbf{\Pi}\right)$ under conditions: $\Pi_j = 0$ for all $j \neq i$. Assumption of risk aversion at wealth level $\mathbf{w}$ with respect to gamble $\mathbf{x}$ (given by (3.7)) is equivalent to positivity of $\Pi_i^{s,i}$ for all $s \in S$, because u is increasing in each variable. Let

$$\overline{\Pi}_i = \min_{s \in S} \Pi_i^{s,i}.$$

Using (3.3) and assumption that u is increasing in each variable, it is easy to show that $f^s(\mathbf{\Pi}) = |E_{\mathbf{x}} u(\mathbf{w}^s + \mathbf{x}) - E_{\mathbf{x}} u\left(\mathbf{w}^s - \mathbf{\Pi}\right)|$ is a decreasing function in variable Π_i on $(-\infty, \overline{\Pi}_i)$ for all $s \in S$ under conditions: $\Pi_j = 0$ for all $j \neq i$. Therefore $\widehat{\Pi}_i \geq \overline{\Pi}_i > 0$. □

3.5 Generalizations of multiperiod risk premium

First, we will assume that there can be some legislative restrictions (or other reasons) such that there is no insurance possibility in some time periods or an investor is not interested in the insurance possibility in these time periods. Let A be the set of considered time periods and m be the number of considered time periods in multiperiod risk premium construction. If $i \in A$ then let $y_i = -\Pi_i$ else $y_i = x_i$. We will denote by S_A the subset of S which consist of the scenarios with the property that if $i \in \{1, 2, \ldots, n\} \setminus A$ then $k_i^s = 1$. With this notation, similarly to (3.6), we define the *partial multiperiod risk premium* $\mathbf{\Pi}^A = \{\Pi_i\}_{i \in A}$ as an m-dimensional vector which solves the problem:

$$\begin{aligned}
&\min_{\mathbf{\Pi}^A} \quad d && (3.9)\\
\text{s.t.} \quad -d \le E_{\mathbf{x}} u(\mathbf{w}^s + \mathbf{x}) &- E_{\mathbf{x}} u(\mathbf{w}^s + \mathbf{y}) \le d \quad \forall s \in S_A\\
w_j^s &= w_1 + \sum_{i=1}^{j-1} [k_i^s x_i + (1 - k_i^s) y_i] \quad j = 2, 3, \ldots, n\\
y_i &= -\Pi_i \quad i \in A\\
y_i &= x_i \quad i \notin A
\end{aligned}$$

We will illustrate the computation of multiperiod risk premium, directional multiperiod risk premium and partial multiperiod risk premium in the following example.

Example 3.3:

Consider $u(w_1, w_2, w_3) = \log(w_1 + w_2 + w_3)$. Let x_1, x_2, x_3 be an independent random variables: $x_i = \pm\frac{1}{2}$ with equal probabilities, $i = 1, 2, 3$. Finally, set $w_1 = 2$.

First, we evaluate the multiperiod risk premium given by (3.6). Any scenario s is determined by vector $K = (k_1^s, k_2^s)$ e.g. if $K = (1, 0)$ then an investor will accept the first gamble and he will pay Π_2 in the second period to insure against x_2. It is clear that S consists of four scenarios: $s_1 \sim (1, 1)$, $s_2 \sim (1, 0)$, $s_3 \sim (0, 1)$ and $s_4 \sim (0, 0)$. It is easy to check that:

$$\begin{aligned}
f^1(\mathbf{\Pi}) &= |E_{\mathbf{x}} u(w_1 + x_1, w_1 + x_1 + x_2, w_1 + x_1 + x_2 + x_3)\\
&\quad - E_{\mathbf{x}} u(w_1 - \Pi_1, w_1 + x_1 - \Pi_2, w_1 + x_1 + x_2 - \Pi_3)|\\
&= \left| \frac{1}{8} \log(1088640) - \frac{1}{4} \log \left[(\frac{15}{2} - \Pi_1 - \Pi_2 - \Pi_3) \right.\right.\\
&\quad \left.\left. (\frac{13}{2} - \Pi_1 - \Pi_2 - \Pi_3)(\frac{11}{2} - \Pi_1 - \Pi_2 - \Pi_3)(\frac{9}{2} - \Pi_1 - \Pi_2 - \Pi_3) \right] \right|
\end{aligned}$$

$$
\begin{aligned}
f^2(\mathbf{\Pi}) &= |E_{\mathbf{x}}u(w_1+x_1, w_1+x_1+x_2, w_1+x_1-\Pi_2+x_3) \\
&\quad -E_{\mathbf{x}}u(w_1-\Pi_1, w_1+x_1-\Pi_2, w_1+x_1-\Pi_2-\Pi_3)| \\
&= \left| \frac{1}{8}\log\left[(\frac{17}{2}-\Pi_2)(\frac{15}{2}-\Pi_2)(\frac{15}{2}-\Pi_2)(\frac{13}{2}-\Pi_2)(\frac{11}{2}-\Pi_2)\right.\right. \\
&\quad \left.(\frac{9}{2}-\Pi_2)(\frac{9}{2}-\Pi_2)(\frac{7}{2}-\Pi_2)\right] \\
&\quad \left. -\frac{1}{2}\log\left[(7-\Pi_1-2\Pi_2-\Pi_3)(5-\Pi_1-2\Pi_2-\Pi_3)\right]\right|
\end{aligned}
$$

$$
\begin{aligned}
f^3(\mathbf{\Pi}) &= |E_{\mathbf{x}}u(w_1+x_1, w_1-\Pi_1+x_2, w_1-\Pi_1+x_2+x_3) \\
&\quad -E_{\mathbf{x}}u(w_1-\Pi_1, w_1-\Pi_1-\Pi_2, w_1-\Pi_1+x_2-\Pi_3)| \\
&= \left| \frac{1}{8}\log\left[(8-2\Pi_1)(7-2\Pi_1)(6-2\Pi_1)(5-2\Pi_1)(7-2\Pi_1)\right.\right. \\
&\quad \left.(6-2\Pi_1)(5-2\Pi_1)(4-2\Pi_1)\right] \\
&\quad \left. -\frac{1}{2}\log\left[(\frac{13}{2}-3\Pi_1-\Pi_2-\Pi_3)(\frac{11}{2}-3\Pi_1-\Pi_2-\Pi_3)\right]\right|
\end{aligned}
$$

$$
\begin{aligned}
f^4(\mathbf{\Pi}) &= |E_{\mathbf{x}}u(w_1+x_1, w_1-\Pi_1+x_2, w_1-\Pi_1-\Pi_2+x_3) \\
&\quad -E_{\mathbf{x}}u(w_1-\Pi_1, w_1-\Pi_1-\Pi_2, w_1-\Pi_1-\Pi_2-\Pi_3)| \\
&= \left| \frac{1}{8}\log\left[(\frac{15}{2}-2\Pi_1-\Pi_2)(\frac{13}{2}-2\Pi_1-\Pi_2)(\frac{13}{2}-2\Pi_1-\Pi_2)\right.\right. \\
&\quad (\frac{11}{2}-2\Pi_1-\Pi_2)(\frac{13}{2}-2\Pi_1-\Pi_2)(\frac{11}{2}-2\Pi_1-\Pi_2) \\
&\quad \left.\left.(\frac{11}{2}-2\Pi_1-\Pi_2)(\frac{9}{2}-2\Pi_1-\Pi_2)\right] - \log(6-3\Pi_1-2\Pi_2-\Pi_3)\right|
\end{aligned}
$$

and the multiperiod risk premium (optimal solution of (3.5)) is:

$$\mathbf{\Pi} = (1.252, 1.27, -2.319) \text{ and } d^* = 6.10^{-4}.$$

Let us compute the directional multiperiod risk premiums given by (3.8). The first directional multiperiod risk premium is an optimal solution of the problem:

$$\begin{aligned} & \min_{\Pi_1} \quad d \\ \text{s.t.} \quad |E_{\mathbf{x}}u(w_1+x_1, w_2^s+x_2, w_3^s+x_3) &- E_{\mathbf{x}}u\left(w_1-\Pi_1, w_2^s, w_3^s\right)| \leq d \quad \forall s \in S \\ w_2^s &= w_1 + k_1^s x_1 - (1-k_1^s\Pi_1) \\ w_3^s &= w_1 + k_1^s x_1 - (1-k_1^s\Pi_1) + k_2^s x_2 \end{aligned}$$

where the last two conditions are concluded from (3.3). Hence $\widehat{\Pi}_1 = 0.1367$ and $d^* = 0.0124$. By analogy, solving the following problem, we obtain the second directional multiperiod risk premium.

$$\begin{aligned} & \min_{\Pi_2} \quad d \\ \text{s.t.} \quad |E_{\mathbf{x}}u(w_1+x_1, w_2^s+x_2, w_3^s+x_3) &- E_{\mathbf{x}}u\left(w_1, w_2^s-\Pi_2, w_3^s\right)| \leq d \quad \forall s \in S \\ w_2^s &= w_1 + k_1^s x_1 \\ w_3^s &= w_1 + k_1^s x_1 + k_2^s x_2 - (1-k_2^s\Pi_2). \end{aligned}$$

Thus $\widehat{\Pi}_2 = 0.1368$ and $d^* = 0.0124$. In the same manner we can see that the third directional multiperiod risk premium can be evaluated from the problem:

$$\begin{aligned} & \min_{\Pi_3} \quad d \\ \text{s.t.} \quad |E_{\mathbf{x}}u(w_1+x_1, w_2^s+x_2, w_3^s+x_3) &- E_{\mathbf{x}}u\left(w_1, w_2^s, w_3^s-\Pi_3\right)| \leq d \quad \forall s \in S \\ w_2^s &= w_1 + k_1^s x_1 \\ w_3^s &= w_1 + k_1^s x_1 + k_2^s x_2. \end{aligned}$$

Therefore $\widehat{\Pi}_3 = 0.1368$ and $d^* = 0.0124$. We can see that all the directional multiperiod risk premiums are approximately equal.

Finally, let us compute the partial multiperiod risk premium. We assume that the insurance possibility does not exist in the second period, i.e. $A = \{1,3\}$. Thus $\mathbf{y} = (-\Pi_1, x_2, -\Pi_3)$. Since only two scenarios are possible in this situation ($S_A = \{K^1, K^2\}$ where $K^1 = (1,1)$ and $K^2 = (0,1)$), applying (3.9), we obtain the partial multiperiod risk premium as a solution of the following system of two equations:

$$\begin{aligned} E_{\mathbf{x}}u(w_1+x_1, w_1+x_1+x_2, w_1+x_1+x_2+x_3) - & \\ E_{\mathbf{x}}u\left(w_1-\Pi_1, w_1+x_1+x_2, w_1+x_1+x_2-\Pi_3\right) &= 0 \\ E_{\mathbf{x}}u(w_1+x_1, w_1-\Pi_1+x_2, w_1-\Pi_1+x_2+x_3) - & \\ E_{\mathbf{x}}u\left(w_1-\Pi_1, w_1-\Pi_1+x_2, w_1-\Pi_1+x_2-\Pi_3\right) &= 0 \end{aligned}$$

Thus $\Pi_1 = 1.638$ and $\Pi_3 = -1.5$.

Another generalization of the multiperiod approach is based on the assumption that w_1 is a random variable. Suppose that the scenarios K^s and their unknown probabilities p^s do not depend on w_1. We follow Ambarish & Kallberg [2] in adopting the possibility of non-zero correlation between w_1 and $\mathbf{x}$. By analogy to the univarite case with random initial wealth developed in Kihlstrom & Romer & Williams [24] and Ross [49], we define the *multiperiod risk premium for random* w_1 as a solution of the problem:

$$\begin{aligned} &\min_{\mathbf{\Pi}} \quad d \\ \text{s.t.} \quad & -d \leq E_{\mathbf{x},w_1} u(\mathbf{w}^s + \mathbf{x}) \; - \; E_{\mathbf{x},w_1} u\left(\mathbf{w}^s - \mathbf{\Pi}\right) \leq d \quad \forall s \in S \end{aligned} \tag{3.10}$$

where the elements of $\mathbf{w}^s$ are given by (3.3). We can see that the only difference between (3.10) and (3.6) is in considering expected value with respect to both random variables: $\mathbf{x}$ and w_1. If we apply expected value with respect to $\mathbf{x}$ and w_1 instead of expected value only with respect to $\mathbf{x}$ then we can also define directional and partial multiperiod risk premiums for random w_1.

Finally we will modify the assumption of independence between history of decisions and $\mathbf{x}$. There can exist investment possibilities, which can have non-random yields in some time periods and the distribution of yields can depend on the history of realization of $\mathbf{x}$. As an example of such investment possibility a bond can be considered. If the bond default comes in t-th time period then $x_i = 0$ a.s. for all $i > t$. In this case, there is no risk in i-th time period. Therefore the value of risk premium in i-th time period has to be equal to zero.

In general, we assume that if x_i is non-random then i-th element of multiperiod risk premium is equal to zero. A *conditional multiperiod risk premium* $\widetilde{\mathbf{\Pi}}$ represents the price of insurance against all risks with an additional condition: if the realization of $\mathbf{x}$ is such that the investment possibility in i-th time period is not risky, then no insurance is available in this time period (i.e. $\widetilde{\Pi}_i = 0$). If $x_i = Ex_i$ a.s. then let $y_i = 0$ else $y_i = \widetilde{\Pi}_i$. With this notation, we define the *conditional multiperiod risk premium* $\widetilde{\mathbf{\Pi}}$ as a solution of the problem:

$$\begin{aligned}
&\min_{\widetilde{\mathbf{\Pi}}} \quad d \\
-d \leq E_{\mathbf{x}} u(\mathbf{w}^s + \mathbf{x}) \quad &- \quad E_{\mathbf{x}} u\left(\mathbf{w}^s - \mathbf{y}\right) \leq d \quad \forall s \in S \\
w_j^s \quad &= \quad w_1 + \sum_{i=1}^{j-1} [k_i^s x_i - (1 - k_i^s) y_i] \qquad j = 2, 3, \ldots, n \\
y_i \quad &= \quad \widetilde{\Pi}_i \qquad \text{if} \quad P(x_i = E x_i) < 1 \\
y_i \quad &= \quad 0 \qquad \text{if} \quad P(x_i = E x_i) = 1.
\end{aligned}$$

By analogy to non-conditional approach, we can consider also directional conditional multiperiod risk premiums.

Chapter 4

Second-order stochastic dominance and efficient portfolios

4.1 Preliminaries

The portfolio selection problem may be regarded as a two-step procedure. Firstly, an efficient set among all available portfolios is chosen and then the risk preferences of decision maker to this set are applied. This chapter deals with the first step. Section 4.2 recalls the basic ideas and results of stochastic dominance approach for pairwise comparisons. A given portfolio is efficient in the considered set of assets if there exists no other convex combination of the assets which strictly dominates the portfolio.

As was demonstrated in Chapter 2 and Chapter 3, the risk preferences of decision maker can be described by a von Neumann-Morgenstern utility function or risk premium. Applying value-at-risk (VaR) or conditional value-at-risk (CVaR) is another way how to express the risk attitude of decision makers. If the yields or losses of assets in the portfolio are described by discrete probabilistic distributions then CVaR can be computed as a solution of linear programming problem. This property will be used in the sequel.

In section 4.3, following Ogryczak & Ruszczyński [40], we recall the basic properties of CVaR in context of stochastic dominance. The relationship between risk premium and CVaR is shown. Finally, CVaR for the case of discrete probability distribution is analyzed. These results are employed in section 4.4 where a necessary and sufficient condition for SSD portfolio efficiency is derived and compared with conditions in Post [43] and Kuosmanen [34]. Also a necessary condition based on CVaR is presented. This necessary condition can detect SSD portfolio inefficiency especially when assets returns are highly correlated.

Summarizing conditions from section 4.4 we formulate linear programming algorithm for testing SSD efficiency of a given portfolio in section 4.5. If a tested portfolio is SSD inefficient then this test always identifies a dominating SSD efficient portfolio.

Following the idea of Post [43], in section 4.6, we introduced a measure of portfolio inefficiency. However, this measure is based on CVaR and uses solution of linear program in necessary and sufficient condition for SSD efficiency presented in section 4.4. We prove the consistency of this measure with SSD relation and we analyze its convexity. Finally, we illustrate these results on a simple numerical example.

We apply the derived results to test SSD efficiency of mean-VaR optimal portfolios in numerical application presented in section 4.7. We compute SSD portfolio inefficiency measures of all tested portfolios.

4.2 Stochastic dominance

For two random variables X_1 and X_2 with respective cumulative probability distributions functions $F_1(x)$, $F_2(x)$ we say that X_1 *dominates* X_2 *by first degree stochastic dominance*: $X_1 \succeq_{FSD} X_2$ if

$$\mathbb{E}_{F_1} u(x) - \mathbb{E}_{F_2} u(x) \geq 0$$

for every utility function u, i.e. for every continuous nondecreasing function u, such that these expected values exist. Let us denote by U_1 the set of all such functions. We say that X_1 *dominates* X_2 *by second degree stochastic dominance*: $X_1 \succeq_{SSD} X_2$ if

$$\mathbb{E}_{F_1} u(x) - \mathbb{E}_{F_2} u(x) \geq 0$$

for every $u \in U_2$ where $U_2 \subset U_1$ denotes the set of all concave utility functions such that these expected values exist. The corresponding strict dominance relations $\succ_{FSD}$ and $\succ_{SSD}$ are defined in the usual way: $X_1 \succ_{FSD} X_2$ ($X_1 \succ_{SSD} X_2$) if and only if $X_1 \succeq_{FSD} X_2$ ($X_1 \succeq_{SSD} X_2$) and $X_2 \not\succeq_{FSD} X_1$

($X_2 \not\succeq_{SSD} X_1$). According to Russel & Seo [50], $u \in U_2$ may be represented by simple utility functions in the following sense:

$$\mathbb{E}_{F_1} u(x) - \mathbb{E}_{F_2} u(x) \geq 0 \quad \forall u \in U_2 \iff \mathbb{E}_{F_1} u(x) - \mathbb{E}_{F_2} u(x) \geq 0 \quad \forall u \in V$$

where $V = \{u_\eta(x) : \eta \in \mathbb{R}\}$ and $u_\eta(x) = \min\{x - \eta, 0\}$.

For the development of the third or higher degree of stochastic dominance see Levy [36], Whitmore [55] and Whitmore [56]. Set

$$F_i^{(2)}(t) = \int_{-\infty}^{t} F_i(x)dx \qquad i = 1, 2.$$

The following necessary and sufficient conditions for stochastic dominance were proved in Hanoch & Levy [18].

Lemma 4.1:

Let $F_1(x)$ and $F_2(x)$ be cumulative distribution functions of X_1 and X_2. Then

- $X_1 \succeq_{FSD} X_2 \quad \Leftrightarrow \quad F_1(x) \leq F_2(x) \quad \forall x \in \mathbb{R}$
- $X_1 \succeq_{SSD} X_2 \quad \Leftrightarrow \quad F_1^{(2)}(t) \leq F_2^{(2)}(t) \quad \forall t \in \mathbb{R}$
- $X_1 \succ_{FSD} X_2 \quad \Leftrightarrow \quad F_1(x) \leq F_2(x) \quad \forall x \in \mathbb{R}$ where at least one strict inequality holds
- $X_1 \succ_{SSD} X_2 \quad \Leftrightarrow \quad F_1^{(2)}(t) \leq F_2^{(2)}(t) \quad \forall t \in \mathbb{R}$ with at least one strict inequality.

Lemma 4.1 can be used as an alternative definition of stochastic dominance.

Consider now the quantile model of stochastic dominance Ogryczak & Ruszczyński [40]. The *first quantile function* $F_X^{(-1)}$ corresponding to a real random variable X is defined as the left continuous inverse of its cumulative probability distribution function F_X:

$$F_X^{(-1)}(v) = \min\{u : F_X(u) \geq v\}. \tag{4.1}$$

The following result follows directly from Lemma 4.1.

Lemma 4.2:

$$X_1 \succeq_{FSD} X_2 \quad \Leftrightarrow \quad F_1^{(-1)}(p) \le F_2^{(-1)}(p) \quad \forall p \in (0,1\rangle.$$

The *second quantile function* $F_X^{(-2)}$ is defined as

$$\begin{aligned} F_X^{(-2)}(p) &= \int_{-\infty}^{p} F_X^{(-1)}(t)dt && \text{for } 0 < p \le 1 \\ &= 0 && \text{for } p = 0 \\ &= +\infty && \text{otherwise.} \end{aligned}$$

The function $F_X^{(-2)}$ is convex and it is well defined for any random variable X satisfying the condition $\mathbb{E}|X| < \infty$. For the proof of the following basic properties of the second quantile function and more details about dual stochastic dominance see Ogryczak & Ruszczyński [40].

Theorem 4.3:

For every random variable X with $\mathbb{E}|X| < \infty$ we have:

(i)

$$F_X^{(-2)}(p) = \sup_{\nu}\{\nu p - \mathbb{E}\max(\nu - X, 0)\}$$

(ii)

$$X_1 \succeq_{SSD} X_2 \quad \Leftrightarrow \quad \frac{F_1^{(-2)}(p)}{p} \ge \frac{F_2^{(-2)}(p)}{p} \quad \forall p \in \langle 0,1\rangle.$$

4.3 VaR and CVaR

Let Y be a random loss variable corresponding to the yield described by random variable X, i.e. $Y = -X$. We assume that $\mathbb{E}|Y| < \infty$. For a fixed level α, the *value-at-risk* VaR is defined as the α-quantile of the cumulative distribution function F_Y :

$$\mathrm{VaR}_\alpha(Y) = F_Y^{(-1)}(\alpha). \tag{4.2}$$

We follow Pflug [42] in defining *conditional value-at-risk* CVaR as the solution of the optimization problem

$$\mathrm{CVaR}_\alpha(Y) = \min_{a \in \mathbb{R}}\{a + \frac{1}{1-\alpha}\mathbb{E}[Y - a]^+\} \tag{4.3}$$

where $[x]^+ = \max(x, 0)$. This problem has always a solution and one of minimizers is $\mathrm{VaR}_\alpha(Y)$. See Pflug [42] for proof and more details. It was shown in Uryasev & Rockafellar [53] that the CVaR can be also defined as the conditional expectation of Y, given that $Y > \mathrm{VaR}_\alpha(Y)$, i.e.

$$\mathrm{CVaR}_\alpha(Y) = \mathbb{E}(Y|Y > \mathrm{VaR}_\alpha(Y)). \tag{4.4}$$

If we use $-Y$ and $1-\alpha$ instead of X and p in Theorem 4.3, respectively we can directly see from the definition of CVaR that:

$$\begin{aligned}
\frac{F_X^{(-2)}(p)}{p} &= \sup_\nu \{\nu - \frac{1}{p}\mathbb{E}\max(\nu - X, 0)\} \\
&= -\inf_\nu \{-\nu + \frac{1}{p}\mathbb{E}\max(\nu - X, 0)\} \\
&= -\inf_a \{a + \frac{1}{1-\alpha}\mathbb{E}\max(Y - a, 0)\} \\
&= -\mathrm{CVaR}_\alpha(Y).
\end{aligned}$$

Therefore Theorem 4.3 leads to the following result.

Lemma 4.4:

Let $Y_i = -X_i$ and $\mathbb{E}|X_i| < \infty$ for $i = 1, 2$. Then

$$X_1 \succeq_{SSD} X_2 \quad \Leftrightarrow \quad \mathrm{CVaR}_\alpha(Y_1) \leq \mathrm{CVaR}_\alpha(Y_2) \quad \forall \alpha \in \langle 0, 1\rangle.$$

A well known property of CVaR_α is its convexity in the following sense.

Lemma 4.5:

Set $\lambda \in \langle 0, 1\rangle$. Then

$$\mathrm{CVaR}_\alpha(\lambda Y_1 + (1-\lambda)Y_2) \leq \lambda \mathrm{CVaR}_\alpha(Y_1) + (1-\lambda)\mathrm{CVaR}_\alpha(Y_2) \tag{4.5}$$

where Y_1, Y_2 are arbitrary random variables.

Proof:

The proof follows from convexity of $y \to [y-a]^+$.

4.3.1 CVaR for scenario approach

In this subsection we limit our attention to scenario approach, i.e. we will assume that Y is a discrete random variable which takes values $y^t, t = 1, ..., T$ with equal probabilities. Then (4.3) can be rewritten as a linear programming problem. Moreover $\text{CVaR}_\alpha(Y)$ can be calculated using the following formula:

$$\text{CVaR}_\alpha(Y) = \frac{1}{T} \sum_{y^t > \text{VaR}_\alpha(Y)} y^t \tag{4.6}$$

and the assumptions of Theorem 4.3 and Lemma 4.4 are fulfilled. For more details we refer to Pflug [42].

Following Rockafellar & Uryasev [53] and Pflug [42], applying scenario approach in (4.3), CVaR can be obtained by solving the following linear program:

$$\begin{aligned} \text{CVaR}_\alpha(Y) &= \min_{a,w_t} \; a + \frac{1}{(1-\alpha)T} \sum_{t=1}^{T} w_t \\ s.t. \quad w_t &\geq y_t - a \\ w_t &\geq 0. \end{aligned} \tag{4.7}$$

Let $y^{[k]}$ be the k-th smallest element among $y^1, y^2, ..., y^T$, i.e. $y^{[1]} \leq y^{[2]} \leq \ldots \leq y^{[T]}$. In context of stochastic dominance a description of $\text{CVaR}_\alpha(Y)$ as a function of α will be useful.

Lemma 4.6:

If $\alpha \in \left\langle \frac{k}{T}, \frac{k+1}{T} \right\rangle$ and $\alpha \neq 1$ then

$$\text{CVaR}_\alpha(Y) = y^{[k+1]} + \frac{1}{(1-\alpha)T} \sum_{i=k+1}^{T} (y^{[i]} - y^{[k+1]}) \tag{4.8}$$

for k = 0,1,...,T-1 and $\text{CVaR}_1(Y) = y^{[T]}$.

Proof:

Consider a random variable Y which takes values $y^t, t = 1, ..., T$ with probabilities $p_1, p_2, ..., p_T$. For a chosen α define j_α such that

$$\alpha \in \left\langle \sum_{j=1}^{j_\alpha - 1} p_j, \sum_{j=1}^{j_\alpha} p_j \right) .$$

Then the following formula was proved in Rockafellar & Uryasev [53]:

$$\mathrm{CVaR}_\alpha(Y) = \frac{1}{1-\alpha} \left[\left(\sum_{j=1}^{j_\alpha} p_j - \alpha \right) y^{[j_\alpha]} + \sum_{j=j_\alpha+1}^{T} p_j y^{[j]} \right] .$$

Since $p_t = 1/T$, $t = 1, ..., T$ we set: $j_\alpha = k + 1$ and the lemma follows.

Combining Lemma 4.4 with Lemma 4.6 we obtain the necessary and sufficient condition of the second-order stochastic dominance. This conditions can be more easily verified than conditions in Lemma 4.1, Theorem 4.3 or Lemma 4.4.

Theorem 4.7:

Let $Y_1 = -X_1$ and $Y_2 = -X_2$ be discrete random variables which take values y_1^t and y_2^t, $t = 1, ..., T$, respectively, with equal probabilities. Then

$$X_1 \succeq_{SSD} X_2 \quad \Leftrightarrow \quad \mathrm{CVaR}_\alpha(Y_1) \le \mathrm{CVaR}_\alpha(Y_2) \quad \forall \alpha \in \{0, \frac{1}{T}, \frac{2}{T}, \dots, \frac{T-1}{T}\}.$$

Proof:

Let $\alpha_k = k/T$, $k = 0, 1, ..., T-2$. Lemma 4.1 implies:

$$\mathrm{CVaR}_{\beta_1}(Y_i) = \mathrm{CVaR}_{\beta_2}(Y_i), \quad i = 1, 2 \quad \text{for all} \quad \beta_1, \beta_2 \in \left\langle \frac{T-1}{T}, 1 \right\rangle .$$

Thus it suffices to show that if

$$\mathrm{CVaR}_{\alpha_k}(Y_1) \le \mathrm{CVaR}_{\alpha_k}(Y_2) \tag{4.9}$$

and

$$\mathrm{CVaR}_{\alpha_{k+1}}(Y_1) \le \mathrm{CVaR}_{\alpha_{k+1}}(Y_2) \tag{4.10}$$

then $\mathrm{CVaR}_\alpha(Y_1) \leq \mathrm{CVaR}_\alpha(Y_2)$ for all $\alpha \in \langle \alpha_k, \alpha_{k+1} \rangle$. To obtain a contradiction, suppose that (4.9) and (4.10) holds and there exists $\widetilde{\alpha} \in \langle \alpha_k, \alpha_{k+1} \rangle$ such that $\mathrm{CVaR}_{\widetilde{\alpha}}(Y_1) > \mathrm{CVaR}_{\widetilde{\alpha}}(Y_2)$. From continuity of CVaR in α there exists $\alpha^1 \in \langle \alpha_k, \alpha_{k+1} \rangle$ and $\alpha^2 \in \langle \alpha_k, \alpha_{k+1} \rangle$, $\alpha^1 \neq \alpha^2$ such that

$$\mathrm{CVaR}_{\alpha^1}(Y_1) = \mathrm{CVaR}_{\alpha^1}(Y_2) \tag{4.11}$$

$$\mathrm{CVaR}_{\alpha^2}(Y_1) = \mathrm{CVaR}_{\alpha^2}(Y_2). \tag{4.12}$$

Substituting (4.8) into (4.11) and (4.12) we conclude that $\alpha^1 = \alpha^2$, contrary to $\alpha^1 \neq \alpha^2$ and the proof is complete.

4.3.2 Relationship between risk premium, VaR and CVaR

In Chapter 2, absolute (relative) risk aversion measure and univariate risk premium as the examples of measures of risk were considered. In these measures, the risk attitude of decision maker is expressed using utility functions. The value-at-risk and the conditional value-at-risk are risk measures of another type, where the decision maker's risk attitude is expressed by level α. We will show that for a suitable choice of utility function and for any absolutely continuous random variable X, risk premium is equal to convex combination of CVaR and VaR.

Theorem 4.8:

Let X be an absolutely continuous random variable and $\mathbb{E}(X) = 0$. Let $Y = -X$. If

$$u(z) = \min(z + F_Y^{(-1)}(\alpha), w), \qquad \alpha \in \langle 0, 1) \tag{4.13}$$

then

$$\pi(w, P_X) = (1-\alpha)\mathrm{CVaR}_\alpha(Y) + \alpha \mathrm{VaR}_\alpha(Y).$$

Proof:

From (2.2) and (4.13) we have:

$$\begin{aligned} \mathbb{E}\min(w - Y + F_Y^{(-1)}(\alpha), w) &= \min(w + F_Y^{(-1)}(\alpha) - \pi(w, P_X), w) \\ \mathbb{E}\min(-Y + F_Y^{(-1)}(\alpha), 0) &= \min(F_Y^{(-1)}(\alpha) - \pi(w, P_X), 0). \end{aligned}$$

Since Y has a smooth distribution function $\mathbb{E}\min(-Y + F_Y^{(-1)}(\alpha), 0)$ is negative. Hence

$$\begin{aligned}\pi(w, P_X) &= F_Y^{(-1)}(\alpha) - \mathbb{E}\min(-Y + F_Y^{(-1)}(\alpha), 0) \\ &= F_Y^{(-1)}(\alpha) + \mathbb{E}\max(Y - F_Y^{(-1)}(\alpha), 0) \\ &= \mathbb{E}\max(Y, F_Y^{(-1)}(\alpha))\end{aligned}$$

and it is easy to see that

$$\begin{aligned}\mathbb{E}\max(Y, F_Y^{(-1)}(\alpha)) &= \mathbb{P}(Y > F_Y^{(-1)}(\alpha))\mathbb{E}(Y|Y > F_Y^{(-1)}(\alpha)) \\ &\quad + \mathbb{P}(Y \leq F_Y^{(-1)}(\alpha))F_Y^{(-1)}(\alpha).\end{aligned}$$

Combining it with (4.1),(4.2) and (4.4) the proof is complete.

In the case that $\mathbb{E}(X) \neq 0$ we can consider initial wealth $w' = w + \mathbb{E}(X)$, the gamble $X' = X - \mathbb{E}(X)$ and Theorem 4.8 can be formulated for w' and X' instead of w and X, respectively.

The utility function $u(z)$ given by (4.13) is a linear transformation of a representative utility function in the sense of Russel & Seo [50]. All wealth levels higher than w give the same utility and utility of losses is modified by the risk term represented by $\text{VaR}_\alpha(Y) = F_Y^{(-1)}(\alpha)$.

4.4 SSD portfolio efficiency criteria

Consider a random vector $\mathbf{r} = (r_1, r_2, ..., r_N)'$ of yields of N assets and T equiprobable scenarios. The yields of the assets for the various scenarios are given by

$$X = \begin{pmatrix} \mathbf{x}^1 \\ \mathbf{x}^2 \\ \vdots \\ \mathbf{x}^T \end{pmatrix}$$

where $\mathbf{x}^t = (x_1^t, x_2^t, \dots, x_N^t)$ is the t-th row of matrix X. Without loss of generality we can assume that the columns of X are linearly independent. In addition to the individual choice alternatives, the decision maker may also combine the alternatives into a portfolio. We will use $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, ..., \lambda_N)'$ for a vector of portfolio weights and the portfolio possibilities are given by

$$\Lambda = \{\boldsymbol{\lambda} \in R^N | \mathbf{1}'\boldsymbol{\lambda} = 1, \quad \lambda_n \geq 0, \quad n = 1, 2, \dots, N\}.$$

The tested portfolio is denoted by $\boldsymbol{\tau} = (\tau_1, \tau_2, ..., \tau_N)'$. In finance data, the yields of assets are usually significantly correlated. A special interesting case of X which may occur for strongly correlated yields of assets is defined as follows.

Definition 4.9:

Matrix X is called portfolio-monotone if there exists permutation $\Pi : \{1, 2, ..., T\} \rightarrow \{1, 2, ..., T\}$ such that $\mathbf{x}^t\boldsymbol{\tau} = (X\boldsymbol{\tau})^{[\Pi(t)]}$ for all $\boldsymbol{\tau} \in \Lambda$, $t = 1, 2, ..., T$.

Lemma 4.10:

If X is portfolio-monotone matrix of scenarios then

$\mathrm{CVaR}_\alpha\left(-\mathbf{r}'[\eta\boldsymbol{\tau}_1 + (1-\eta)\boldsymbol{\tau}_2]\right) = \eta\mathrm{CVaR}_\alpha(-\mathbf{r}'\boldsymbol{\tau}_1) + (1-\eta)\mathrm{CVaR}_\alpha(-\mathbf{r}'\boldsymbol{\tau}_2)$

for any $\boldsymbol{\tau}_1, \boldsymbol{\tau}_2 \in \Lambda$ and for any $\eta, \alpha \in \langle 0, 1\rangle$.

Proof:

If X is portfolio-monotone then $-X$ is portfolio-monotone and the proof follows directly from Lemma 4.1. $\square$

Following Ruszczyński & Vanderbei [51] and Kuosmanen [34] we will define SSD efficiency of a given portfolio $\boldsymbol{\tau}$.

Definition 4.11:

A given portfolio $\boldsymbol{\tau} \in \Lambda$ is SSD inefficient if and only if there exists portfolio $\boldsymbol{\lambda} \in \Lambda$ such that $\mathbf{r}'\boldsymbol{\lambda} \succ_{SSD} \mathbf{r}'\boldsymbol{\tau}$. Otherwise, portfolio $\boldsymbol{\tau}$ is SSD efficient.

This definition classifies portfolio as SSD efficient if and only if no other portfolio is better for all risk averse and risk neutral decision makers. In Post [43], more stringent definition of SSD efficiency was introduced.

Definition 4.12:

A given portfolio $\boldsymbol{\tau} \in \Lambda$ is SSD strict inefficient if and only if there exists portfolio $\boldsymbol{\lambda} \in \Lambda$ satisfying the following inequality

$$Eu(\mathbf{r}'\boldsymbol{\lambda}) > Eu(\mathbf{r}'\boldsymbol{\tau})$$

for all $u \in U_2^s$ where $U_2^s \in U_2$ is the set of all strictly concave utility functions. Otherwise, portfolio $\boldsymbol{\tau}$ is SSD strict efficient.

Comparing the Post definition (Definition 4.12) with our definition (Definition 4.11), these definitions coincide from empirical point of view as was argued in Post [43]. However, one can construct an example where a portfolio is classified as SSD efficient only for the Post definition, i.e. it is SSD strict efficient but SSD inefficient. Hence the Post linear programming test in the following proposition gives a necessary condition for SSD efficiency.

Proposition 4.13:

Let

$$\theta^* = \min_{\theta,\beta_t} \theta \tag{4.14}$$

$$\begin{aligned} s.t. \qquad \sum_{t=1}^{T} \beta_t(\mathbf{x}^t\boldsymbol{\tau} - x_n^t) + T\theta &\geq 0 \qquad n = 1, 2, ..., N \\ \beta_t - \beta_{t+1} &\geq 0 \qquad t = 1, 2, ..., T-1 \\ \beta_t &\geq 0 \qquad t = 1, 2, ..., T-1 \\ \beta_T &= 1. \end{aligned}$$

If portfolio $\boldsymbol{\tau}$ is SSD efficient then $\theta^* = 0$.

If some ties in elements of $X\boldsymbol{\tau}$ occur, then the constraints can be modified. See Post [43] for more details. Anyway, this criterion failed in comparing portfolios with identical means. It does not detect the presence of SSD dominating portfolio if mean of its yields equals to mean of $X\boldsymbol{\tau}$. It is caused by differences in definitions. From now on, we will deal with SSD efficiency in the sense of Definition 4.11. Following Kuosmanen [34] we can improve the Post criterion in order to obtain a necessary and sufficient condition for SSD efficiency. It depends on "ties" in $X\boldsymbol{\tau}$. We say that k-way tie occurs if k elements of $X\boldsymbol{\tau}$ are equal.

Proposition 4.14:

Let

$$
\begin{aligned}
\theta^{**} = \min_{W,\boldsymbol{\lambda},S^+,S^-} & \sum_{j=1}^{T}\sum_{i=1}^{T}(s_{ij}^+ + s_{ij}^-) \qquad (4.15) \\
s.t. \qquad X\boldsymbol{\lambda} &= WX\boldsymbol{\tau} \\
s_{ij}^+ - s_{ij}^- &= w_{ij} - \frac{1}{2} \quad i,j = 1,2,...,T \\
s_{ij}^+, s_{ij}^-, w_{ij} &\geq 0 \quad i,j = 1,2,...,T \\
\sum_{j=1}^{T} w_{ij} &= 1 \quad i = 1,2,...,T \\
\sum_{i=1}^{T} w_{ij} &= 1 \quad j = 1,2,...,T \\
\boldsymbol{\lambda} &\in \Lambda
\end{aligned}
$$

where $S^+ = \{s_{ij}^+\}_{i,j=1}^T$, $S^- = \{s_{ij}^-\}_{i,j=1}^T$ and $W = \{w_{ij}\}_{i,j=1}^T$. Let ϵ_k denote the number of k-way ties in $X\boldsymbol{\tau}$. Then portfolio $\boldsymbol{\tau}$ is SSD efficient if and only if

$$
\theta^{**} = \frac{T^2}{2} - \sum_{k=1}^{T} k\epsilon_k \quad \wedge \quad \theta^* = 0
$$

where θ^* is given by (4.14).

These criteria are based on applications of Lemma 4.1. We will derive sufficient and necessary conditions for SSD efficiency of $\boldsymbol{\tau}$ based on quantile model of second order stochastic dominance, in particular the relationship between CVaR and SSD will be employed. This new test will use smaller linear program than problem (4.15). We start with necessary condition using the following theorem. To simplify the notation, set $\Gamma = \left\{0, \frac{1}{T}, \frac{2}{T}, \ldots, \frac{T-1}{T}\right\}$.

Theorem 4.15:

Let $\alpha_k = k/T, \quad k = 0, 1, \ldots, T-1$. Let

$$d^* = \max_{\lambda_n} \sum_{k=0}^{T-1} \sum_{n=1}^{N} \lambda_n \left[\mathrm{CVaR}_{\alpha_k}(-\mathbf{r}'\boldsymbol{\tau}) - \mathrm{CVaR}_{\alpha_k}(-r_n)\right] \tag{4.16}$$

$$\text{s.t.} \quad \sum_{n=1}^{N} \lambda_n \left[\mathrm{CVaR}_{\alpha_k}(-\mathbf{r}'\boldsymbol{\tau}) - \mathrm{CVaR}_{\alpha_k}(-r_n)\right] \geq 0, \quad k = 0, 1, \ldots, T-1$$

$$\lambda \in \Lambda$$

If $d^* > 0$ then $\boldsymbol{\tau}$ is SSD inefficient. Optimal solution $\boldsymbol{\lambda}^*$ of (4.16) is an SSD efficient portfolio such that $\mathbf{r}'\boldsymbol{\lambda}^* \succ_{SSD} \mathbf{r}'\boldsymbol{\tau}$.

Proof:

If $d^* > 0$ then there is feasible solution $\boldsymbol{\lambda}$ of problem (4.16) satisfying

$$\sum_{n=1}^{N} \lambda_n \left[\mathrm{CVaR}_{\alpha_k}(-\mathbf{r}'\boldsymbol{\tau}) - \mathrm{CVaR}_{\alpha_k}(-r_n)\right] \geq 0, \quad \forall \alpha_k \in \Gamma$$

where at least one strict inequality holds. For this $\boldsymbol{\lambda}$ we have

$$\sum_{n=1}^{N} \lambda_n \mathrm{CVaR}_{\alpha_k}(-r_n) \leq \mathrm{CVaR}_{\alpha_k}(-\mathbf{r}'\boldsymbol{\tau}), \quad \forall \alpha_k \in \Gamma$$

with at least one strict inequality. From Lemma 4.5 we obtain

$$\mathrm{CVaR}_{\alpha_k}(-\mathbf{r}'\boldsymbol{\lambda}) \leq \sum_{n=1}^{N} \lambda_n \mathrm{CVaR}_{\alpha_k}(-r_n) \quad \forall \alpha_k \in \Gamma.$$

Hence

$$\mathrm{CVaR}_{\alpha_k}(-\mathbf{r}'\boldsymbol{\lambda}) \leq \mathrm{CVaR}_{\alpha_k}(-\mathbf{r}'\boldsymbol{\tau}) \quad \forall \alpha_k \in \Gamma$$

with at least one strict inequality. Applying Theorem 4.7 we can conclude that $\mathbf{r}'\boldsymbol{\lambda} \succeq_{SSD} \mathbf{r}'\boldsymbol{\tau}$. Since the last inequality is strict for at least one $\alpha_k \in \Gamma$, $\mathbf{r}'\boldsymbol{\lambda} \not\preceq_{SSD} \mathbf{r}'\boldsymbol{\tau}$ and according to Definition 4.11, $\boldsymbol{\tau}$ is SSD inefficient. The SSD efficiency of optimal solution $\boldsymbol{\lambda}^*$ follows directly from the formulation of objective function in (4.16), which completes the proof. $\square$

Problem (4.16) is a linear program with N variables and $N + T + 1$ constraints. Since, in SSD portfolio efficiency testing, N is usually much more smaller than T , in comparison with test suggested in Post [43] (Proposition 4.13), problem (4.16) is smaller. Moreover, contrary to (4.14), if (4.16) shows SSD inefficiency it also identifies the dominating SSD efficient portfolio. The power of necessary condition in Theorem 4.15 depends on correlation between random variables r_n, $n = 1, 2, ..., N$. In finance data, the yields of assets are often strongly correlated. In this case, according to Lemma 4.1. the convexity gap of CVaR, i.e. the difference between RHS and LHS in (4.5) is not very large. Thus the condition in Theorem 4.15 can identify the corresponding SSD efficient dominating portfolio very fast. Moreover, according to Lemma 4.1, if X is portfolio-monotone then Theorem 4.15 presents necessary and sufficient condition for SSD efficiency.

In general, Theorem 4.15 presents only necessary condition for SSD efficiency of $\boldsymbol{\tau}$ and portfolio $\boldsymbol{\tau}$ can be SSD inefficient even if (4.16) has no feasible solution or $d^* = 0$. If $d^* = 0$ then two possibilities may occur:

(1) Problem (4.16) has a unique solution $\boldsymbol{\lambda}^* = \boldsymbol{\tau}$. If this is the case then $\boldsymbol{\tau}$ is SSD efficient.

(2) Problem (4.16) has an optimal solution $\boldsymbol{\lambda}^* \neq \boldsymbol{\tau}$. In this case, $\boldsymbol{\tau}$ is SSD inefficient and $\mathbf{r}'\boldsymbol{\lambda}^* \succ_{SSD} \mathbf{r}'\boldsymbol{\tau}$. Moreover, $\boldsymbol{\lambda}^*$ is an SSD efficient portfolio.

The situation when $d^* = 0$, $\boldsymbol{\lambda}^* \neq \boldsymbol{\tau}$ and $\boldsymbol{\tau}$ is SSD efficient would imply

$$X\boldsymbol{\lambda}^* = X\boldsymbol{\tau}$$

which contradicts the assumption of linearly independent columns of X.

If problem (4.16) has no feasible solution then we can employ the following necessary and sufficient condition for SSD efficiency. This result was obtained thanks to a personal consultation with Petr Chovanec.

Theorem 4.16:

Let $\alpha_k = k/T, \quad k = 0, 1, \dots, T-1$. Let

$$D^*(\boldsymbol{\tau}) = \max_{D_k, \lambda_n, b_k} \sum_{k=0}^{T-1} D_k \tag{4.17}$$

$$\begin{aligned} \text{s.t. } \mathrm{CVaR}_{\alpha_k}(-\mathbf{r}'\boldsymbol{\tau}) - b_k - \frac{1}{1-\alpha_k}\mathbb{E}\max(-\mathbf{r}'\boldsymbol{\lambda} - b_k, 0) &\geq D_k, & k = 0, 1, \dots, T-1 \\ D_k &\geq 0, & k = 0, 1, \dots, T-1 \\ \boldsymbol{\lambda} &\in \Lambda \end{aligned}$$

If $D^*(\boldsymbol{\tau}) > 0$ then $\boldsymbol{\tau}$ is SSD inefficient and $\mathbf{r}'\boldsymbol{\lambda}^* \succ_{SSD} \mathbf{r}'\boldsymbol{\tau}$. Otherwise, $D^*(\boldsymbol{\tau}) = 0$ and $\boldsymbol{\tau}$ is SSD efficient.

Proof:

Let $\boldsymbol{\lambda}^*, b_k^*$, $k = 0, 1, \dots, T-1$ be an optimal solution of (4.17). If $D^*(\boldsymbol{\tau}) > 0$ then

$$b_k^* + \frac{1}{1-\alpha_k}\mathbb{E}\max(-\mathbf{r}'\boldsymbol{\lambda}^* - b_k^*, 0) \leq \mathrm{CVaR}_{\alpha_k}(-\mathbf{r}'\boldsymbol{\tau}) \quad \forall \alpha_k \in \Gamma \tag{4.18}$$

where at least one inequality holds strict. Since from the definition of CVaR we have

$$\mathrm{CVaR}_{\alpha_k}(-\mathbf{r}'\boldsymbol{\lambda}^*) = \min_{b_k} \left\{ b_k + \frac{1}{1-\alpha_k}\mathbb{E}\max(-\mathbf{r}'\boldsymbol{\lambda}^* - b_k, 0) \right\}$$

we conclude from (4.18) that

$$\mathrm{CVaR}_{\alpha_k}(-\mathbf{r}'\boldsymbol{\lambda}^*) \leq \mathrm{CVaR}_{\alpha_k}(-\mathbf{r}'\boldsymbol{\tau})$$

with at least one strict inequality. By analogy to the proof of Theorem 4.15, it is easily seen that $\boldsymbol{\tau}$ is SSD inefficient and $\mathbf{r}'\boldsymbol{\lambda}^* \succ_{SSD} \mathbf{r}'\boldsymbol{\tau}$.

If $D^*(\boldsymbol{\tau}) = 0$ then problem (4.17) has unique optimal solution: $\boldsymbol{\lambda}^* = \boldsymbol{\tau}$, because the presence of another optimal solution contradicts the assumption of linearly independent columns of X. Thus there is no strictly dominating portfolio and hence $\boldsymbol{\tau}$ is SSD efficient, similarly as for (4.16). Since $\boldsymbol{\tau}$ is always a feasible solution of (4.17), D^* can not be negative and the proof is complete. $\square$

Problem (4.17) has $N + 2T + 1$ constraints and $N + 2T$ variables. Inspired by (4.7) and following Pflug [42], Rockafellar & Uryasev [53], it can be rewritten as a linear programming problem with $2T(T + 1) + N + 1$ constraints and $T(T + 2) + N$ variables:

$$
\begin{aligned}
D^*(\boldsymbol{\tau}) = \max_{D_k, \lambda_n, b_k, w_k^t} & \sum_{k=1}^{T} D_k && && (4.19)\\
\text{s.t.}\ \ \mathrm{CVaR}_{\frac{k-1}{T}}(-\mathbf{r}'\boldsymbol{\tau}) - b_k - \frac{1}{(1-\frac{k-1}{T})T} & \sum_{t=1}^{T} w_k^t \ \geq \ D_k, && k = 1, \ldots, T \\
w_k^t & \geq -\mathbf{x}^t\boldsymbol{\lambda} - b_k, && t, k = 1, \ldots, T \\
w_k^t & \geq 0, && t, k = 1, \ldots, T \\
D_k & \geq 0, && k = 1, \ldots, T \\
\boldsymbol{\lambda} & \in \Lambda
\end{aligned}
$$

Using (4.19) instead of (4.17) in Theorem 4.16 we obtain a linear programming criterion for SSD efficiency.

This sufficient and necessary condition requires solution of a smaller linear program than it is in the Kuosmanen test (see Theorem 4.13). Moreover, it identifies SSD efficient dominating portfolio. In comparison with necessary conditions in Proposition 4.13 and Theorem 4.15, the number of variables is approximately equal to square of T.

4.5 Algorithm for testing SSD portfolio efficiency

Employing results derived in Section 4.4 we have an algorithm for testing SSD portfolio efficiency of portfolio $\boldsymbol{\tau}$ in the set of assets. In the first step, we check some special convex combinations. In the next steps, we use necessary conditions derived in Theorem 4.15 and Proposition 4.13. Finally, we use test in Theorem 4.16. The steps are sorted from the easiest to the most demanding in computational perspective. If the SSD efficiency or SSD inefficiency is detected in Step 1, Step 2 or Step 4 then we obtain a dominating SSD efficient portfolio as a by-product.

Step 1: If $r_n \succ_{SSD} \mathbf{r}'\boldsymbol{\tau}$ for some $n \in \{1, 2, \ldots, N\}$ or $\frac{1}{N}\sum_{n=1}^{N} r_n \succ_{SSD} \mathbf{r}'\boldsymbol{\tau}$ then go to Step 5.

Step 2: Solve (4.16). If $d^* > 0$ then go to Step 5. If $d^* = 0$ and (4.16) has an unique optimal solution then go to Step 6. If $d^* = 0$ and (4.16) has multiple optimal solution then go to Step 5.

Step 3: Solve (4.14). If $\theta^* > 0$ then go to Step 5.

Step 4: Solve (4.17) or (4.19). If $D^* > 0$ then go to Step 5 else go to Step 6.

Step 5: Stop the algorithm, portfolio $\boldsymbol{\tau}$ is SSD inefficient.

Step 6: Stop the algorithm, $\boldsymbol{\tau}$ is SSD efficient.

4.6 SSD portfolio inefficiency measure

Inspired by Post [43] and Kopa & Post [32], $D^*(\boldsymbol{\tau})$ from (4.17) or (4.19) can be considered as a measure of inefficiency of portfolio $\boldsymbol{\tau}$, because it expresses the distance between a given tested portfolio and its dominating SSD efficient portfolio. To be able to compare SSD inefficiency of two portfolios we need to consider such a measure, which is "consistent" with SSD relation. In Ogryczak & Ruszczyński [40], a consistency of risk measure with SSD relation in mean-risk models was analyzed. By analogy, we define the consistency of a measure of SSD portfolio inefficiency with SSD relation.

Definition 4.17:

Let ξ be a measure of SSD portfolio inefficiency. We say that ξ is consistent with SSD if and only if

$$\mathbf{r}'\boldsymbol{\tau}_1 \succeq_{SSD} \mathbf{r}'\boldsymbol{\tau}_2 \Rightarrow \xi(\boldsymbol{\tau}_2) \geq \xi(\boldsymbol{\tau}_1)$$

for any $\boldsymbol{\tau}_1, \boldsymbol{\tau}_2 \in \Lambda$.

The property of consistency guarantees that if a given portfolio is worse than the other one for every risk averse or risk neutral investor then it has larger measure of inefficiency. Let $\Lambda^*(\boldsymbol{\tau}) \in \Lambda$ be a set of optimal solutions $\boldsymbol{\lambda}^*$ of (4.17) or (4.19).

Theorem 4.18:

(i) The measure of SSD portfolio inefficiency D^* given by (4.17) or (4.19) is consistent with SSD.

(ii) If $\mathbf{r}'\boldsymbol{\tau}_1 \succeq_{SSD} \mathbf{r}'\boldsymbol{\tau}_2$ and both $\boldsymbol{\tau}_1$, $\boldsymbol{\tau}_2$ are SSD inefficient then

$$D^*(\boldsymbol{\tau}_2) = D^*(\boldsymbol{\tau}_1) + \sum_{k=1}^{T} \left[\mathrm{CVaR}_{\frac{k-1}{T}}(-\mathbf{r}'\boldsymbol{\tau}_2) - \mathrm{CVaR}_{\frac{k-1}{T}}(-\mathbf{r}'\boldsymbol{\tau}_1) \right].$$

(iii) If $\mathbf{r}'\boldsymbol{\tau}_1 \succeq_{SSD} \mathbf{r}'\boldsymbol{\tau}_2$ then

$$D^*(\boldsymbol{\tau}_2) \geq D^*(\boldsymbol{\tau}_1) + \sum_{k=1}^{T} \left[\mathrm{CVaR}_{\frac{k-1}{T}}(-\mathbf{r}'\boldsymbol{\tau}_2) - \mathrm{CVaR}_{\frac{k-1}{T}}(-\mathbf{r}'\boldsymbol{\tau}_1) \right].$$

Proof:

Applying Theorem 4.7, if $\mathbf{r}'\boldsymbol{\tau}_1 \succeq_{SSD} \mathbf{r}'\boldsymbol{\tau}_2$ then

$$\sum_{k=1}^{T} \left[\mathrm{CVaR}_{\frac{k-1}{T}}(-\mathbf{r}'\boldsymbol{\tau}_2) - \mathrm{CVaR}_{\frac{k-1}{T}}(-\mathbf{r}'\boldsymbol{\tau}_1) \right] \geq 0.$$

Hence it suffices to prove (ii) and (iii).

Let $\mathbf{r}'\boldsymbol{\tau}_1$ be SSD inefficient. It is easily seen that (4.17) can be rewritten in the following way:

$$\begin{aligned} D^*(\boldsymbol{\tau}) = \max_{\lambda_n} \sum_{k=0}^{T-1} & \mathrm{CVaR}_{\frac{k}{T}}(-\mathbf{r}'\boldsymbol{\tau}) - \mathrm{CVaR}_{\frac{k}{T}}(-\mathbf{r}'\boldsymbol{\lambda}) \\ \text{s.t.} \quad & \mathrm{CVaR}_{\frac{k}{T}}(-\mathbf{r}'\boldsymbol{\tau}) - \mathrm{CVaR}_{\frac{k}{T}}(-\mathbf{r}'\boldsymbol{\lambda}) \geq 0, \quad k = 0, 1, \ldots, T-1 \\ & \boldsymbol{\lambda} \in \Lambda. \end{aligned} \tag{4.20}$$

Let $\boldsymbol{\lambda}^*(\boldsymbol{\tau}_1) \in \Lambda^*(\boldsymbol{\tau}_1)$, $\boldsymbol{\lambda}^*(\boldsymbol{\tau}_2) \in \Lambda^*(\boldsymbol{\tau}_2)$. Using Theorem 4.7 and $\mathbf{r}'\boldsymbol{\tau}_1 \succeq_{SSD} \mathbf{r}'\boldsymbol{\tau}_2$,

$$\mathrm{CVaR}_{\frac{k}{T}}(-\mathbf{r}'\boldsymbol{\tau}_2) - \mathrm{CVaR}_{\frac{k}{T}}(-\mathbf{r}'\boldsymbol{\tau}_1) \geq 0 \quad k = 0, 1, \ldots, T-1.$$

Since the sum of these differences does not depend on the choice of $\boldsymbol{\lambda}^*(\boldsymbol{\tau}_1)$, the dominating portfolio $\boldsymbol{\lambda}^*(\boldsymbol{\tau}_1)$ is also an optimal solution of (4.17) when deriving $D^*(\boldsymbol{\tau}_2)$, i.e. $\boldsymbol{\lambda}^*(\boldsymbol{\tau}_1) \in \Lambda^*(\boldsymbol{\tau}_2)$. Hence

$$\begin{aligned} D^*(\boldsymbol{\tau}_2) &= \sum_{k=0}^{T-1} \mathrm{CVaR}_{\frac{k}{T}}(-\mathbf{r}'\boldsymbol{\tau}_2) - \mathrm{CVaR}_{\frac{k}{T}}(-\mathbf{r}'\boldsymbol{\lambda}^*(\boldsymbol{\tau}_2)) \\ &= \sum_{k=0}^{T-1} \left[\mathrm{CVaR}_{\frac{k}{T}}(-\mathbf{r}'\boldsymbol{\tau}_2) - \mathrm{CVaR}_{\frac{k}{T}}(-\mathbf{r}'\boldsymbol{\tau}_1)\right] \\ &\quad + \sum_{k=0}^{T-1} \left[\mathrm{CVaR}_{\frac{k}{T}}(-\mathbf{r}'\boldsymbol{\tau}_1) - \mathrm{CVaR}_{\frac{k}{T}}(-\mathbf{r}'\boldsymbol{\lambda}^*(\boldsymbol{\tau}_1))\right] \\ &= D^*(\boldsymbol{\tau}_1) + \sum_{k=0}^{T-1} \left[\mathrm{CVaR}_{\frac{k}{T}}(-\mathbf{r}'\boldsymbol{\tau}_2) - \mathrm{CVaR}_{\frac{k}{T}}(-\mathbf{r}'\boldsymbol{\tau}_1)\right] \end{aligned}$$

which completes the proof of (ii).

Let $\mathbf{r}'\boldsymbol{\tau}_1$ be SSD efficient. From Theorem 4.16, we have $D^*(\boldsymbol{\tau}_1) = 0$. According to (4.20),

$$\begin{aligned} D^*(\boldsymbol{\tau}_2) = \max_{\lambda_n} \sum_{k=0}^{T-1} & \mathrm{CVaR}_{\frac{k}{T}}(-\mathbf{r}'\boldsymbol{\tau}_2) - \mathrm{CVaR}_{\frac{k}{T}}(-\mathbf{r}'\boldsymbol{\lambda}) \\ \text{s.t. } & \mathrm{CVaR}_{\frac{k}{T}}(-\mathbf{r}'\boldsymbol{\tau}_2) - \mathrm{CVaR}_{\frac{k}{T}}(-\mathbf{r}'\boldsymbol{\lambda}) \geq 0, \quad k = 0, 1, \ldots, T-1 \\ & \boldsymbol{\lambda} \in \Lambda. \end{aligned}$$

Since $\mathbf{r}'\boldsymbol{\tau}_1 \succeq_{SSD} \mathbf{r}'\boldsymbol{\tau}_2$, portfolio $\boldsymbol{\tau}_1$ is a feasible solution of (4.20). Hence

$$D^*(\boldsymbol{\tau}_2) \geq \sum_{k=0}^{T-1} \mathrm{CVaR}_{\frac{k}{T}}(-\mathbf{r}'\boldsymbol{\tau}_2) - \mathrm{CVaR}_{\frac{k}{T}}(-\mathbf{r}'\boldsymbol{\tau}_1)$$

and combining it with (ii), the proof is complete. □

Since SSD relation is not complete, i.e. there exist incomparable pairs of portfolios, the strict inequality of values of any portfolio inefficiency measure can not imply SSD relation. Also for the measure D^* some pair of portfolios $\boldsymbol{\tau}_1, \boldsymbol{\tau}_2$ can be found such that $D^*(\boldsymbol{\tau}_2) \geq D^*(\boldsymbol{\tau}_1)$ and $\mathbf{r}'\boldsymbol{\tau}_1 \not\succeq_{SSD} \mathbf{r}'\boldsymbol{\tau}_2$. In the following theorem, a convexity property of portfolio inefficiency measure D^* is analyzed.

Theorem 4.19:

Let $\boldsymbol{\tau}_1$, $\boldsymbol{\tau}_2$, $\boldsymbol{\tau}_3 \in \Lambda$.

(i) If $\mathbf{r}'\boldsymbol{\tau}_1 \succeq_{SSD} \mathbf{r}'\boldsymbol{\tau}_2$ then

$$D^*(\eta\boldsymbol{\tau}_1 + (1-\eta)\boldsymbol{\tau}_2) \leq \eta D^*(\boldsymbol{\tau}_1) + (1-\eta)D^*(\boldsymbol{\tau}_2)$$

for any $\eta \in \langle 0, 1\rangle$.

(ii) If $\mathbf{r}'\boldsymbol{\tau}_1 \succeq_{SSD} \mathbf{r}'\boldsymbol{\tau}_2$ and $\mathbf{r}'\boldsymbol{\tau}_1 \succeq_{SSD} \mathbf{r}'\boldsymbol{\tau}_3$ then $\mathbf{r}'\boldsymbol{\tau}_1 \succeq_{SSD} \mathbf{r}'(\eta\boldsymbol{\tau}_2 + (1-\eta)\boldsymbol{\tau}_3)$ and

$$D^*(\eta\boldsymbol{\tau}_2 + (1-\eta)\boldsymbol{\tau}_3) \leq \eta D^*(\boldsymbol{\tau}_2) + (1-\eta)D^*(\boldsymbol{\tau}_3)$$

for any $\eta \in \langle 0, 1\rangle$.

Proof:

(i) Applying Lemma 4.1 for equiprobable scenario approach, we obtain

$$\mathbf{r}'\boldsymbol{\tau}_1 \succeq_{SSD} \mathbf{r}'\boldsymbol{\tau}_2 \Rightarrow \mathbf{r}'\boldsymbol{\tau}_1 \succeq_{SSD} \mathbf{r}'\left(\eta\boldsymbol{\tau}_1 + (1-\eta)\boldsymbol{\tau}_2\right) \succeq_{SSD} \mathbf{r}'\boldsymbol{\tau}_2$$

for any $\eta \in \langle 0, 1\rangle$. By analogy to the proof of previous theorem, if $\boldsymbol{\lambda}^*(\boldsymbol{\tau}_1) \in \Lambda^*(\boldsymbol{\tau}_1)$ then $\boldsymbol{\lambda}^*(\boldsymbol{\tau}_1) \in \Lambda^*(\boldsymbol{\tau}_2)$ and $\boldsymbol{\lambda}^*(\boldsymbol{\tau}_1) \in \Lambda^*(\eta\boldsymbol{\tau}_1 + (1-\eta)\boldsymbol{\tau}_2)$. Hence

$$\begin{aligned}
D^*(\eta\boldsymbol{\tau}_1 + (1-\eta)\boldsymbol{\tau}_2) &= \sum_{k=0}^{T-1} \mathrm{CVaR}_{\frac{k}{T}}\left(-\mathbf{r}'[\eta\boldsymbol{\tau}_1 + (1-\eta)\boldsymbol{\tau}_2]\right) - \mathrm{CVaR}_{\frac{k}{T}}(-\mathbf{r}'\boldsymbol{\lambda}^*(\boldsymbol{\tau}_1)) \\
D^*(\boldsymbol{\tau}_1) &= \sum_{k=0}^{T-1} \mathrm{CVaR}_{\frac{k}{T}}(-\mathbf{r}'\boldsymbol{\tau}_1) - \mathrm{CVaR}_{\frac{k}{T}}(-\mathbf{r}'\boldsymbol{\lambda}^*(\boldsymbol{\tau}_1)) \\
D^*(\boldsymbol{\tau}_2) &= \sum_{k=0}^{T-1} \mathrm{CVaR}_{\frac{k}{T}}(-\mathbf{r}'\boldsymbol{\tau}_2) - \mathrm{CVaR}_{\frac{k}{T}}(-\mathbf{r}'\boldsymbol{\lambda}^*(\boldsymbol{\tau}_1))
\end{aligned}$$

Combining it with convexity of CVaR (see Lemma 4.5), we obtain

$$\begin{aligned} D^*(\eta\boldsymbol{\tau}_1 + (1-\eta)\boldsymbol{\tau}_2) &= \sum_{k=0}^{T-1} \mathrm{CVaR}_{\frac{k}{T}}\left(-\mathbf{r}'[\eta\boldsymbol{\tau}_1 + (1-\eta)\boldsymbol{\tau}_2]\right) - \mathrm{CVaR}_{\frac{k}{T}}(-\mathbf{r}'\boldsymbol{\lambda}^*(\boldsymbol{\tau}_1)) \\ &\leq \eta \sum_{k=0}^{T-1} \mathrm{CVaR}_{\frac{k}{T}}\left(-\mathbf{r}'\boldsymbol{\tau}_1\right) + (1-\eta) \sum_{k=0}^{T-1} \mathrm{CVaR}_{\frac{k}{T}}\left(-\mathbf{r}'\boldsymbol{\tau}_2\right) \\ &\quad -\eta \sum_{k=0}^{T-1} \mathrm{CVaR}_{\frac{k}{T}}(-\mathbf{r}'\boldsymbol{\lambda}^*(\boldsymbol{\tau}_1)) - (1-\eta)\mathrm{CVaR}_{\frac{k}{T}}(-\mathbf{r}'\boldsymbol{\lambda}^*(\boldsymbol{\tau}_1)) \\ &\leq \eta D^*(\boldsymbol{\tau}_1) + (1-\eta) D^*(\boldsymbol{\tau}_2). \end{aligned}$$

(ii) Applying Lemma 4.1 for scenario approach, we obtain:

$$\mathbf{r}'\boldsymbol{\tau} \succeq_{SSD} \mathbf{r}'\boldsymbol{\lambda} \Leftrightarrow \sum_{t=1}^{T} (\mathbf{x}^t\boldsymbol{\tau} - \mathbf{x}^t\boldsymbol{\lambda}) \geq 0 \quad \forall t = 1, 2, ..., T. \tag{4.21}$$

Hence

$$\begin{aligned} \sum_{t=1}^{T} (\mathbf{x}^t\boldsymbol{\tau}_1 - \mathbf{x}^t\boldsymbol{\tau}_2) &\geq 0 \quad \forall t = 1, 2, ..., T \\ \sum_{t=1}^{T} (\mathbf{x}^t\boldsymbol{\tau}_1 - \mathbf{x}^t\boldsymbol{\tau}_3) &\geq 0 \quad \forall t = 1, 2, ..., T \end{aligned}$$

and therefore

$$\sum_{t=1}^{T} (\mathbf{x}^t\boldsymbol{\tau}_1 - \eta\mathbf{x}^t\boldsymbol{\tau}_2 - (1-\eta)\mathbf{x}^t\boldsymbol{\tau}_3) \geq 0 \quad \forall t = 1, 2, ..., T$$

for any $\eta \in \langle 0, 1\rangle$. Thus, according to Lemma 4.1,

$$\mathbf{r}'\boldsymbol{\tau}_1 \succeq_{SSD} \mathbf{r}'(\eta\boldsymbol{\tau}_2 + (1-\eta)\boldsymbol{\tau}_3) \text{ for any } \eta \in \langle 0, 1\rangle.$$

Similarly to the proof of previous theorem, if $\boldsymbol{\lambda}^*(\boldsymbol{\tau}_1) \in \Lambda^*(\boldsymbol{\tau}_1)$ then $\boldsymbol{\lambda}^*(\boldsymbol{\tau}_1) \in \Lambda^*(\boldsymbol{\tau}_2)$, $\boldsymbol{\lambda}^*(\boldsymbol{\tau}_1) \in \Lambda^*(\boldsymbol{\tau}_3)$ and $\boldsymbol{\lambda}^*(\boldsymbol{\tau}_1) \in \Lambda^*(\eta\boldsymbol{\tau}_2 + (1-\eta)\boldsymbol{\tau}_3)$ for any $\eta \in \langle 0, 1\rangle$ and the rest of the proof follows by analogy to (i). □

Let $I(\boldsymbol{\tau})$ be a set of all portfolios whose yields are SSD dominated by yield of $\boldsymbol{\tau}$, i.e.

$$I(\boldsymbol{\tau}) = \{\boldsymbol{\lambda} \in \Lambda | \mathbf{r}'\boldsymbol{\tau} \succeq_{SSD} \mathbf{r}'\boldsymbol{\lambda}\}.$$

Theorem 4.19 shows that $I(\boldsymbol{\tau})$ is convex and D^* is convex on $I(\boldsymbol{\tau})$ for any $\boldsymbol{\tau} \in \Lambda$. Both these properties are consequences of convexity of CVaR. The following example illustrates these results and we stress the fact that the set of SSD efficient portfolios is not convex.

Example 4.20:

Consider three assets with three scenarios:

$$X = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 2 & 7 & 5 \end{pmatrix}.$$

It is easy to check that $\boldsymbol{\lambda}_1 = (1,0,0)'$, $\boldsymbol{\lambda}_2 = (0,1,0)'$ and $\boldsymbol{\lambda}_3 = (0,0,1)'$ are SSD efficient. Let $\boldsymbol{\tau}_1 = \boldsymbol{\lambda}_3$, $\boldsymbol{\tau}_2 = (\frac{1}{2}, \frac{1}{2}, 0)'$ and let $\boldsymbol{\tau}_3 = (\frac{1}{3}, \frac{2}{3}, 0)'$. Then $X\boldsymbol{\tau}_2 = (-\frac{1}{2}, \frac{1}{2}, \frac{9}{2})$ and according to (4.21), $\mathbf{r}'\boldsymbol{\tau}_1 \succ_{SSD} \mathbf{r}'\boldsymbol{\tau}_2$. Hence the set of SSD efficient portfolios is not convex. Similarly, $\mathbf{r}'\boldsymbol{\tau}_1 \succ_{SSD} \mathbf{r}'\boldsymbol{\tau}_3$ and $\mathbf{r}'\boldsymbol{\tau}_1 \succeq_{SSD} \mathbf{r}'\boldsymbol{\tau}_1$. Applying Theorem 4.19, a set of convex combinations of $\boldsymbol{\tau}_1$, $\boldsymbol{\tau}_2$, $\boldsymbol{\tau}_3$ is a subset of $I(\boldsymbol{\tau}_1)$. We will show that $I(\boldsymbol{\tau}_1)$ consists only of convex combinations of $\boldsymbol{\tau}_1$, $\boldsymbol{\tau}_2$ and $\boldsymbol{\tau}_3$, i.e.

$$I(\boldsymbol{\tau}_1) = \{\boldsymbol{\lambda} \in \Lambda | \lambda = \eta_1\boldsymbol{\tau}_1 + \eta_2\boldsymbol{\tau}_2 + \eta_3\boldsymbol{\tau}_3, \quad \eta_i \geq 0, \quad i = 1,2,3, \quad \sum_{i=1}^{3} \eta_i = 1\}$$

Substituting into (4.21) we can see that only portfolios $\boldsymbol{\lambda} \in \Lambda$ satisfying the following system of inequalities can be included in $I(\boldsymbol{\tau}_1)$:

$$\begin{aligned} -\lambda_2 &\leq 0 \\ \lambda_1 - \lambda_2 &\leq 0 \\ 3\lambda_1 + 6\lambda_2 + 5(1 - \lambda_1 - \lambda_2) &\leq 5 \end{aligned}$$

The grafical solution of this system is illustrated on the following figure and we can see that the set of portfolios which yields are SSD dominated by yield of portfolio $\boldsymbol{\tau}_1$ is equal to the set of all convex combinations of portfolios $\boldsymbol{\tau}_1$, $\boldsymbol{\tau}_2$, $\boldsymbol{\tau}_3$. Points A, B and C correspond to portfolios $\boldsymbol{\tau}_2$, $\boldsymbol{\tau}_3$, $\boldsymbol{\tau}_1$, respectively.

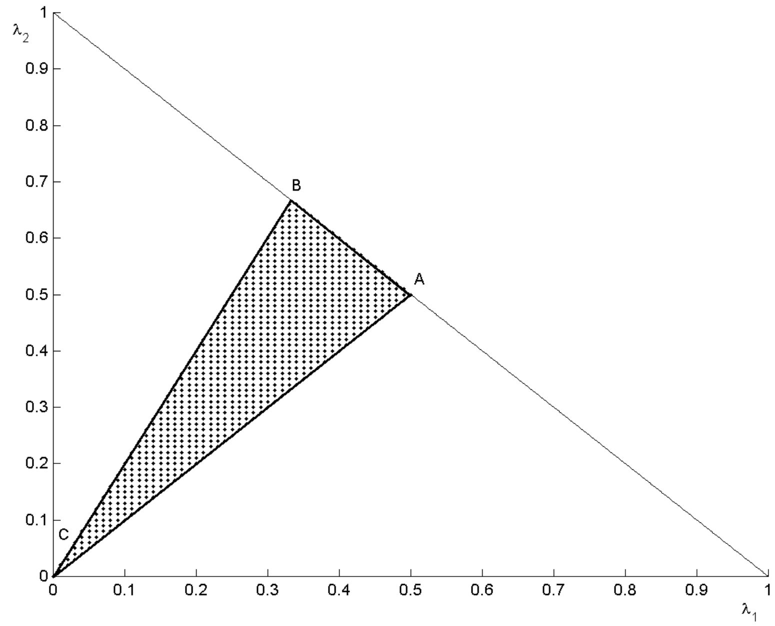

Figure 4.1: The set $I(\boldsymbol{\tau}_1)$ of portfolios whose yields are SSD dominated by yield of portfolio $\boldsymbol{\tau}_1 = (0, 0, 1)$.

As was shown in Theorem 4.19 (ii), SSD portfolio inefficiency measure D^* is convex on $I(\boldsymbol{\tau}_1)$. The following figure shows the graph of D^* on $I(\boldsymbol{\tau}_1)$. Since $\boldsymbol{\tau}_1$ is SSD efficient, $D^*(\boldsymbol{\tau}_1) = 0$ and $D^*(\boldsymbol{\tau}) > 0$ for all $\boldsymbol{\tau} \in I(\boldsymbol{\tau}_1)\backslash\{\boldsymbol{\tau}_1\}$. It is easy to check that X is portfolio-monotone with identical permutation. Hence, according to Lemma 4.10, (4.19) can be considered as a parametric linear problem where the parameters $\mathrm{CVaR}_{\frac{k-1}{T}}(-\mathbf{r}'\boldsymbol{\tau}), k = 1, 2, ..., T$ are only in the right hand side of the constraints. The duality theory in parametric linear programming implies linearity of $D^*(\boldsymbol{\tau})$ on $I(\boldsymbol{\tau}_1)$, because $I(\boldsymbol{\tau}_1)$ is a subset of the area of stability for $\boldsymbol{\lambda}^*(\boldsymbol{\tau}_1)$, i.e. $\boldsymbol{\lambda}^*(\boldsymbol{\tau}_1) \in \Lambda^*(\boldsymbol{\tau})$ for all $\boldsymbol{\tau} \in I(\boldsymbol{\tau}_1)$. See Grygarová [17] for more details about parametric linear programming.

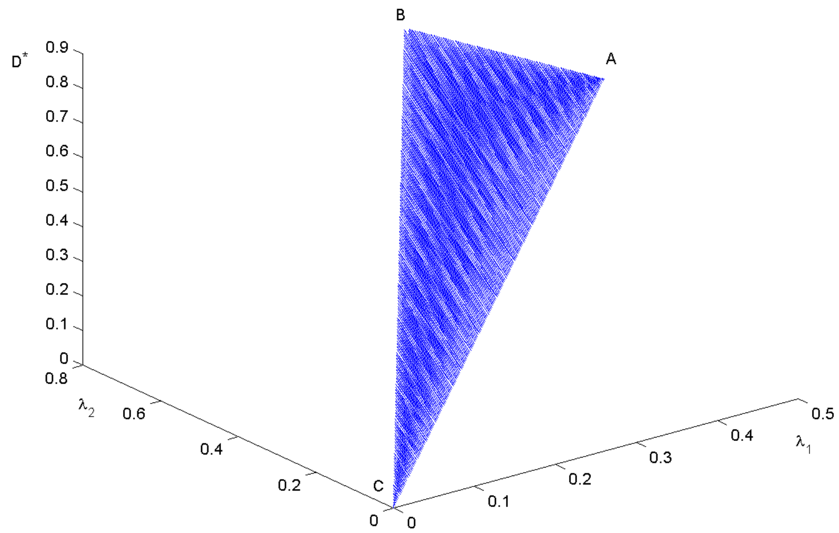

Figure 4.2: The graph of D^* on $I(\boldsymbol{\tau}_1)$.

4.7 Numerical application: SSD efficiency of mean-VaR optimal portfolios

According to Lemma 4.2 and (4.2) we can see that a portfolio with minimal VaR is FSD efficient. When searching for portfolio with minimal VaR under condition of a minimal level of expected yield this property may disappear.

We define *mean-VaR optimal portfolio* $\boldsymbol{\lambda}^{\mathrm{VaR}}(\mu) \in \Lambda$ as a portfolio with minimal VaR and a prescribed minimal level of expected yield μ, i.e. $\boldsymbol{\lambda}^{\mathrm{VaR}}(\mu)$ is an optimal solution of the problem:

$$\begin{aligned} \min_{\boldsymbol{\lambda}} \mathrm{VaR}_{\alpha}(-\mathbf{r}'\boldsymbol{\lambda}) & \\ \text{s.t.} \quad \frac{1}{T}\sum_{t=1}^{T} \mathbf{x}^t \boldsymbol{\lambda} & \geq \mu \\ \boldsymbol{\lambda} \in \Lambda. & \end{aligned}$$

Inspired by Gaivoronski & Pflug [14] we rewrite this problem as the mixed-integer linear program:

$$\begin{aligned} & \min_{\delta,\boldsymbol{\lambda},\zeta^t} \delta & & & (4.22) \\ \text{s.t.} \quad - \mathbf{x}^t\boldsymbol{\lambda} & \leq \delta + M\zeta^t, & & t = 1, \dots, T \\ \sum_{t=1}^{T} \zeta^t & = \lfloor (1-\alpha)T \rfloor \\ \frac{1}{T}\sum_{t=1}^{T} \mathbf{x}^t\boldsymbol{\lambda} & \geq \mu \\ \boldsymbol{\lambda} & \in \Lambda \\ \zeta^t & \in \{0,1\}, & & t = 1, \dots, T, \end{aligned}$$

where M is a sufficiently large constant:

$$M \geq \max_{i,j} y_j^i - \min_{i,j} y_j^i$$

and $\lfloor z \rfloor$ denotes the largest integer number which does not exceed z. Parameter μ represents a prescribed minimal level of expected yield of the portfolio. We shall examine SSD efficiency of these mean-VaR optimal portfolios using Theorem 4.16 and (4.19).

The data were obtained from http://finance.yahoo.com and consisted of 530 observations (07.1.1995–28.1.2005) of weakly yields of five U.S. stocks:IBM (International Business Machines), UTX (United Technologies), MMM (3M Company), JNJ (Johnson and Johnson) and CAT (Caterpillar Inc).

We move a window through the data with bandwith 210 and step 20. Thus we have 17 partial data sets. The number of observations in a partial data set corresponds to 4 years history. To track at least partly the behavior of the optimal mean-VaR portfolios in dependence on the parameter μ – the minimal required expected yield of the portfolio – we choose 5 levels of parameter μ for each partial data set. Thus we have to solve (4.22) 85 times.

Let ν_j^l denote the expected yield of j-th asset for l-th data set. Set

$$\underline{\nu}^l = \min_j \nu_j^l, \qquad \overline{\nu}^l = \max_j \nu_j^l, \qquad l = 1, 2, \dots, 17.$$

We set the levels of parameter μ using the following formulas:

$$
\begin{aligned}
\mu_1^l &= \underline{\nu}^l & l = 1, 2, \ldots, 17, \\
\mu_2^l &= \underline{\nu}^l + 0.5(\overline{\nu}^l - \underline{\nu}^l) & l = 1, 2, \ldots, 17, \\
\mu_3^l &= \underline{\nu}^l + 0.6(\overline{\nu}^l - \underline{\nu}^l) & l = 1, 2, \ldots, 17, \\
\mu_4^l &= \underline{\nu}^l + 0.7(\overline{\nu}^l - \underline{\nu}^l) & l = 1, 2, \ldots, 17, \\
\mu_5^l &= \underline{\nu}^l + 0.8(\overline{\nu}^l - \underline{\nu}^l) & l = 1, 2, \ldots, 17.
\end{aligned}
$$

Problem (4.22) has 210 integer variables (210 scenarios), 6 other variables and 218 constraints. The computations were done in GAMS solver CoinCbc and CPLEX. Using 2 GHz computer with 512 MB RAM, solving of problem (4.22) took at most 30 seconds for each data set and we obtained 65 different mean-VaR optimal portfolios. Then we tested the SSD efficiency of these portfolios applying (4.19). Since we consider 210 scenarios these linear programs have more than 40000 variables and constraints. The computation took approximately 10 minutes. Applying criterion for testing SSD efficiency suggested in Kuosmanen [34], we solved linear program with more than 40000 constraints and 130000 variables (see Proposition 4.14). Using the same computer as in the case of our test, the computation took approximately 40 minutes. We can see the results in Table 4.1 where "E" denotes SSD efficient portfolios and "I" SSD inefficient ones. From this table we can see that only 25 of 85 (29 %) mean-VaR optimal portfolios are SSD efficient. Especially for small required minimal expected yield of portfolio ($\mu_1^l, \mu_2^l, \mu_3^l, \mu_4^l$) mean-VaR optimal portfolios are SSD inefficient in 78 % cases. If the following portfolio selection problem with $u \in U_2$

$$
\begin{aligned}
&\max_{\boldsymbol{\lambda}} Eu(\mathbf{r}'\boldsymbol{\lambda}) \\
&\quad \text{s.t.} \quad \boldsymbol{\lambda} \in \Lambda
\end{aligned}
\tag{4.23}
$$

has unique solution then SSD inefficient portfolio cannot be an optimal solution of this problem. Thus mean-VaR optimal portfolios are not very suitable for risk averse investors.

If we compare time period before and after September 11, 2001, we have 43 % SSD efficient portfolios before the date and only 18 % after the date. This is caused by greater fluctuation of yields and losses after this date because VaR method does not take into account the magnitude of large losses.

Finally, we can see that mean-VaR optimal portfolios with high level of required minimal expected yield (μ_5^l) are more often SSD efficient than the others. This can be explained by the fact that investor accepts higher risk in this case, i.e. the requirement of minimal risk measured by VaR has

Time period	Minimal expected yield				
	μ_1^l	μ_2^l	μ_3^l	μ_4^l	μ_5^l
07. 01. 1995 – 31. 12. 1998	E	E	I	I	I
27. 05. 1995 – 21. 05. 1999	I	I	I	I	E
14. 10. 1995 – 08. 10. 1999	E	I	I	I	E
02. 03. 1996 – 25. 02. 2000	I	E	E	E	E
20. 07. 1996 – 14. 07. 2000	E	I	E	E	E
07. 12. 1996 – 01. 12. 2000	I	I	I	I	E
19. 04. 1997 – 20. 04. 2001	I	E	I	E	E
06. 09. 1997 – 07. 09. 2001	I	I	I	I	I
17. 01. 1998 – 18. 01. 2002	I	E	E	E	I
06. 06. 1998 – 07. 06. 2002	I	I	I	I	E
24. 10. 1998 – 25. 10. 2002	I	I	I	I	I
13. 03. 1999 – 07. 03. 2003	I	I	I	I	E
31. 07. 1999 – 25. 07. 2003	I	I	I	I	E
18. 12. 1999 – 12. 12. 2003	I	I	I	E	E
06. 05. 2000 – 23. 04. 2004	I	I	I	I	I
23. 09. 2000 – 10. 09 .2004	I	I	I	I	I
10. 02. 2001 – 28. 01. 2005	I	I	I	I	I
Total number of SSD efficient portfolios	3	4	3	5	10

Table 4.1: SSD efficiency of mean-VaR optimal portfolios

less important impact than in the case of smaller required minimal expected yield. In Table 4.2, we show the values of SSD portfolio inefficiency measure D^* for all tested portfolios.

Time period	**Minimal expected yield**				
	μ_1^l	μ_2^l	μ_3^l	μ_4^l	μ_5^l
07. 01. 1995 – 31. 12. 1998	0	0	5.5927	27.7726	15.8304
27. 05. 1995 – 21. 05. 1999	6.8648	6.8648	17.6124	3.4762	0
14. 10. 1995 – 08. 10. 1999	0	7.2607	3.5876	6.2085	0
02. 03. 1996 – 25. 02. 2000	7.0251	0	0	0	0
20. 07. 1996 – 14. 07. 2000	0	10.7063	0	0	0
07. 12. 1996 – 01. 12. 2000	4.068	25.493	28.4394	1.9612	0
19. 04. 1997 – 20. 04. 2001	4.0144	0	5.4213	0	0
06. 09. 1997 – 07. 09. 2001	5.1081	5.1081	5.1081	15.0719	24.6189
17. 01. 1998 – 18. 01. 2002	14.6595	0	0	0	22.995
06. 06. 1998 – 07. 06. 2002	4.9033	4.9033	4.9033	42.2749	0
24. 10. 1998 – 25. 10. 2002	58.7302	59.2872	35.3060	37.9927	22.7875
13. 03. 1999 – 07. 03. 2003	13.4106	13.4106	13.4106	13.4106	0
31. 07. 1999 – 25. 07. 2003	10.9355	10.9355	10.9355	10.9355	0
18. 12. 1999 – 12. 12. 2003	12.0750	12.0750	11.2118	0	0
06. 05. 2000 – 23. 04. 2004	40.7849	40.7849	48.8619	49.8263	16.2724
23. 09. 2000 – 10. 09 .2004	41.9353	41.9353	45.302	45.302	43.5411
10. 02. 2001 – 28. 01. 2005	25.8776	61.6235	57.9779	32.1731	6.1322

Table 4.2: SSD portfolio inefficiency measure D^*.

Chapter 5

A portfolio efficiency test based on the first-order stochastic dominance optimality

5.1 Preliminaries

In Chapter 4, we analyzed portfolio efficiency with respect to the second-order stochastic dominance. This concept is based on the assumption that decision maker is risk averse. Since market portfolios turned out to be SSD inefficient (see e.g. Post [43]) the presence of non-risk averse decision makers has to be involved. A complication in testing FSD portfolio efficiency is that we must distinguish between efficiency criteria based on "admissibility" and "optimality". There is a subtle difference between these two concepts. According to Kopa & Post [32], an alternative is FSD admissible if and only if no other alternative is preferred by all nonsatiable decision-makers. A FSD admissibility test was presented in Kuosmanen [34]. Following an FSD optimality idea in Bawa *et al.* [3], an alternative is FSD optimal if and only if it is an optimal choice for at least some increasing utility function. For pairwise comparisons, the two concepts are identical. However, more generally, when multiple alternatives are available, FSD admissibility is a necessary but not sufficient condition for FSD optimality.

Section 5.2 presents basic assumptions and definitions. In section 5.3, we reformulate the FSD optimality criterion in terms of piecewise-constant representative utility functions. Section 5.4 develops a linear programming test for searching over all such functions in order to test FSD portfolio optimality and suggests several approaches to identifying the input to this test. To obtain a necessary and sufficient condition for FSD optimality we employ mixed-integer linear problems. Section 5.5 presents a mixed-integer linear programming algorithm for testing FSD optimality. Section 5.6 uses a numerical example to illustrate our test and compare it with two existing tests presented in Bawa *et al.* [3] and Kuosmanen [34].

5.2 FSD optimality versus FSD admissibility

We hold the notation from Chapter 4. The evaluated portfolio, denoted by $\boldsymbol{\tau} \in \Lambda$, is assumed to be risky. Testing optimality for a riskless portfolio is trivial, because we then only need to check if there exists some portfolio that achieves a higher minimum return than the riskless rate. If no such portfolio exists, the riskless alternative is the optimal solution for extreme risk averters and hence FSD optimal. Let

$$\underline{m} = \min_{t,n} x_n^t, \quad \overline{m} = \max_{t,n} x_n^t \quad \text{and} \quad k(\boldsymbol{\tau}) = \min\{t : (X\boldsymbol{\tau})^{[t]} > (X\boldsymbol{\tau})^{[1]}\}.$$

Since a positive linear transformation of an utility function does not change the set of optimal solutions of (4.23), without loss of generality, we may focus on the following set of standardized utility functions:

$$U_1(\boldsymbol{\tau}) \;=\; \{u \in U_1 : u(\underline{m}) = 0; \;\; u((X\boldsymbol{\tau})^{[t]}) - u((X\boldsymbol{\tau})^{[k(\boldsymbol{\tau})]}) = 1\}. \qquad (5.1)$$

Note that the standardization depends on the evaluated portfolio and hence will differ for evaluating different portfolios. Furthermore, the standardization requires utility to be strictly increasing at least somewhere in the interior of the range for the evaluated portfolio. This requirement is natural, because, testing optimality relative to all $u \in U_1$ is trivial. Specifically, every portfolio $\boldsymbol{\lambda} \in \Lambda$ is an optimal solution for $u_0 = I(x \geq (X\boldsymbol{\tau})^{[1]})$. Thus $U_1(\boldsymbol{\tau})$ is the largest subset of U_1 for which testing optimality is non-trivial.

Definition 5.1:

Portfolio $\boldsymbol{\tau} \in \Lambda$ is FSD optimal if and only if it is the optimal solution of (4.23) for at least some utility function $u \in U_1(\boldsymbol{\tau})$, i.e., there exists $u \in U_1(\boldsymbol{\tau})$

such that

$$\sum_{t=1}^{T} u(\mathbf{x}^t\boldsymbol{\tau}) - \sum_{t=1}^{T} u(\mathbf{x}^t\boldsymbol{\lambda}) \geq 0 \quad \forall \boldsymbol{\lambda} \in \Lambda.$$

Otherwise, $\boldsymbol{\tau}$ is FSD non-optimal.

According to Kuosmanen [34], we recall FSD admissibility definition based on existence of an alternative which is better than a given portfolio for all decision makers. FSD admissibility is a necessary condition for FSD optimality.

Definition 5.2:

Portfolio $\boldsymbol{\tau} \in \Lambda$ is FSD admissible if and only if there exists no $\boldsymbol{\lambda} \in \Lambda$ such that $(X\boldsymbol{\lambda})^{[t]} \geq (X\boldsymbol{\tau})^{[t]}$ for all $t = 1, 2, ..., T$ with strong inequality for at least some t.

The following necessary and sufficient condition for FSD admissibility using mixed-integer linear programming was derived in Kuosmanen [34].

Theorem 5.3:

Let $\boldsymbol{\tau} \in \Lambda$ and Π be the set of permutation matrices, i.e.

$$\Pi = \left\{ [P_{ij}]_{T\times T} \ : \ P_{ij} \in \{0,1\}, \sum_{i=1}^{T} P_{ij} = \sum_{j=1}^{T} P_{ij} = 1, \quad i,j = 1,2,...,T \right\}$$

Consider

$$\begin{aligned} \theta^1(\boldsymbol{\tau}) \quad &= \quad \max_{\boldsymbol{\lambda},P} \mathbf{1}'(X\boldsymbol{\lambda} - X\boldsymbol{\tau}) \qquad (5.2) \\ s.t. \quad & \quad \sum_{i=1}^{N} x_i^t \lambda_i \geq \sum_{j=1}^{T} P_{tj} \sum_{i=1}^{N} x_i^j \tau_i \qquad t = 1,2,...,T \\ & \quad P \in \Pi \\ & \quad \boldsymbol{\lambda} \in \Lambda. \end{aligned}$$

Portfolio $\boldsymbol{\tau}$ is FSD admissible if and only if $\theta^1(\boldsymbol{\tau}) = 0$.

5.3 Representative utility functions

This section reformulates the optimality criterion in terms of a set of elementary representative utility functions. For pairwise FSD comparisons, the set of three-piece linear utility functions is representative for all admissible utility functions, see Russel & Seo [50] for more details. In our portfolio context, with diversification allowed, a more general class of piecewise constant utility functions is relevant:

$$R_1(\boldsymbol{\tau}) = \{u \in U_1 | u(y) = \sum_{t=1}^{T} a_t I(y \geq (X\boldsymbol{\tau})^{[t]}), \quad \mathbf{a} \in A(\boldsymbol{\tau})\} \quad (5.3)$$

$$A(\boldsymbol{\tau}) = \{\mathbf{a} \in \mathbb{R}_+^T : \sum_{t=k(\boldsymbol{\tau})}^{T} a_t = 1, \quad (X\boldsymbol{\tau})^{[t]} = (X\boldsymbol{\tau})^{[s]} \quad \wedge \quad t < s \Rightarrow a_s = 0 \quad t, s = 1, 2, \ldots, T\} \quad (5.4)$$

where

$$I(y \geq y_0) = 1 \text{ for } y \geq y_0$$
$$= 0 \text{ otherwise.}$$

Theorem 5.4:

Portfolio $\boldsymbol{\tau} \in \Lambda$ is FSD optimal if and only if it is the optimal solution of (4.23) for at least some utility function $u \in R_1(\boldsymbol{\tau})$, i.e., there exists $u \in R_1(\boldsymbol{\tau})$ such that

$$\sum_{t=1}^{T} u(\mathbf{x}^t\boldsymbol{\tau}) - \sum_{t=1}^{T} u(\mathbf{x}^t\boldsymbol{\lambda}) \geq 0 \quad \forall \boldsymbol{\lambda} \in \Lambda.$$

Otherwise, $\boldsymbol{\tau}$ is FSD non-optimal.

Proof:

The sufficient condition follows directly from $R_1(\boldsymbol{\tau}) \subset U_1(\boldsymbol{\tau})$. To establish the necessary condition, suppose that $\boldsymbol{\tau}$ is optimal for $u(y) \in U_1(\boldsymbol{\tau})$ and let

$$u_R(y) = \sum_{t=1}^{T} a_t I(y \geq (X\boldsymbol{\tau})^{[t]}),$$

with $a_1 = u(X\boldsymbol{\tau})^{[1]}$, $a_t = 0$, $t = 2, \dots, k(\boldsymbol{\tau}) - 1$ and

$$a_t = u(X\boldsymbol{\tau})^{[t]} - u(X\boldsymbol{\tau})^{[t-1]}, \quad t \;=\; k(\boldsymbol{\tau}), \dots, T.$$

By construction, $u_R(y) \in R_1(\boldsymbol{\tau})$. Furthermore, $u_R(y) \leq u(y)$, $\forall y \in \langle \underline{m}, \overline{m} \rangle$ and $u_R(y) \;=\; u(y)$, for $y = (X\boldsymbol{\tau})^{[1]}, (X\boldsymbol{\tau})^{[2]}, \dots, (X\boldsymbol{\tau})^{[T]}$. Therefore,

$$\sum_{t=1}^{T} u_R(\mathbf{x}^t\boldsymbol{\tau}) - \sum_{t=1}^{T} u_R(\mathbf{x}^t\boldsymbol{\lambda}) \geq \sum_{t=1}^{T} u(\mathbf{x}^t\boldsymbol{\tau}) - \sum_{t=1}^{T} u(\mathbf{x}^t\boldsymbol{\lambda}) \qquad \forall \boldsymbol{\lambda} \in \Lambda.$$

Since $\boldsymbol{\tau}$ is optimal for $u(y) \in U_1(\boldsymbol{\tau})$, the RHS is nonnegative for all $\boldsymbol{\lambda} \in \Lambda$, and hence $\boldsymbol{\tau}$ is also optimal for $u_R(y) \in R_1(\boldsymbol{\tau})$, which completes the proof. □

The proof makes use of the fact that for a given portfolio $\boldsymbol{\tau}$ any utility function can be transformed into a piecewise constant function with increments only at $\mathbf{x}^t\boldsymbol{\tau}$, $t = 1, \dots, T$. This transformation doesn't affect the expected utility for the evaluated portfolio but it may lower the expected utility of other portfolios. Since the objective is to analyze if the evaluated portfolio is optimal for some utility function, only the representative utility functions need to be checked; all other utility functions are known to put the evaluated portfolio in a worse perspective than some representative utility function.

To illustrate the representation theorem, consider the cubic utility function $u(y) = 10 + y - 0.1y^2 + 0.05y^3$ and a portfolio with returns $(X\boldsymbol{\tau})^{[1]} = -5$, $(X\boldsymbol{\tau})^{[2]} = 1$ and $(X\boldsymbol{\tau})^{[3]} = 6$. Figure 1 shows a version of this function that is transformed such that it belongs to $U_1(\boldsymbol{\tau})$: $u_0(y) = 2.6 + 0.04y - 0.004y^2 + 0.002y^3$ (the solid line). Since the latter function is obtained after a positive linear transformation, it yields the same results as the former function. The dashed line gives the piecewise-constant function $u_R(y) = 2.087I(y \geq -5) + 0.546I(y \geq 1) + 0.454I(y \geq 6)$. This function is constructed such that it yields exactly the same utility levels for the evaluated portfolio as $u_0(y)$ does. Furthermore, the utility levels for all other portfolios are smaller than or equal to those for $u_0(y)$. Thus, if the evaluated portfolio is optimal for $u_0(y)$, then it is also optimal for $u_R(y)$. A similar analysis applies for every admissible utility function $u(y) \in U_1(\boldsymbol{\tau})$.

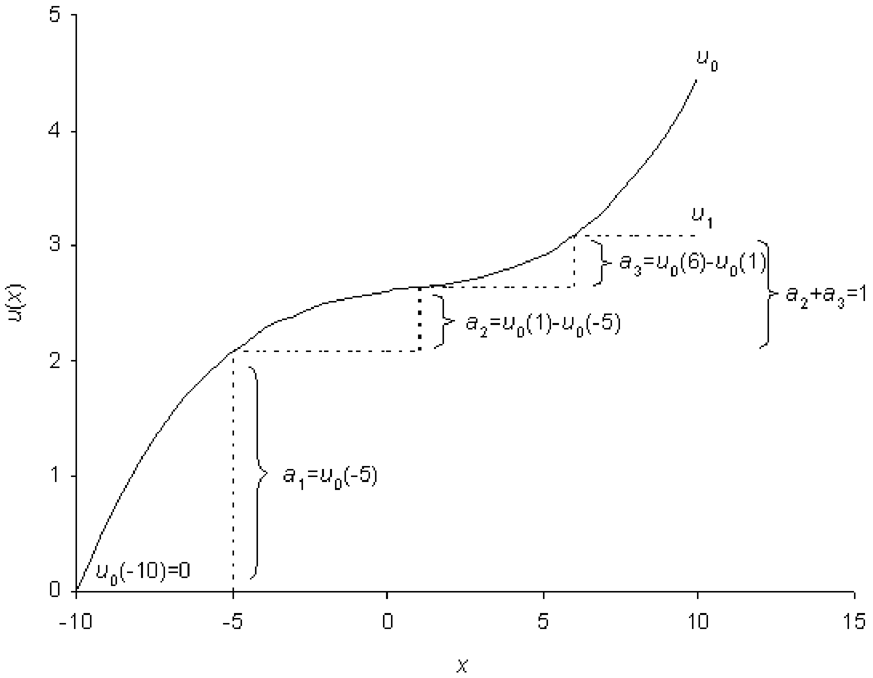

Figure 5.1: Representative utility function. The figure shows the original utility function u_0 and the associated representative utility function u_1.

Apart from replacing $U_1(\boldsymbol{\tau})$ with $R_1(\boldsymbol{\tau})$, we may also replace Λ with a reduced portfolio set that considers only portfolios with a higher minimum than the evaluated portfolio:

$$\Lambda^* = \left\{ \boldsymbol{\lambda} \in \Lambda : (X\boldsymbol{\tau})^{[1]} \leq (X\boldsymbol{\lambda})^{[1]} \right\}.$$

Using the representative utility functions and the reduced portfolio set, we can construct the following FSD inefficiency measure for any $\Lambda_0 \subseteq \Lambda^*$:

$$\xi(\boldsymbol{\tau}, \Lambda_0) = \frac{1}{T} \min_{u \in R_1(\boldsymbol{\tau})} \max_{\boldsymbol{\lambda} \in \Lambda_0} \sum_{t=1}^{T} \left(u(\mathbf{x}^t \boldsymbol{\lambda}) - u(\mathbf{x}^t \boldsymbol{\tau}) \right). \tag{5.5}$$

Replacing Λ with Λ^* reduces the parameter space and it causes no harm, because

$$\max_{\boldsymbol{\lambda} \in \Lambda} \sum_{t=1}^{T} \left(u(\mathbf{x}^t \boldsymbol{\lambda}) - u(\mathbf{x}^t \boldsymbol{\tau}) \right) = \max_{\boldsymbol{\lambda} \in \Lambda^*} \sum_{t=1}^{T} \left(u(\mathbf{x}^t \boldsymbol{\lambda}) - u(\mathbf{x}^t \boldsymbol{\tau}) \right)$$

for all $u \in R_1(\boldsymbol{\tau})$ with sufficiently large a_1 and we minimize the maximum of expected utility differences. If the evaluated portfolio has the maximal minimum then we can directly conclude that $\xi(\boldsymbol{\tau}, \Lambda^*) = 0$, i.e., the evaluated portfolio is FSD optimal (see the following Corollary).

Corollary 5.5:

(i) Portfolio $\boldsymbol{\tau}$ is FSD optimal if and only if $\xi(\boldsymbol{\tau}, \Lambda^*) = 0$. Otherwise, $\xi(\boldsymbol{\tau}, \Lambda^*) > 0$.

(ii) If $\Lambda_0 \subseteq \Lambda^*$ then $\xi(\boldsymbol{\tau}, \Lambda_0) \leq \xi(\boldsymbol{\tau}, \Lambda^*)$.

The next section will show that $\xi(\boldsymbol{\tau}, \Lambda^*)$ can be computed by solving a linear programming problem.

5.4 Mathematical programming formulation

There exist well-known, simple algorithms for establishing FSD-dominance relationships between a pair of choice alternatives; see, e.g., Levy [37]. Bawa *et al.* [3] derive a linear programming algorithm for FSD optimality relative to a discrete set of alternatives. Kuosmanen's [34] test for FSD admissibility in the portfolio context is computationally more demanding, because we need to account for changes to the ranking of the portfolio returns as the portfolio weights change, a task that requires integer programming. A similar complication arises for testing FSD optimality in a portfolio context. This section develops a linear programming test for testing portfolio optimality. However, the input to the linear programming test may require an initial phase of mixed integer linear programming (MILP) or subsampling.

Before presenting the algorithm, we stress that in some cases, simple necessary or sufficient conditions will suffice to classify the evaluated portfolio as efficient or inefficient. For example, a pairwise dominance relationship or an inefficiency classification by the Bawa *et al.* or the Kuosmanen tests suffice to conclude that the portfolio is FSD nonoptimal. Similarly, if the evaluated portfolio is classified as efficient according to a mean-variance test or a SSD test, we can conclude that the portfolio is FSD optimal.

Let

$$\begin{aligned}
h_s(\boldsymbol{\lambda},\boldsymbol{\tau}) &= \sum_{t=1}^{T} I(\mathbf{x}^t\boldsymbol{\lambda} \geq (X\boldsymbol{\tau})^{[s]}), \qquad s=1,\dots,T && (5.6)\\
\mathbf{h}(\boldsymbol{\lambda},\boldsymbol{\tau}) &= (h_1(\boldsymbol{\lambda},\boldsymbol{\tau}),\dots,h_T(\boldsymbol{\lambda},\boldsymbol{\tau})) && (5.7)\\
H(\boldsymbol{\tau}) &= \{\mathbf{h}\in\{0,\dots,T\}^T : \mathbf{h}=\mathbf{h}(\boldsymbol{\lambda},\boldsymbol{\tau}), \quad \boldsymbol{\lambda}\in\Lambda^*\}. && (5.8)
\end{aligned}$$

Since $h_s(\boldsymbol{\lambda},\boldsymbol{\tau})$ can take at most $T+1$ values $(0,1,\dots,T)$ for any $s=1,\dots,T$, the set $H(\boldsymbol{\tau})$ has a finite number of elements. For small-scale applications, identifying all elements is a fairly trivial task. However, for large-scale applications, the task is more challenging and can become computationally demanding. Some computational strategies to identifying the elements of $H(\boldsymbol{\tau})$ are discussed below. Interestingly, given $H(\boldsymbol{\tau})$, the test statistic $\xi(\boldsymbol{\tau},\Lambda^*)$ can be computed using simple linear programming. To see this, consider the following chain of equalities:

$$\begin{aligned}
\xi(\boldsymbol{\tau},\Lambda^*) &= \frac{1}{T}\min_{u\in R_1(\boldsymbol{\tau})}\max_{\boldsymbol{\lambda}\in\Lambda^*}\sum_{t=1}^{T}\left(u(\mathbf{x}^t\boldsymbol{\lambda})-u(\mathbf{x}^t\boldsymbol{\tau})\right)\\
&= \frac{1}{T}\min_{\mathbf{a}\in A(\boldsymbol{\tau})}\max_{\boldsymbol{\lambda}\in\Lambda^*}\sum_{t=1}^{T}\sum_{s=1}^{T}a_s\left(I(\mathbf{x}^t\boldsymbol{\lambda}\geq(X\boldsymbol{\tau})^{[s]})-I(\mathbf{x}^t\boldsymbol{\tau}\geq(X\boldsymbol{\tau})^{[s]})\right)\\
&= \frac{1}{T}\min_{\mathbf{a}\in A(\boldsymbol{\tau})}\max_{\boldsymbol{\lambda}\in\Lambda^*}\sum_{t=1}^{T}\sum_{s=k(\boldsymbol{\tau})}^{T}a_s\left(I(\mathbf{x}^t\boldsymbol{\lambda}\geq(X\boldsymbol{\tau})^{[s]})-I(\mathbf{x}^t\boldsymbol{\tau}\geq(X\boldsymbol{\tau})^{[s]})\right)\\
&= \frac{1}{T}\min_{\mathbf{a}\in A(\boldsymbol{\tau})}\max_{\boldsymbol{\lambda}\in\Lambda^*}\sum_{s=k(\boldsymbol{\tau})}^{T}a_s\left(\sum_{t=1}^{T}I(\mathbf{x}^t\boldsymbol{\lambda}\geq(X\boldsymbol{\tau})^{[s]})-\sum_{t=1}^{T}I(\mathbf{x}^t\boldsymbol{\tau}\geq(X\boldsymbol{\tau})^{[s]})\right)\\
&= \frac{1}{T}\min_{\mathbf{a}\in A(\boldsymbol{\tau})}\max_{\boldsymbol{\lambda}\in\Lambda^*}\sum_{s=k(\boldsymbol{\tau})}^{T}a_s(h_s(\boldsymbol{\lambda},\boldsymbol{\tau})-h_s(\boldsymbol{\tau},\boldsymbol{\tau}))\\
&= \frac{1}{T}\min_{\mathbf{a}\in A(\boldsymbol{\tau}),\delta}\left\{\delta : \sum_{s=k(\boldsymbol{\tau})}^{T}a_s(\overline{h}_s-h_s(\boldsymbol{\tau},\boldsymbol{\tau}))\leq\delta \quad \forall\,\overline{\mathbf{h}}\in H(\boldsymbol{\tau})\right\}.
\end{aligned}$$

The RHS of the final equality involves the minimization of a linear objective under a finite number of linear constraints. Thus, testing FSD optimality requires solving a simple linear programming problem and Corollary 5.5(i) implies the following sufficient and necessary condition for FSD optimality.

Theorem 5.6:

Let $H_0 \subseteq H(\boldsymbol{\tau})$. Let

$$\delta^*(H_0) = \min_{\mathbf{a} \in A(\boldsymbol{\tau})} \quad \delta \tag{5.9}$$

$$\text{s.t.} \quad \sum_{s=k(\boldsymbol{\tau})}^{T} a_s(\overline{h}_s - h_s(\boldsymbol{\tau}, \boldsymbol{\tau})) \quad \leq \quad \delta \qquad \forall \, \overline{\mathbf{h}} \in H_0. \tag{5.10}$$

Portfolio $\boldsymbol{\tau}$ is FSD optimal if and only if $\delta^*(H(\boldsymbol{\tau})) = 0$. If $\delta^*(H_0) > 0$ for some $H_0 \subseteq H(\boldsymbol{\tau})$ then $\boldsymbol{\tau}$ is FSD nonoptimal.

Note that $\xi(\boldsymbol{\tau}, \Lambda^*) = \delta^*/T$. Since $a \in A(\boldsymbol{\tau})$ and $\mathbf{h} \in \{0, \dots, T\}^T$ for all $\mathbf{h} \in H(\boldsymbol{\tau})$, using Corollary 5.5(i), we have $0 \leq \xi(\boldsymbol{\tau}, \Lambda^*) \leq 1$. A remaining problem is identifying elements of the set $H(\boldsymbol{\tau})$. We may adopt several strategies for this task. The next section provides a mixed-integer linear programming (MILP) algorithm that identifies a set of candidate vectors $\widetilde{H}(\boldsymbol{\tau}) \supseteq H(\boldsymbol{\tau})$, and checks if $\mathbf{h} \in H(\boldsymbol{\tau})$ for every candidate $\mathbf{h} \in \widetilde{H}(\boldsymbol{\tau})$. A drawback of this approach is that the number of candidates increases exponentially with the number of scenarios (T). Hence, for large numbers of scenarios, this strategy may become computationally prohibitive. Some sort of approximation may then be required, e.g. based on Corollary 5.5(ii).

For example, we may form a representative sample of elements $\mathbf{h} \in H(\boldsymbol{\tau})$ by using a sample $\Lambda_s \in \Lambda^*$ and constructing the associated values for $\mathbf{h}(\boldsymbol{\lambda}, \boldsymbol{\tau})$. According to Corollary 5.5(ii), this will lead to a necessary condition for FSD optimality. There exist various techniques for performing this task, ranging from a regular grid to Monte Carlo methods and Quasi-Monte Carlo methods (see, e.g., Jackel [21], and Glasserman [15]). Using regular grid in Kopa & Post [32], FSD optimality of US stock market portfolio relative to benchmark portfolios formed on market capitalization and book-to-market equity ratio was analyzed.

While the MILP algorithm starts from a large set of candidate vectors and checks feasibility for every candidate, sampling from the portfolio space avoids searching over infeasible candidates. Of course, the limitation of this strategy is that the critical sample size needed to obtain an accurate approximation increases exponentially as the number of individual choice alternatives (N) increases. Still, this approach can yield an accurate approximation in an efficient manner if N is low. This is true especially when the correlation between the individual choice alternatives is high and hence small changes in the portfolio weights do not lead to large changes in the values of $\mathbf{h}(\boldsymbol{\lambda}, \boldsymbol{\tau})$.

5.5 Mixed-integer Programming Algorithm for Testing FSD Optimality

This section provides a MILP algorithm for identifying the elements of $H(\boldsymbol{\tau})$ and suggests some stopping rules for testing FSD optimality of portfolios.

STEP 1: *Perform a FSD admissibility test*

Test FSD admissibility of $\boldsymbol{\tau}$, for example using the MILP test from Theorem 5.3. If $\boldsymbol{\tau}$ is FSD inadmissible then stop the algorithm; $\boldsymbol{\tau}$ is FSD non-optimal.

STEP 2: *Identify candidates for $H(\boldsymbol{\tau})$*

For all $j = k(\boldsymbol{\tau}), ..., T$ solve the following MILP problem:

$$\begin{array}{llll} \max & h_j + \frac{1}{T^2}\sum_{t\,=\,k(\boldsymbol{\tau})}^{T} h_t & & (5.11) \\ \text{s.t.} \quad (v_{s,t}-1)(\overline{m}-\underline{m}) \leq & \mathbf{x}^s\boldsymbol{\lambda} - (X\boldsymbol{\tau})^{[t]} \leq v_{s,t}(\overline{m}-\underline{m}) & s = 1,\ldots,T; & \\ & & t = k(\boldsymbol{\tau}),\ldots,T & \\ h_t = & \sum_{s\,=\,1}^{T} v_{s,t} & t = k(\boldsymbol{\tau}),\ldots,T & \\ v_{s,t} \in & \{0,1\} & s = 1,\ldots,T; & \\ & & t = k(\boldsymbol{\tau}),\ldots,T & \\ \boldsymbol{\lambda} \in & \Lambda^* & & \end{array}$$

Denote $(h_t^{*j}, \lambda_t^{*j}, v_{s,t}^{*j})$ the optimal solution of this problem. Let $\Lambda_1 \in \Lambda^*$ be a set of pairwise different $\boldsymbol{\lambda}^{*j}$ (all redundancy is excluded). Set

$$\begin{aligned} h_t^{max} &= \max_j h_t^{*j} \\ H_1 &= \{\mathbf{h}(\boldsymbol{\lambda}, \boldsymbol{\tau}): \quad \boldsymbol{\lambda} \in \Lambda_1\}. \end{aligned}$$

STEP 3: *Stopping rules*

Consider $\mathbf{h}(\boldsymbol{\tau}, \boldsymbol{\tau})$ as defined by (5.6)-(5.7). If there exists $t \in \{k(\boldsymbol{\tau}), \ldots, T\}$ such that $h_t^{max} \leq h_t(\boldsymbol{\tau}, \boldsymbol{\tau})$ then stop the algorithm; $\boldsymbol{\tau}$ is FSD optimal. Otherwise, solve problem (5.9)-(5.10) for $H_0 = H_1$. If $\delta^*(H_1) > 0$ then stop the algorithm; $\boldsymbol{\tau}$ is FSD non-optimal.

STEP 4: *Reduce the candidate set using a dominance rule*

Let $\overline{H}_t = \{0, 1, \ldots, h_t^{max}\}$. Denote by $\overline{H}$ the cartesian product of sets $\overline{H}_t$, i.e. $\overline{H} = \bigotimes_{k(\boldsymbol{\tau})}^{T} \overline{H}_t$. It is clear that $H(\boldsymbol{\tau}) \subseteq \overline{H}$. Let

$$\widetilde{H} = \left\{ \mathbf{h} \in \overline{H} | h_t \leq \xi h_j(\boldsymbol{\tau}, \boldsymbol{\tau}) + (1-\xi) \sum_{j = k(\boldsymbol{\tau})}^{T} \eta_j h_t^{*j}, \quad \forall t \in \{k(\boldsymbol{\tau}), \ldots, T\}, \right.$$

$$\left. \forall \mathbf{h}^{*j} \in H_1, \quad 0 \leq \xi \leq 1, \quad \sum_{j = k(\boldsymbol{\tau})}^{T} \eta_j = 1, \quad \eta_j \geq 0, \quad \forall j \in \{k(\boldsymbol{\tau}), \ldots, T\} \right\}.$$

Set $p = 1$.

STEP 5: *Check feasibility of the remaining candidates*

If $\overline{H} \setminus \widetilde{H}$ is empty, i.e. all possible $\mathbf{h} \in \overline{H}$ have been considered, then stop the algorithm; portfolio $\boldsymbol{\tau}$ is FSD optimal. Otherwise, choose $\mathbf{h} \in \overline{H} \setminus \widetilde{H}$ and add it to $\widetilde{H}$. If there exists a feasible solution of the system:

$$\begin{aligned}
(v_{s,t} - 1)(\overline{m} - \underline{m}) &\leq \mathbf{x}^s \boldsymbol{\lambda} - (X\boldsymbol{\tau})^{[t]} \leq v_{s,t}(\overline{m} - \underline{m}) & s &= 1, \ldots, T; \\
& & t &= t_1, \ldots, T \\
h_t &= \textstyle\sum_{s = 1}^{T} v_{s,t} & t &= t_1, \ldots, T \\
v_{s,t} &\in \{0, 1\} & s &= 1, \ldots, T; \\
& & t &= t_1, \ldots, T \\
\boldsymbol{\lambda} &\in \Lambda^*
\end{aligned} \tag{5.12}$$

put $p = p + 1$, $H_p = H_{p-1} \cup \mathbf{h}$ and go to the next step. Otherwise, repeat this step.

STEP 6: *Test optimality using the feasible candidates*

Solve problem (5.9)-(5.10) for $H_0 = H_p$. If $\delta^*(H_p) > 0$ then stop the algorithm; $\boldsymbol{\tau}$ is FSD non-optimal. Otherwise, go to Step 5.

5.6 Numerical example

A numerical example can illustrate our test and the difference with the Bawa *et al.* test and the Kuosmanen test. We focus on an example with five scenarios ($T = 5$), because FSD optimality is equivalent to FSD admissibility for ($T \leq 4$). To show this, let $T = 4$ and let $\boldsymbol{\tau}$ be FSD admissible. Since a dominated $\mathbf{h}(\boldsymbol{\lambda}, \boldsymbol{\tau})$ can not change the solution of (5.9)-(5.10) consider all possible $\mathbf{h}(\boldsymbol{\lambda}, \boldsymbol{\tau})$ which are not dominated by each other:

$$\begin{aligned} \mathbf{h}^1(\boldsymbol{\lambda}, \boldsymbol{\tau}) &= (4, 2, 2, 2) \\ \mathbf{h}^2(\boldsymbol{\lambda}, \boldsymbol{\tau}) &= (4, 3, 3, 0) \\ \mathbf{h}^3(\boldsymbol{\lambda}, \boldsymbol{\tau}) &= (4, 4, 2, 0) \\ \mathbf{h}^4(\boldsymbol{\lambda}, \boldsymbol{\tau}) &= (4, 4, 1, 1). \end{aligned}$$

Entering these candidates in the linear programming test in Theorem 5.6, we can see that $\boldsymbol{\tau}$ is the optimal portfolio for a representative utility function with $a_2 = a_3 = a_4 = 1/3$, and hence $\boldsymbol{\tau}$ is FSD optimal.

Table 5.1 shows the returns to three alternatives (X_1, X_2, X_3) and the tested portfolio $Z = 0.16X_1 + 0.21X_2 + 0.63X_3$ in the five scenarios $(1, 2, 3, 4, 5)$.

t	X_1	X_2	X_3	Z
1	-1	6	-4	-1.42
2	-2	5.90	2	2.18
3	3.50	2.20	3	2.91
4	8.70	2	5	4.96
5	10	7	7.50	7.80
Mean	3.84	4.62	2.70	3.29
St. dev.	5.46	2.34	4.30	3.42

Table 5.1: Scenarios and descriptive statistics for three alternatives and the tested portfolio

By comparing the means and standard deviations, we can immediately see that no individual alternative (X_1, X_2 or X_3) FSD dominates Z. However, this does not mean that Z is an efficient portfolio. Therefore, it is interesting to employ the three efficiency tests.

To implement the Kuosmanen test, we need to solve the following LP problem for each of the $5! = 120$ permutations of Z, say $\mathbf{y}_j = (y_j^1, y_j^2, y_j^3, y_j^4, y_j^5)$, $j = 1, 2, \ldots, 120$, or an equivalent mixed-integer linear problem:

$$\begin{aligned} \Psi_j = \max_{\lambda_1,\lambda_2,\lambda_3} \quad & \frac{1}{5}\sum_{t=1}^{5}(\lambda_1 x_1^t + \lambda_2 x_2^t + \lambda_3 x_3^t - y_j^t) \\ \text{s.t.} \quad \lambda_1 x_1^t + \lambda_2 x_2^t + \lambda_3 x_3^t &\geq y_j^t \quad t = 1, 2, 3, 4, 5 \\ \lambda_1 + \lambda_2 + \lambda_3 &= 1 \\ \lambda_1, \lambda_2, \lambda_3 &\geq 0 \end{aligned}$$

We find $\Psi_j = 0$ for every $j = 1, 2, \ldots, 120$, and hence Z is in the FSD admissible set (not FSD dominated by any convex combination of X_1, X_2 and X_3).

To implement the Bawa *et al.* test, we need to establish if some convex combination of the CDFs of X_1, X_2 and X_3 dominates the CDF of Z (see Bawa *et al.* [3]). Table 5.2 shows the CDFs of the three alternatives (Φ_{X_1}, Φ_{X_2}, Φ_{X_3}) and the CDF of Z (Φ_Z). Note that these CDFs need to be evaluated only at the observed return levels: $\{z_j\}_{j=1}^{19}$.

To test FSD optimality according to Bawa *et al.* [3] , we need to solve the following LP problem:

$$\eta = \max_{\lambda_1,\lambda_2,\lambda_3} \sum_{j=1}^{19}(\Phi_Z(z_j) - \lambda_1\Phi_{X_1}(z_j) - \lambda_2\Phi_{X_2}(z_j) - \lambda_3\Phi_{X_3}(z_j))$$

$$\begin{aligned} \text{s.t.} \quad \lambda_1\Phi_{X_1}(z_j) + \lambda_2\Phi_{X_2}(z_j) + \lambda_3\Phi_{X_3}(z_j) &\leq \Phi_Z(z_j) \quad j = 1, \ldots, 19 \\ \lambda_1 + \lambda_2 + \lambda_3 &= 1 \\ \lambda_1, \lambda_2, \lambda_3 &\geq 0. \end{aligned}$$

Solving this problem, we find $\eta = 0$, and hence Z is classified as efficient; not every nonsatiable decision-maker will prefer X_1 or X_2 or X_3 to Z. Based on the positive outcomes of the two tests, we may be tempted to conclude that Z is the optimal portfolio for some increasing utility function, i.e. FSD optimal. Perhaps surprisingly, this conclusion is wrong. The application of our MILP algorithm in section 5.5 will demonstrate this.

Since we have already tested FSD admissibility, we start with the second step: "Identify candidates for $H(\boldsymbol{\tau})$". For $j = 2, 3, 4, 5$, we solve (5.11)

j	z_j	Φ_{X_1}	Φ_{X_2}	Φ_{X_3}	Φ_Z
1	-4	0	0	1/5	0
2	-2	1/5	0	1/5	0
3	-1.42	1/5	0	1/5	1/5
4	-1	2/5	0	1/5	1/5
5	2	2/5	1/5	2/5	1/5
6	2.18	2/5	1/5	2/5	2/5
7	2.2	2/5	2/5	2/5	2/5
8	2.91	2/5	2/5	2/5	3/5
9	3	2/5	2/5	3/5	3/5
10	3.5	3/5	2/5	3/5	3/5
11	4.962	3/5	2/5	3/5	4/5
12	5	3/5	2/5	4/5	4/5
13	5.9	3/5	3/5	4/5	4/5
14	6	3/5	4/5	4/5	4/5
15	7	3/5	1	4/5	4/5
16	7.5	3/5	1	1	4/5
17	7.795	3/5	1	1	1
18	8.7	4/5	1	1	1
19	10	1	1	1	1

Table 5.2: Cumulative distribution functions of the three individual alternatives (X_1, X_2, X_3) and the tested portfolio Z for all observed return levels.

where $k(\boldsymbol{\tau}) = 2$, $T = 5$, $\underline{m} = -4$, $\overline{m} = 10$ and $X\boldsymbol{\tau} = Z$. Table 5.3 shows the optimal $\mathbf{h}(\boldsymbol{\lambda}, \boldsymbol{\tau})$ and optimal $\boldsymbol{\lambda}$. From Table 5.3, we can see that $\mathbf{h}^{max} = (5, 5, 4, 3, 2)$. In the third step we apply the stopping rules. Since $\mathbf{h}(\boldsymbol{\tau}, \boldsymbol{\tau}) = (5, 4, 3, 2, 1)$, $h_t^{max} > h_t(\boldsymbol{\tau}, \boldsymbol{\tau})$ for all $t = k(\boldsymbol{\tau}), ..., T$, hence the sufficient condition of FSD optimality is not fulfilled. Table 5.3 shows: $\Lambda_1 = \{(0.1483, 0.8517, 0), \quad (0.1187, 0.8813, 0), \quad (0.9266, 0.0734, 0)\}$. Let $H_1(\boldsymbol{\tau})$ be the set of corresponding values of $\mathbf{h}^*$, i.e., $H_1(\boldsymbol{\tau}) = \{(5, 5, 4, 2, 0), (5, 5, 3, 3, 0), \quad (5, 3, 3, 2, 2)\}$. Since $\xi(\boldsymbol{\tau}, \Lambda_1) = 0$, the necessary condition of FSD optimality is not fulfilled either. Thus we proceed with fourth step. Since $h_i(\boldsymbol{\lambda}, \boldsymbol{\tau}) \geq h_j(\boldsymbol{\lambda}, \boldsymbol{\tau})$ for all $i < j$, we can easily identify all candidates which satisfy the following conditions:

(i) are non-dominated by any convex combination of all $\mathbf{h} \in H_1(\boldsymbol{\tau}) \bigcup \mathbf{h}(\boldsymbol{\tau}, \boldsymbol{\tau})$

t	h_1^*	h_2^*	h_3^*	h_4^*	h_5^*	λ_1^*	λ_2^*	λ_3^*
2	5	5	4	2	0	0.1483	0.8517	0
3	5	5	4	2	0	0.1483	0.8517	0
4	5	5	3	3	0	0.1187	0.8813	0
5	5	3	3	2	2	0.9266	0.0734	0

Table 5.3: The initial candidates $H_1(\boldsymbol{\tau})$ and the associated $\Lambda_1(\boldsymbol{\tau})$ obtained in Step 2 of our algorithm.

(ii) are smaller than $\mathbf{h}(\boldsymbol{\tau},\boldsymbol{\tau})$ in at least one element (because $\boldsymbol{\tau}$ is FSD admissible)

(iii) are feasible for (5.11), i.e., the sum of elements of a candidate does not exceed the sum of elements of appropriate $\mathbf{h} \in H_1(\boldsymbol{\tau})$ and a candidate does not exceed $\mathbf{h}^{max}$ in any element.

The relevant candidates are:

$$\begin{aligned}
\mathbf{h}_c^1 &= (5,5,4,1,1)\\
\mathbf{h}_c^2 &= (5,5,2,2,2)\\
\mathbf{h}_c^3 &= (5,5,2,2,1)\\
\mathbf{h}_c^4 &= (5,5,2,1,1)\\
\mathbf{h}_c^5 &= (5,5,1,1,1)\\
\mathbf{h}_c^6 &= (5,4,4,1,1)\\
\mathbf{h}_c^7 &= (5,4,2,2,2)\\
\mathbf{h}_c^8 &= (5,3,3,3,1).
\end{aligned}$$

For these 8 candidates, we employ the last two steps of our algorithm. Step 5 tests feasibility of a candidate using (5.12). If the candidate is infeasible then we choose the next one. If the candidate is feasible then we add it to $H_1(\boldsymbol{\tau})$ and we recompute $\xi(\boldsymbol{\tau}, H_1(\boldsymbol{\tau}))$. Let us start with $\mathbf{h}_c^1 = (5,5,4,1,1)$. This candidate is feasible as it corresponds to $\boldsymbol{\lambda} = (0.265, 0.735, 0)$. Adding this candidate, we consider $\Lambda_2 = \Lambda_1 \cup (0.265, 0.735, 0)$ and $H_2(\boldsymbol{\tau}) = H_1(\boldsymbol{\tau}) \cup (5,5,4,1,1)$. Applying Theorem 5.6, we solve the following linear problem:

$$
\begin{array}{llllllll}
 & & & \min \delta & & & & \\
\text{s.t.} & a_2 & +a_3 & & -a_5 & \leq & \delta \\
 & a_2 & & +a_4 & -a_5 & \leq & \delta \\
 & -a_2 & & & +a_5 & \leq & \delta \\
 & a_2 & +a_3 & -a_4 & & \leq & \delta \\
 & a_2 & +a_3 & +a_4 & +a_5 & = & 1
\end{array}
$$

Since the optimal objective value of this problem $\delta^* = 1/9$,

$$\xi(\boldsymbol{\tau}, \Lambda_2) = \delta^*/5 = 1/45 > 0$$

and hence portfolio $\boldsymbol{\tau}$ is FSD non-optimal, which completes the algorithm. Thus, in this example, Z is classified as efficient according to the Bawa *et al.* and the Kuosmanen tests. Yet, it can be demonstrated to be not optimal for any increasing utility function.

We may repeat this exercise for more portfolios $\boldsymbol{\tau} \in \Lambda \cap \{0, 0.01, \ldots, 1\}^3$, i.e., when using a grid with step size 0.01 for the portfolio weights. Figure 5.2 illustrates the comparison between FSD admissibility and FSD optimality. The Kuosmanen test recognizes that many diversified portfolios are FSD dominated by other diversified portfolio, most notably those that assign a high weight to X_3. In this example, only 22 % of the considered portfolios are FSD admissible (the union of the grey and black dots). The FSD optimal set is even smaller than the admissible set. The set of grey dots, including Z, is now excluded, leaving only the black dots. The reduction in the efficient set to 16 % of all considered portfolios (a 26 % reduction) is possible because the optimality test acknowledges that an alternative may not be optimal for all investors even if no single other alternative is preferred by all. Note that the efficient regions are not convex, witness for example the small isolated optimal area near $\boldsymbol{\lambda} = (0, 0.7, 0.3)$.

A similar analysis can be done for FSD efficiency according to Bawa *et al.* [3]. Figure 5.3 shows that 93 % of all portfolios is classified as efficient. Only 17 % of these portfolios are FSD optimal.

The efficient set is substantially larger than ours, because the Bawa efficiency test does not account for diversification. Interestingly, only a few of inefficient portfolios according to the Bawa *et al.* test are FSD inadmissible. This suggest that one may use the Bawa *et al.* test as a complementary tool

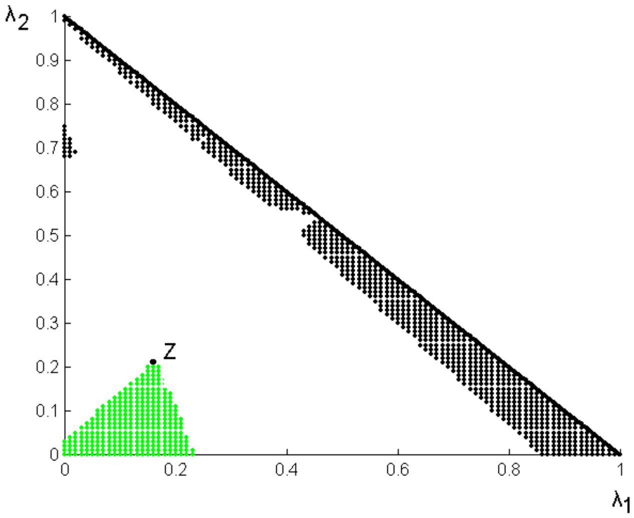

Figure 5.2: The FSD optimal set is represented by the black dots. The FSD admissible set is the union of the black dots and the grey dots.

to the FSD admissibility test. Still, portfolio Z proves that the FSD optimal set is even smaller than the intersection of these two FSD efficiency sets, i.e., a portfolio may be FSD non-optimal even if both of these tests classify it as efficient. Figure 5.4 shows all such portfolios in our example. The reduction of the efficient set (set of grey dots) is still quite large (8 %).

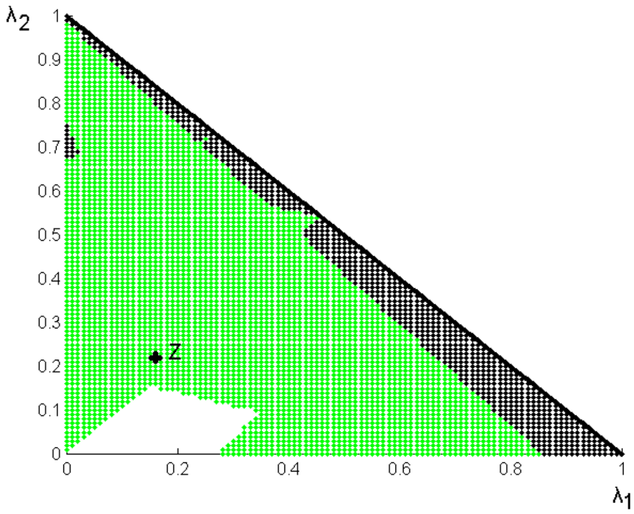

Figure 5.3: The FSD optimal set is represented by the black dots. The Bawa *et al.* efficient set is the union of the black dots and the grey dots.

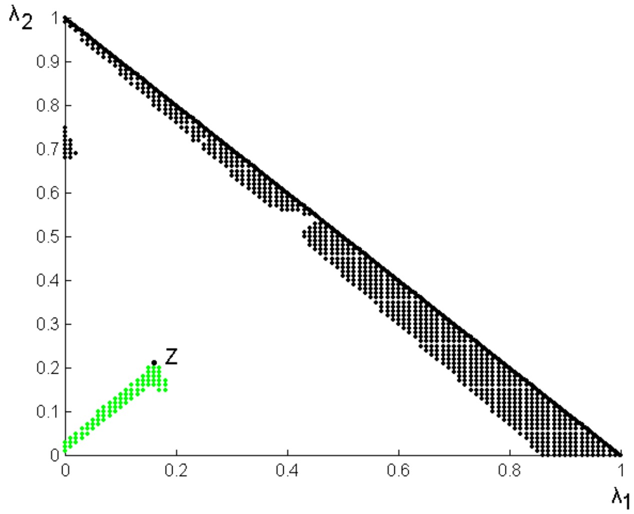

Figure 5.4: The FSD optimal set is represented by the black dots. The intersection of the Bawa *et al.* efficient and the Kuosmanen efficient set is the union of the black dots and the grey dots.

Chapter 6

Summary and open problems

In this thesis, utility functions in context of portfolio selection problems were analyzed. In practical studies, the perfect information about decision maker's utility function is usually not known. Therefore, we considered three the following situations.

Firstly, we assumed that an approximate information about utility function of a decision maker was known. Under assumption of twice differentiability of a utility function, we analyzed the stability of optimal solutions and optimal objective values of portfolio selection problem with respect to changes in Arrow – Pratt absolute risk aversion measure. Applying the theory of variational analysis, under assumption of hypoconvergence of utility functions, the limit set of optimal portfolios was analyzed. In comparison with general stability results in stochastic programming, we analyzed the stability with respect to perturbations of utility functions instead of changes in probability measures. These results allow us to apply approximate utility functions in solving portfolio selection problem and to judge the quality of these approximations.

We introduced a multiperiod risk premium as a measure of multiperiod risks. By analogy to classical univariate and multidimensional risk premiums, we analyzed its properties.

Secondly, we only assumed risk aversion of decision maker. We applied a concept of the second-order stochastic dominance and we were interested to classify a portfolio as SSD efficient or SSD inefficient. We said that portfolio had been SSD efficient if there was no better portfolio for all risk averse and risk neutral investors. Employing quantile model of the second-order stochastic dominance, we derived a linear programming algorithm for testing SSD efficiency of a given portfolio. This algorithm consisted of necessary conditions and a necessary and sufficient condition based on relationship between CVaR and SSD. It was faster than the Kuosmanen test and contrary to the Post criterion, it always detected the presence of SSD dominating portfolio which was SSD efficient. We introduced a SSD portfolio inefficiency measure which was consistent with SSD relation. It means that if an alternative was worse than the other alternative for all risk averse and risk neutral investors then it had a higher value of this measure. We also explored the convexity property of this measure.

Finally, we dropped all the assumptions about decision maker's risk attitude. We employed the first-order stochastic dominance approach. We discussed the differences between FSD admissibility and FSD optimality when any diversification across the assets was allowed. We derived a necessary and sufficient condition for FSD optimality via introducing the representative class of utility functions in the case of FSD with diversification. We suggested a mixed-integer linear programming algorithm and some subsampling techniques.

Dealing with stochastic dominance criteria in the context of portfolio efficiency, there are still some open problems. In this thesis, we assumed that the probability distribution of yields is known. However, we usually only approximate the unknown true probability distribution. Therefore a stability of SSD efficiency tests and the FSD optimality test with respect to perturbations in underlying probability measures are of interest. Another open area is connected with convexity of the set of efficient portfolios. It is known, that the set of SSD efficient portfolios is not convex. Of course, the set of FSD admissible or FSD optimal portfolios is not convex either. Therefore a new stochastic dominance relation which will guarantee convexity of the set of efficient portfolios can be another point of future research.

Bibliography

[1] K. J. Arrow: Aspects of the Theory of Risk-Bearing, Yrjo Jahnsson Foundation, Helsinki, 1965.

[2] R. Ambarish, J. G. Kallberg: Multivariate risk premiums, *Theory and Decision* **22** (1987), 77–96.

[3] V.S. Bawa, J.N. Bodurtha Jr., M.R. Rao and H.L. Suri: On Determination of Stochastic Dominance Optimal Sets, *Journal of Finance* **40** (1985), 417–431.

[4] D. Bell: One switch utility function and a measure of risk, *Management Science* **34**, 12 (1988), 1416–1424.

[5] J. A. Chalfant, I. Finkelshtain: Multivariate risk premiums, Portfolio Choices in the Presence of Other Risks, *Management Science* **39**, 8 (1993), 925–936.

[6] M. Černý, M. Vlach, K. Zimmermann: Axiomatic utility theory, SPN Praha, 1975. (in Czech)

[7] G. B. Dantzig: Linear programming and extensions, Princeton University Press, Princeton, New Jersey, 1963.

[8] D. Dentcheva, A Ruszczyński: Portfolio Optimization with stochastic dominance constraints, *Journal of Banking and Finance* **30**, 2 (2006), 433–451.

[9] G. T. Duncan: A matrix measure of multivariate local risk aversion, *Econometrica* **45** (1977), 895–903.

[10] J. Dupačová, J. Hurt, J. Štěpán: Stochastic Modeling in Economics and Finance, Kluwer Academic Publishers, Dordrecht 2002.

[11] A. Eichhorn, W. Rőmisch: Polyhedral risk measures in stochastic programming, *SIAM Journal on Optimization* **16** (2005), 69–95.

[12] P. C. Fishburn: A General Theory of Subjective Probabilities and Expected Utilities, *The annals of Mathematical Statistics* **40** (1969), 1419–1429.

[13] P.C. Fishburn: Convex Stochastic Dominance with Continuous Distribution Functions, *Journal of Economic Theory* **7** (1974), 143–158.

[14] A. A. Gaivoronski, G. Ch. Pflug: Value-at-Risk in portfolio optimization: properties and computational approach, *Journal of Risk* **7**, 2 (2005), 1–31.

[15] P. Glasserman: *Monte Carlo Methods in Financial Engineering*, Springer Verlag, 2004.

[16] D. Glűckaufová, M. Černý: Indirect Utility in Axiomatic Treatment, Research Paper No. 27, Czech Academy of Science Prague, 1970.

[17] L. Grygarová: Introduction to parametric programming, Karolinum MFF UK, Praha, 1994. (in Czech)

[18] G. Hanoch, H. Levy: The efficiency analysis of choices involving risk, *Review of Economic Studies* **36** (1969), 335–346.

[19] I. N. Herstein, J. Milnor: An Axiomatic Approach to Measurable Utility, *Econometrica* **21** (1953), 291–297.

[20] J. E. Ingersoll: Theory of financial decision making, Rowman & Littlefield, New Jersey, 1987.

[21] P. Jackel: *Monte Carlo Methods in Finance*, John Wiley & Sons, 2002.

[22] J. G. Kallberg, W. T. Ziemba: Comparison of alternative utility functions in portfolio selection problems, *Management Science* **29** (1983), 1257–1276.

[23] R. E. Kihlstrom, L. J. Mirman: Risk aversion with many commodities, *Journal of Economic Theory* **8** (1974), 361-388.

[24] R. E. Kihlstrom, D. Romer, S. Williams: Risk aversion with random initial wealth, *Econometrica* **49**, 4 (1981), 911–920.

[25] M. Kopa: Importance of utility function in stochastic programming, Diploma Thesis MFF UK, 2002. (in Czech)

[26] M. Kopa: Utility functions and portfolio selection problem, *Bulletin of the Czech Econometric Society* **19** (2003), 87 – 108.

[27] M. Kopa: Portfolio selection problem and stability of optimal portfolio, *WDS'03 Proceedings*, Part I (2003), 13 – 18.

[28] M. Kopa: Stability of optimal solution in portfolio selection problem, *Mathematical Methods in Economics 2003 Proceedings*, pp. 166 – 170.

[29] M. Kopa: Risk premiums for multiperiod risks, *Bulletin of the Czech Econometric Society* **21** (2004), 67 – 79.

[30] M. Kopa: Stochastic dominance and efficient portfolios, KPMS MFF UK preprint 52, 2006.

[31] M. Kopa: Stochastic dominance and CVaR in portfolio selection problem, *Mathematical Methods in Economics 2005 Proceedings*, pp. 211 – 216.

[32] M. Kopa, T. Post: A portfolio efficiency test based on FSD optimality, *Available at SSRN*: http://ssrn.com/abstract=720882.

[33] M. Kopa, T. Post: On portfolio FSD efficiency tests: admissibility versus optimality, *Proceedings of Prague Stochastics 2006*, to appear.

[34] T. Kuosmanen: Efficient diversification according to stochastic dominance criteria, *Management Science* **50**, 10 (2004), 1390–1406.

[35] M. Landsberger, I. Meilijson: Lotteries, insurance, and star-shaped utility functions, *Journal of economic theory* **52** (1990), 1–17.

[36] H. Levy: Stochastic dominance and expected utility: Survey and analysis, *Management Science* **38**, 4 (1992), 555–593.

[37] H. Levy: Stochastic Dominance, Kluwer Academic Publishers, Norwell, MA 1998.

[38] O. Linton, E. Maasoumi and Y.-J. Whang: Consistent testing for stochastic dominance under general sampling schemes, *Review of Economic Studies* **72** (2005), 735–765.

[39] H. M. Markowitz: Portfolio selection, *Journal of Finance* **7**, 1 (1952), 77–91.

[40] W. Ogryczak, A. Ruszczyński: Dual stochastic dominance and related mean-risk models, *SIAM J. Optim.* **13** (2002), 60–78.

[41] C. Pedersen, S. Satchel: Risk, utility and switching between gambles, Trinity College, University of Cambridge, working paper, 1997.

[42] G. Ch. Pflug: Some remarks on the value-at-risk and the conditional value-at-risk, Probabilistic Constrained Optimization: Methodology and Applications (S. Uryasev ed.), Kluwer Academic Publishers, Norwell, MA, (2000), pp. 278–287.

[43] T. Post: Empirical tests for stochastic dominance efficiency, *Journal of Finance* **58** (2003) 1905–1932.

[44] T. Post: On the Dual Test for SSD Efficiency: With an Application to Momentum Investment Strategies, *European Journal of Operational Research*, to appear.

[45] J. W. Pratt: Risk aversion in the small and in the large, *Econometrica* **32** (1964), 122–136.

[46] S. F. Richard: Multivariate risk aversion, utility dependence and separable utility functions, *Management Science* **21** (1975), 12-21.

[47] R. T. Rockafellar, R. J.-B. Wets: Variational Analysis, Springer Verlag, Berlin, 2004

[48] W. Rőmisch: Stability of Stochastic Programming Problems, Stochastic Programming (A. Ruszczyński and A. Shapiro eds.), Chapter 8 in Handbooks in Operations Research and Management Science, Elsevier, Amsterdam 2003.

[49] S. A. Ross: Some stronger measures of risk aversion in the small and the large with applications, *Econometrica* **49**, 3 (1981), 621–638.

[50] W. R. Russell, T. K. Seo: Representative sets for stochastic dominance rules, Studies in the Economics of Uncertainty: In honor of Josef Hadar (Thomas B. Fomby and Tae Kun Seo eds.), Springer Verlag, New York, (1989), 59–76.

[51] A. Ruszczyński, R. J. Vanderbei: Frontiers of stochastically nondominated portfolios, *Econometrica* **71**, 4 (2003), 1287–1297.

[52] M. Sion: On general minimax theorems, *Pacific Journal of Mathematics* **8** (1958), 171–176.

[53] S.Uryasev, R. T. Rockafellar: Conditional value-at-risk for general loss distributions, *Journal of Banking* & *Finance* **26** (2002), 1443–1471.

[54] J. von Neumann, O. Morgenstern: Theory of Games and Economic Behavior, Princeton University Press, 1944.

[55] G. A. Whitmore: Third degree stochastic dominance, *American Economic Review* **60** (1970), 457–459.

[56] G. A. Whitmore: Stochastic dominance for the class of completely monotonic utility functions, Studies in the Economics of Uncertainty (Thomas B. Fomby and Tae Kun Seo eds.), Springer Verlag, New York, (1970), 77–88.

[57] W. T. Ziemba, R. G. Vickson (eds.): Stochastic Optimization Models in Finance, Academic Press, New York, 1975.