STOCHASTIC DOMINANCE IN PORTFOLIO ANALYSIS AND ASSET PRICING

ISBN: 978 90 3610 187 5

Cover design: Crasborn Graphic Designers bno, Valkenburg a.d. Geul

This book is no. **487** of the Tinbergen Institute Research Series, established through cooperation between Thela Thesis and the Tinbergen Institute. A list of books which already appeared in the series can be found in the back.

STOCHASTIC DOMINANCE IN PORTFOLIO ANALYSIS AND ASSET PRICING

Stochastische dominatie in portefeuilleanalyse en de prijsvorming van effecten

Thesis

to obtain the degree of Doctor from the
Erasmus University Rotterdam
by command of the rector magnificus

Prof.dr. H.G. Schmidt

and in accordance with the decision of Doctoral Board.

The public defence shall be held on

Friday 12 November 2010 at 11:30 hrs

by

Andrey M. Lizyayev

Born in Sverdlovsk, USSR

ERASMUS UNIVERSITEIT ROTTERDAM

Doctoral Committee

Promoter:	Prof.dr. W.F.C. Verschoor
Other Members:	Prof.dr. A. Ruszczyńnski Prof.dr. J. Spronk Prof.dr. H.T.J. Smit
Copromoter:	Dr. P.J.P.M. Versijp

Acknowledgements

First of all I would like to thank my supervisor, Willem Verschoor: I certainly wouldn't have finished my Ph.D. without his involvement – thanks for your tremendous support and talking Dutch to me throughout! It was my greatest pleasure to collaborate with all my (co-)supervisors and co-authors from diverse scientific backgrounds. I benefited a lot from the mathematical rigour of Andrzej Ruszczyński, as well as the economic insights of Willem Verschoor and Philippe Versijp. I truly enjoyed the year of co-supervision with Timo Kuosmanen and sincerely value his indispensable feedback, even though we didn't always see eye to eye. I'm also thankful to Thierry Post for showing a great deal of trust in me by first selecting me for this Ph.D. project and subsequently for leaving me on my own in this burgeoning and constantly evolving world of academia. I'm also indebted to Kees Roos and Dick den Hertog for sharing their expertise in the Operational Research with me. Finally, I am truly grateful to Jaap Spronk for his advice and support during the turbulent period of the project.

Andrzej Ruszczyński's supervision period was short but very productive and insightful due to his wide-ranging expertise and his kindness and patience in sharing it with me. My working visit to Rutgers University has played a crucial role in the development of this Ph.D. project and I am sincerely grateful to the Business Economics section of Erasmus School of Economics and the Trustfonds association for sponsoring the visit. The standard financial support from Tinbergen Institute is also gratefully acknowledged.

I'd like to thank the supporting staff and secretaries for their assistance: Betty, Trudy, Hélèné, Ursula, Daphne, Cia, Shirley and Terry. A special thankyou goes to (Dr.) Lynn Agre for all her help in New Brunswick and kind assistance during my working visit. I can't forget to mention Tony, the best chef of a university dining hall – ever!

I enjoyed working in the multi-cultural environment of Tinbergen Institute among its Ph.D. students: Alexei, Andrei, Carlos, Diana, Felix, Ghulam, Jairo, Oleg, Ron and Simon, to name just a few.

I also thank my friends outside Rotterdam for making my life more enjoyable: Andrey (×3), Kostya, Afsal, Stephan, Chunfang, Winston, Dmitry, Lukas, Roman, Xavier, Ivan, Jay, Steve, Alex, Fred, Marianne, Gert, Waeil, Sela, Ru, Tamara, Ted, Svetlana, Bill, Ewie, Rogier, Edwin, Rochy, Francisco, Joel, Floor, Piet-Jan, Helena and Catherina.

A special mention goes to *mi bonito corazón*. Thanks for your immense support and putting up with me all these years! We've discovered and experienced so many nice things together, and that has inevitably left its mark on my life and transformed my tastes and outlook. I'm so delighted that our paths have crossed! Every success in the next chapter of your life! Whatever happens, remember Proposition XI.

Finally, the last in order but the foremost in importance, I owe the deepest gratitude to my mum Nina, aunt Tanja, Sergey, Lena, Sasha and the rest of my family for their support and encouragement. You're always on my mind!

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The problem of comparing and ordering various random outcomes represents a huge challenge in theoretical and applied research that occurs in numerous instances: in agriculture, crops yields need to be compared across several locations; in medicine, the most effective treatment or drug has to be selected from available alternatives; in poverty research, countries have to be classified according to the distribution of wealth in them. A common factor present in all such instances is uncertainty, as a result of which random vectors need to be compared rather than fixed indicators. For example, crop yield in each location is a result of (among other factors) weather conditions, available fertilizers and technology; medical treatments may affect specific subgroups (e.g. based on age, sex, body-mass index etc.) differently; and the degree of poverty of a particular country depends on the distribution of wealth among all its citizens, rather than its total or per capita wealth. Financial decision making is no exception. Indeed, an investor has to allocate her wealth among available assets based on their joint distribution over all possible states of nature.

If full information concerning the distribution of underlying variables of interest is available (as is assumed in the instances above), it is natural to use this full information in the decision-making process, rather than only certain characteristics of that distribution, such as its mean or variance. Stochastic Dominance (further SD) relation is a decision-making rule which uses that full information for the ordering of uncertain prospects. The idea of SD dates back to at least 1738, when Daniel Bernoulli (1738) suggested comparing random outcomes by transforming them into their utilities, before calculating expectations. The formulation of SD in terms of cumulative distribution functions and modern applications of the concept were introduced in mathematical literature by Mann and Whitney (1947) and Lehmann (1955), and in economics by Quirk and Saposnik (1962), Hadar and Russell (1969), Hanoch and Levy (1969) and Rothschild and Stiglitz (1970, 1971). Subsequently, for more than three decades, the SD approach was used for comparing mutually exclusive choice alternatives without the possibility of diversification. It was not until the beginning of the 21st century that the full diversification setting was solved. Applied in the area of finance and investments, the first algorithms for identifying if a given portfolio is efficient (non-dominated) relative to an infinite portfolio possibilities set, or for computing an efficient portfolio relative to such a set, were proposed in Dentcheva and Ruszczyński (2003), Post (2003) and Kuosmanen (2004). In Chapter 2 we review this development and classify the methods into three main categories: majorization, revealed preference and distribution-based approaches. Unfortunately, some of these schools of thought are developing independently, with little interaction or cross-referencing among them. Moreover, the methods differ in terms of their objectives, the information content of the results and their computational complexity. As a result, the relative merits of alternative approaches are difficult to compare. The same Chapter presents the first systematic review of all three approaches in a unified methodological framework. We examine the main developments in this emerging literature, critically evaluating the advantages and disadvantages of the alternative approaches. We also point out some misleading arguments and propose corrections and improvements to some of the methods considered.

Next to portfolio efficiency testing, Stochastic Dominance finds an important application in theoretical asset pricing, as it may lead to heterogeneous investors models which are much more realistic than the conventional Capital Asset Pricing Model. In that context the convexity of portfolio efficient sets turns out to play an important role. For this reason, in Chapter 3 we point out the importance of Stochastic Dominance (SD) efficient sets being convex. We review classic convexity and efficient set characterization results on the SD efficiency of a given portfolio relative to a diversified set of assets and generalize them in the following aspects. First, we broaden the class of individual utilities in Rubinstein (1974) which lead to two-fund separation. Secondly, we propose a linear programming SSD test that is more efficient than that of Post (2003) and expand the SSD efficiency criteria developed by Dybvig and Ross (1982) onto the Third Order Stochastic Dominance and further to Decreasing Absolute and Increasing Relative Risk Aversion Stochastic Dominance. We also elaborate on the structure of efficient sets for those refined classes of utility functions.

The fourth Chapter is devoted to Almost Stochastic Dominance. LL-Almost Stochastic Dominance (LL-ASD) is a relaxation of the Stochastic Dominance (SD) concept proposed by Leshno and Levy (2002) that explains more of realistic preferences observed in practice than SD alone does. Unfortunately, numerical applications of LL-ASD, such as identifying if a given portfolio is LL-ASD efficient, or determining a marketed portfolio that LL-ASD dominates a given benchmark, are computationally prohibitive due to the structure of LL-ASD. We propose a new Almost Stochastic Dominance (ASD) concept that is computationally tractable. For instance, a marketed dominating portfolio can be identified by solving a simple linear program. Moreover, the new ASD performs well on all the intuitive examples from Leshno and Levy (2002) and Levy (2009), and in some cases leads to more realistic predictions than those of LL-ASD. We develop some properties of ASD, formulate efficient optimization programs and apply the concept to analyzing investors' preferences between bonds and stocks for the long run.

In Chapter 5 we empirically apply the results from Chapters 2 through 4 for testing a periodic asset allocation strategy based on the Second order Stochastic Dominance (SSD) efficiency and compare its performance with other portfolio rebalancing strategies, such as Lower Partial Moments, Mean-Variance, Momentum, Value, Alpha, Beta, and passive investing on French's 48 industry portfolios. We observe that the SSD strategy performs reasonably well in terms of realized return and indicates out-of-sample persistence in restricting portfolio risk. Furthermore, extending the results of Grootveld and Hallerbach (1999), we find a substantial difference in efficient portfolios formed on the basis of downside risk criteria (linear lower partial moments and semideviation) and full moments (mean-variance), particularly in the case when short sales restrictions are relaxed.

Finally, in Chapter 6 we draw some conclusions and suggest new ideas for future research. In particular, we discuss another interesting modification of SD, namely: Robust Stochastic Dominance (further RSD). RSD is based on Robust Programming, the area of Operational Research that has recently been gaining considerable popularity and momentum. In addition to the standard SD efficiency, RSD ensures that the efficient portfolio is more robust to data perturbations. Another virtue of RSD lies in its ability to incorporate uncertainty about the probabilities of the states of nature (which are usually assumed equally likely), and for instance to place greater importance on recent events and progressively less emphasis on the historical ones.

Chapter 2

SD Efficiency Analysis of Diversified Portfolios: Classification, Comparison and Refinements

For more than three decades, empirical analysis of stochastic dominance was restricted to settings with mutually exclusive choice alternatives. In recent years, a number of methods for testing efficiency of diversified portfolios have emerged, which can be classified into three main categories: 1) majorization, 2) revealed preference and 3) distribution-based approaches. Unfortunately, some of these schools of thought are developing independently, with little interaction or cross-referencing among them. Moreover, the methods differ in terms of their objectives, the information content of the results and their computational complexity. As a result, the relative merits of alternative approaches are difficult to compare. This chapter presents the first systematic review of all three approaches in a unified methodological framework. We examine the main developments in this emerging literature, critically evaluating the advantages and disadvantages of the alternative approaches. We also point out some misleading arguments and propose corrections and improvements to some of the methods considered.

2.1 Introduction

For more than three decades, empirical analysis of stochastic dominance was restricted to settings with mutually exclusive choice alternatives, appropriate for comparison of income distributions or crop yields in agriculture, for example. These methods include various mean-risk models (see, for example, Hogan and Warren, 1972, Ang, 1975, Shalit and Yitzhaki, 1984) and direct pairwise efficient comparison of distribution functions (Hadar and Russell, 1969, Bawa et al., 1979, Aboudi and Thon, 1994, Anderson, 1996, Annaert et al., 2009, among many others). However, pairwise comparison algorithms are insufficient for identifying dominating portfolios from an infinite set of diversified portfolios, which is a typical setting in finance. Levy (1992) emphasizes this problem by stating:

> "Ironically, the main drawback of the SD framework is found in the area of finance where it is most intensively used, namely, in choosing the efficient diversification strategies. This is because as yet there is no way to find the SD efficient set of diversification strategies as prevailed by the M-V framework. Therefore, the next important contribution in this area will probably be in this direction".

Some authors introduced other SD-related concepts, such as convex SD (Fishburn, 1974) and marginal conditional SD (Shalit and Yitzhaki, 1994). Such methods can only provide a necessary condition for stochastic dominance efficiency when the portfolio possibilities set has a particular structure, but not in general.

In recent years, stochastic dominance literature has developed a number of methods for analyzing efficiency of diversified portfolios, following the works of Kuosmanen (2004, 2001-WP), Post (2003) and Dentcheva and Ruszczyński (2003). Although Dybvig and Ross (1982) propose SSD efficiency criteria that can be developed into an SSD efficiency test with diversification (such as in Lizyayev, 2009), they only provide a useful idea, but not an explicit algorithm.

The first authors to address stochastic dominance relative to an infinite set of choice alternatives after Dybvig and Ross (1982) were Ogryczak and

Ruszczyński (1999, 2001, 2002) in their mean-risk models. Ogryczak and Ruszczyński (1999) proposed an optimization problem that identified mean-risk efficient frontiers of stochastically non-dominated portfolios, and extended it to higher-order semideviations in Ogryczak and Ruszczyński (2001). Subsequently, Ruszczyński and Vanderbei (2003) have explicitly formulated the frontier identification problem for portfolio weights, and suggested an efficient parametric optimization. Although mean-risk models cannot generally solve the problem of identifying whether a given portfolio is SD efficient (which is the formulation usually employed in asset pricing and investment management), they can be used as a necessary condition for SSD efficiency.

To our knowledge, Dentcheva and Ruszczyński (2003) and Kuosmanen (2004) independently developed the first algorithms to identify a portfolio that dominates a given benchmark among an infinite number of diversified portfolios by solving a finite dimensional optimization problem. A preliminary version of Kuosmanen's test appeared in Kuosmanen (2001-WP) working paper. Meanwhile, Post (2003) developed an alternative SSD efficiency test which is simpler and computationally less demanding, but does not generally produce a dominating portfolio. Dentcheva and Ruszczyński (2003) introduced an optimization model with stochastic dominance constraints and developed this model further in Dentcheva and Ruszczyński (2006b) and Rudolf and Ruszczyński (2008). Although this model has an arbitrary objective function and in this respect is more general, we will focus on its use in the most frequently applied setting in finance, namely: identifying the SD efficiency of a given portfolio relative to a diversified portfolio possibilities set. Dentcheva and Ruszczyński (2006a) introduced inverse stochastic dominance constraints, which were later employed in Kopa and Chovanec (2008) refined method for testing stochastic dominance efficiency.

The literature of stochastic dominance currently spans a number of alternative methods. To structure this literature, we propose to classify the present approaches into three categories: 1) majorization, 2) revealed preference and 3) distribution-based approaches. These approaches differ in their objectives, the information content of the results, and their computational complexity. Unfortunately, some of these schools of thought are developing independently, with little interaction or cross-reference to the other schools and as a result the advantages and disadvantages of alternative approaches have not been compared in a fair and systematic fashion. The proponents of each method have a natural tendency to exaggerate the advantages of their favorite method and overlook the advantages of their competitors'.

This Chapter presents the first systematic attempt to bring all three approaches under the common umbrella of a unified methodological framework. We will examine the main developments in this emerging literature, critically evaluating the advantages and disadvantages of the alternative approaches using a number of objective criteria. We will also point out some misleading arguments in this literature and propose corrections and improvements to some of the methods considered.

The Chapter is organized as follows. In Section 2.2 we define the basic general concepts related to stochastic dominance efficiency and state some common assumptions. Since most of the methods are applied to the second order stochastic dominance (SSD), where the efficiency test becomes a linear program[1], we classify, analyze and compare the most important SSD efficiency algorithms to date in Section 2.3. To keep such comparative analysis objective, we use a unified framework of Section 2.2 and adjust each of the methods considered in such a way that they solve the same standardized problem which is commonly and frequently used in practice. In Section 2.4 we consider some extensions to the standardized framework such as first order stochastic dominance (FSD) and unbounded short sales, and analyze the extent to which the existing methods can tackle those modifications. Finally, Section 2.6 gives some concluding remarks and finalizes the Chapter.

[1]Some authors use more demanding non-linear programs (such as Linton et al. (2005, 2010) and the iterative quadratic program of Post and Versijp (2007)) which, in addition to the efficiency outcome, also provide statistical significance scores under some assumptions. Since such programs do not produce a dominated portfolio and are considerably more computationally demanding, we will omit them from our analysis. As statistical significance scores can be more naturally obtained via non-parametric bootstrapping procedures in the framework of this Chapter, we will focus on SSD efficiency tests which are more practical in terms of the computational complexity and the information content of the result.

2.2 Unified framework

As a first step towards bringing alternative approaches under a common umbrella, we need a general framework into which all alternative methods can naturally fit. It is the purpose of this section to describe such a framework. We should note that some of the methods reviewed in the subsequent sections do not necessarily require all of the assumptions imposed in this section. In the interest of clarity, however, we will review all methods from the perspective of the unified framework, duly noting the possible extensions as we proceed.

A canonical model of investment decision making in a static setting can be described as follows. There are n marketed assets, whose returns may vary across different states of nature. From m possible states, one state is randomly drawn as the realized state. Returns of assets in m alternative states of nature are described by m-by-n matrix X. If a riskless asset is available in the market, we can include it as one column of X (a column with equal components). Naturally, all asset returns are assumed to be linearly independent, which implies that $X^{\mathsf{T}}X$ is positive definite. Note that there is no uncertainty about the return matrix X; the investors' risk arises from the random realization of one out of m possible states. Without loss of generality, we assume all states to be equally likely.[2]

Investors may diversify between available assets. We shall use $\lambda \in \mathbb{R}^n$ for a vector of portfolio weights. The portfolio possibilities set (assuming away short sales) is

$$\Lambda = \left\{\lambda \in \mathbb{R}^n : \lambda^{\mathsf{T}} e = 1, \lambda \geq 0\right\}^3,$$

and the set of all available allocations is

$$M_X = \{x \in \mathbb{R}^m : x = X\lambda, \lambda \in \Lambda\}.$$

[2]States with different probabilities can be dealt with by a linear transformation of decision variables so that the resulting program will be equivalent to the one with equally probable states; see Dybvig and Ross (1982) for details.

[3]Unless otherwise stated, we will consider the PPS with short sales restricted. Nonetheless some other restrictions on portfolio possibilities may apply in practice; moreover the use of some methods can be particularly advantageous for certain classes of PPS, as will be shown in subsequent sections.

Each investor has a von Neuman-Morgenstern utility function $u \in U = \{u : \mathbb{R} \to \mathbb{R}\}$ which depends on his final wealth at the end of the holding period. As shown in Pratt (1964), investors' non-satiation and risk attitude can be modeled via the first and second derivative of u, respectively. The class of increasing utility functions which represents all non-satiable investors is denoted by U_1, and the class of increasing and concave utility functions is denoted by U_2 and represents all non-satiable and risk-averse investors. Formally,

$$U_1 \equiv \{u : \mathbb{R} \to \mathbb{R} \quad \text{s.t. } u'(t) \geq 0,\ \forall t\}$$

and

$$U_2 \equiv \{u : \mathbb{R} \to \mathbb{R} \quad \text{s.t. } u'(t) \geq 0, \text{ and } u''(t) \leq 0,\ \forall t\}\,.$$

Due to the uncertainty about which state of the world will occur, investors seek to maximize their expected utility. Portfolio $\tau \in \Lambda$ is the optimal choice for an investor with utility $u \in U$ if and only if

$$\mathbb{E}u(X\tau) = \sup_{\lambda \in \Lambda} \mathbb{E}u(X\lambda), \tag{2.1}$$

where $\mathbb{E}$ denotes the expected value operator. Since all states are equally likely by assumption, equation (2.1) can be equivalently stated as

$$\sum_{i=1}^{m} u(x^i\tau) = \sup_{\lambda \in \Lambda} \sum_{i=1}^{m} u(x^i\lambda). \tag{2.2}$$

Observing a given portfolio τ, our purpose is to evaluate whether τ is the optimal choice for a group of investors. Since the investors' utility functions are unknown, we focus on broad classes of economically meaningful utility functions, U_1 and U_2. To this end, the following definitions prove useful.

Definition 2.1 (dominance)**.** *Portfolio $\lambda \in \Lambda$ dominates portfolio $\tau \in \Lambda$ by First Order Stochastic Dominance, further **FSD** (by Second Order Stochastic Dominance, further **SSD**) if and only if for all utility functions $u \in U_1 (u \in U_2)$*

$$\sum_{i=1}^{m} u(x^i\lambda) \geq \sum_{i=1}^{m} u(x^i\tau), \tag{2.3}$$

with a strict inequality for at least one $u \in U_1 (u \in U_2)$.

Definition 2.2 (super-dominance)**.** *Portfolio $\lambda \in \Lambda$ super-dominates portfolio $\tau \in \Lambda$ by FSD (SSD) if and only if for all strictly increasing utility functions $u \in U_1 (u \in U_2)$*

$$\sum_{i=1}^{m} u(x^i \lambda) > \sum_{i=1}^{m} u(x^i \tau), \tag{2.4}$$

Definition 2.1 is standard in the stochastic dominance literature. The notion of super-dominance is a new term that we have coined for the definition first proposed by Post (2003). Note that super-dominance implies dominance, but the reverse is not true. For example, if τ is a mean-preserving spread of portfolio λ, then τ dominates λ by SSD, but it does not super-dominate it.

Definitions 2.1 and 2.2 can be stated analogously for any given class of utility functions U. Although U_1 and U_2 are the most frequently used, some authors developed tests for refined utility classes, e.g. modeling increasing relative and decreasing absolute risk aversion, such as Vickson (1975, 1977) and Lizyayev (2009).

Using Definitions 2.1 and 2.2, the notions of portfolio efficiency and optimality are defined as follows:

Definition 2.3 (weak efficiency)**.** *Portfolio $\tau \in \Lambda$ is weakly FSD (SSD) efficient if and only if there does not exist another portfolio $\lambda \in \Lambda$ that super-dominates τ in the sense of Definition 2.2.*

Definition 2.4 (strong efficiency)**.** *Portfolio $\tau \in \Lambda$ is strongly FSD (SSD) efficient if and only if there does not exist another portfolio $\lambda \in \Lambda$ that dominates τ in the sense of Definition 2.1.*

Definition 2.5 (optimality)**.** *Portfolio $\tau \in \Lambda$ is FSD (SSD) optimal if and only if there exists a strictly increasing $u \in U_1 (u \in U_2)$ for which τ is the optimal portfolio choice, that is,*

$$\sum_{i=1}^{m} u(x^i \tau) > \sum_{i=1}^{m} u(x^i \lambda), \text{ for all } \lambda \in \Lambda \backslash \{\tau\}.$$

There exist alternative equivalent definitions of stochastic dominance which we state below.

Definition 2.6. *Allocation $x \in M_X$ with cumulative distribution function (CDF) $F_X(z)$ dominates allocation $y \in M_X$ having CDF $F_Y(z)$ by FSD (SSD) if and only if*

$$F_X(z) \leq F_Y(z) \left(F_X^{(2)}(z) \leq F_Y^{(2)}(z) \right), \tag{2.5}$$

for all z, with a strict inequality for at least one z,
where $F_X^{(2)}(z)$ is defined as $F_X^{(2)}(z) \equiv \int\limits_{-\infty}^{z} F_X(t)dt = \mathbb{E}\left(\max\{z - X, 0\}\right).$

Due to the latter representation $F_X^{(2)}(z)$ is also called *the expected shortfall* of X. Similarly, SD relation can be equivalently formulated in terms of (integrated) inverted CDF (quantiles) as follows. Condition (2.5) is equivalent to

$$F_X^{-1}(q) \geq F_Y^{-1}(q) \left(F_X^{-2}(q) \equiv \int\limits_0^q F_X^{-1}(v)dv \geq \int\limits_0^q F_Y^{-1}(v)dv \equiv F_X^{-2}(q) \right), \tag{2.6}$$

for all $q \in [0, 1]$.

SSD condition (2.6) can also be expressed in terms of *conditional value at risk* (CVaR) which is related to $F_X^{-2}(q)$ (see Rockafellar and Uryasev, 2002) as

$$F_X^{-2}(q) = -q\mathrm{CVaR}_{1-q}(-X), q \in (0, 1).$$

Definition 2.7. *Allocation $x \in M_X$ dominates allocation $y \in M_X$ by FSD (SSD) if and only if*

$$\exists P \in \Pi \,(\exists W \in \Xi) : x \geq Py(x \geq Wy),$$

where Π is the class of permutation matrices:

$$\Pi = \left\{ [w_{ij}]_{m \times m} : w_{ij} \in \{0, 1\}, \sum_{i=1}^{m} w_{ij} = \sum_{j=1}^{m} w_{ij} = 1, i, j = 1, \ldots, m \right\}$$

and Ξ *is the class of doubly stochastic matrices:*

$$\Xi = \left\{ [w_{ij}]_{m\times m} : 0 \leq w_{ij} \leq 1, \sum_{i=1}^{m} w_{ij} = \sum_{j=1}^{m} w_{ij} = 1, i, j = 1, \ldots, m \right\}.$$

Definitions 2.1, 2.6 and 2.7 are known to be equivalent. The equivalence of definitions 2.1 and 2.6 is easy to prove by changing variables in the integration of Definition 2.6. For the equivalence of Definition 2.7 see Hardy et al. (1934), Hadar and Russell (1969) and Marshall and Olkin (1979).

For the sake of brevity we will sometimes refer (with a slight abuse of notation) to an allocation by the corresponding portfolio, for instance by stating that portfolio $\tau \in \Lambda$ dominates allocation $y \in M_X$ we mean that $X\tau$ dominates y.

The difference between the FSD and SSD efficiency arises from the assumption of risk aversion: SSD assumes risk aversion, whereas FSD does not. In the case of SSD, the optimality and efficiency definitions (2.2) and (2.3) are equivalent if the portfolio possibilities set Λ is convex. However, FSD optimality is only a necessary condition for FSD efficiency, even with a convex Λ.

Restrictions on the set of utility functions strongly affect the computational complexity of a test, as will be demonstrated below. The computational burden becomes particularly restrictive when it comes to bootstrapping and statistical inference. To assess whether the outcome of a test is statistically significant (and cannot be attributed solely to chance), one needs to simulate a large number of new data sets of asset returns generated by the same distribution as the original, and further to run the same efficiency test on all those samples. With current computing power and the usual dimensionality of the data, only certain types of optimization programs can be tackled within a reasonable time, such as linear or quadratic programs. Mixed integer linear programs (which FSD efficiency tests are in essence) are far too demanding for any rigorous bootstrapping techniques. For that reason, and because the vast majority of the tests used in practice are focused on second order stochastic dominance efficiency, we will analyze them in detail below.

2.3 Second Order (SSD) Efficiency

The extensive literature which suggests SSD efficiency algorithms can be grouped into three main categories: 1) majorization, 2) revealed preference, and 3) distribution-based approaches. The first category is based on optimality conditions in the space of returns given in Definition 2.7; the second on Lagrangean conditions for the marginal utility rationalizing a given portfolio in accordance with Definition 2.5, and the last on various equivalent criteria of SD efficiency formulated directly on cumulative distribution functions of underlying portfolios as in Definition 2.6. Although the categories above are not mutually exclusive (e.g. the dual formulation to distribution-based approach has a revealed preference interpretation), most of the methods are most frequently used either in their primal or dual form, which we will take as the basis for our classification. In this chapter we attempt to cover the most efficient methods of each school of thought.

To characterize and compare all the methods in a fair and systematic fashion we would like to point out the criteria an SSD efficiency test should fulfill. Clearly, the primary goal of every method should be to identify whether a given portfolio is efficient relative to a given convex portfolio possibilities set in the sense of Definition 2.4.[4] The methods therefore should provide necessary and sufficient conditions for such efficiency. In cases when the subject portfolio is inefficient, one would like to have a measure indicating the degree of its inefficiency. A natural choice for such a measure could be the highest possible difference in mean returns between the subject portfolio and an efficient marketed portfolio that dominates it. If there is a dominating portfolio with the same mean return as the subject portfolio but with a tighter spread around the risk-free asset (this dominance is self-evident and formally in accordance with Definition 2.4), one would like to incorporate the maximal feasible spread into the measure of inefficiency as well. For that reason, it is desirable that when a given portfolio is inefficient, an efficiency test identifies a dominating portfolio that is marketed and SSD efficient itself. Another advantage would be if the method could be split into some sequential sub-

[4]Although occasionally we will distinguish the weak efficiency in the sense of Definition 2.3, we adapt the commonly accepted SSD efficiency given by Definition 2.4 throughout, and unless otherwise stated, SSD efficiency will refer to this strong definition.

tests that are less computationally demanding, so that one could identify inefficiency at an earlier stage based on some necessary conditions, in which case running the rest of the test would be unnecessary. Finally, the ability of SSD tests to be easily generalized to FSD efficiency testing would also be of value.

2.3.1 Revealed Preference Approach

The revealed preference approach has its roots in Afriat (1967) celebrated theorem. Analogous to Afriat's test of rational consumer behavior[5], SSD efficiency can be tested based on the first order optimality conditions for the utility function which, provided that such function exists, would rationalize the subject portfolio, in accordance with Definition 2.5 and the fact that SSD optimality is equivalent to SSD efficiency if the portfolio possibilities set is convex. The general idea of the revealed preference approach is to try to find marginal utilities β for some well-behaved von Neumann-Morgenstern utility function for which the evaluated portfolio $y \in M_X$ is the optimal solution maximizing its expected value. If such marginal utilities β exist, then the evaluated portfolio is literally "revealed optimal", at least for some hypothetical decision maker with rational preferences. If such marginal utilities do not exist, then the evaluated portfolio y is SSD inefficient.

While the marginal conditional stochastic dominance introduced in Shalit and Yitzhaki (1994) can, like some other earlier methods, formally be assigned to this category, it uses different settings in which the subject portfolio is tested relative to a set of vertices of a portfolio possibilities set. This test is computationally less demanding but can only be used as a first-stage necessary pre-processing test for our framework, as it can not generally identify SSD efficiency in the case of full diversification. Marginal conditional formulation also appears as duality results in the distribution-based methods, such as Dentcheva and Ruszczyński (2006a,b), Rudolf and Ruszczyński (2008). However, the primal distribution-based method appears to be computationally competitive relative to its dual linear programming formulations.

[5]Varian (1983) has applied Afriat's approach to testing rationality of investor behavior in a somewhat different setting than the one considered in this chapter.

Therefore we will classify these tests as distribution-based and will cover them below in a separate sub-section.

Post (2003) formulates the following revealed preference test for SSD efficiency of a given marketed portfolio $y \in M_X$:

$$\begin{aligned} \xi(y) &= \min \theta \\ \text{s.t. } \frac{1}{m}\sum_{t=1}^{m} \beta_t (y_t - X_{ti}) + \theta &\geq 0, \quad i = 1, \dots, n \\ \beta_1 \geq \beta_2 \geq \cdots \geq \beta_m &= 1 \\ \lambda &\in \Lambda \\ \theta &\text{ free} \end{aligned} \tag{2.7}$$

Parameters β_t can be interpreted as Afriat numbers, which represent the marginal von Neumann-Morgenstern utility of some rational decision maker in state t. If the optimal solution to (2.7) satisfies $\xi^*(y) = 0$, then the evaluated portfolio is an optimal solution that maximizes expected utility to some rational risk-averse decision maker. Thus $\xi^*(y) = 0$ is a necessary and sufficient condition for weak SSD efficiency of y (given in Definition 2.3) and, as Kuosmanen (2004) notes, only a necessary condition for the strong SSD efficiency (in the sense of Definition 2.4).

Kuosmanen (2004, Sec. 4.4) derives a similar test based on the idea of separating hyperplanes. Both methods are only capable of determining the efficiency status of a given portfolio; they do not generally produce a dominating portfolio. The major advantage of the methods is their computational simplicity: (2.7) is a linear program with $m+1$ variable and $n+m$ constraints.

Post (2003) also derives a dual formulation to (2.7) as follows.

$$\begin{aligned} \psi(y) &= \max s_m \\ \text{s.t. } \frac{1}{m}\sum_{i=1}^{k} \left(x^i \lambda - y_i\right) &= s_k, \quad k = 1, \dots, m \\ \lambda &\in \Lambda \\ s &\in \mathbb{R}^m_+ \end{aligned} \tag{2.8}$$

If the optimal solution to (2.8) is $\psi^* = 0$, then y is weakly SSD efficient. Although the optimal portfolio λ^* has an intuitive interpretation as the portfolio with the largest increase in the mean return, it does not necessarily dominate y. To see this, consider (2.8) for $y = [1, 4], x^1 = [9, 0], x^2 = [0, 2], x^3 = y$. Running the tests yields $\xi^*(y) = \psi^*(y) = 2$ which correctly identifies SSD inefficiency of y, however $X\lambda^* = x^1 = [9, 0]$, even though x^1 does not dominate y.

Post (2008) has extended the SSD test for weak efficiency to the standard case of strong efficiency (Definition 2.4) by simply changing the objective function of (2.8) from s_m to the sum $s^\mathsf{T} e$, obtained $s^\mathsf{T} e = 0$ as the necessary and sufficient condition for the strong SSD efficiency and shown that the subject portfolio y is always SSDdominated by a linear combination of $X\lambda^*$ and y. However, a dominating portfolio obtained thus does not necessarily have the highest mean return among all dominating portfolios, and therefore is not suitable as a benchmark for efficiency gauging. Further, the dominating portfolio is not necessarily SSD efficient even in the sense of weak SSD efficiency (Definition 2.3).

2.3.2 Majorization Approach

The majorization approach is based on Definition 2.7, which originates in the mathematical literature on stochastic dominance, where the concept appeared as stochastic ordering. The first majorization-based test in economic literature appeared in Kuosmanen (2001-WP) and was further developed in Kuosmanen (2004).

Kuosmanen (2004) splits SSD efficiency test into necessary and sufficient subtests. The necessary test reads[6]

[6] Kuosmanen (2004) formulates (2.9) with X augmented by y, as it can happen that $y \notin M_X$ is SSD efficient, but is dominated by a linear combination of a marketed portfolio and itself. We omit this augmentation here for the sake of comparability with the other methods.

$$\begin{aligned}
\theta_2^N(y) = \max_{\lambda, W} & (X\lambda - y)^\mathsf{T} e \\
\text{s.t. } & X\lambda \geq Wy \\
& W \in \Xi \\
& \lambda \in \Lambda
\end{aligned} \tag{2.9}$$

Comparing (2.9) with Post's (2008) dual (2.8) reveals that the two problems are structurally similar, except for the doubly stochastic matrix W included in (2.9). Post (2008) sorts the asset returns in ascending order with respect to y, whereas Kuosmanen did not utilize the prior ordering. As a result, the optimal portfolio λ^* of (2.9) always SSD dominates y when the latter is inefficient (provided W^* is not a permutation matrix), contrary to (2.8).

Kuosmanen (2004) shows that $\theta_2^N = 0$ is a necessary condition for the strong SSD efficiency of portfolio y. Note, that θ_2^N/m can be intuitively used as an inefficiency measure that indicates the difference between the mean return of the dominating portfolio λ^* with the highest mean return and the expected return of y. Another possibility considered by Kuosmanen (2004) is to gauge efficiency by using the minimum risk-free premium that needs to be added to y to make it SSD efficient. While such a measure can be intuitive for gauging inefficiency loss, it cannot provide a necessary SSD efficiency condition analogous to (2.9). The same is true for the more general directional distance function formulated in Kuosmanen (2007).

Kuosmanen (2007) derives the dual formulation to (2.9), which can be expressed as

$$\begin{aligned}
\xi^D(y) = \min_{\beta, \theta, a, b} & \theta - (a^\mathsf{T} e + b^\mathsf{T} e) \\
\text{s.t. } & \theta e \geq X^\mathsf{T} \beta \\
& \beta_s y_t \geq y_t + a_t + b_s, \quad \forall s, t = 1, \ldots, m \\
& \beta \geq e \\
& \theta \in \mathbb{R}, \quad a, b, \beta \in \mathbb{R}^m
\end{aligned} \tag{2.10}$$

Clearly, the dual program (2.10) is similar to (2.9) in terms of the computational complexity. However, (2.10) is less intuitive and its coefficients are difficult to interpret. Moreover, it is unclear if (2.10) can be generalized to a sufficient test for SSD efficiency of y in a straightforward way. For that reason we shall focus on the primal formulation (2.9) for which Kuosmanen (2004) proposed the following sufficient test statistic.

$$\begin{aligned}
\theta_2^S(y) = \min & \sum_{i=1}^{m}\sum_{j=1}^{m}\left(s_{ij}^{+}+s_{ij}^{-}\right) \\
\text{s.t. } & X\lambda = Wy \\
& s_{ij}^{+}+s_{ij}^{-} = w_{ij}-\frac{1}{2}, \quad i,j=1,\dots,m \\
& s_{ij}^{+}+s_{ij}^{-} \geq 0, \quad i,j=1,\dots,m \\
& W\in\Xi, \lambda\in\Lambda
\end{aligned} \tag{2.11}$$

Program (2.11) minimizes $\sum_{i=1}^{m}\sum_{j=1}^{m}\left|w_{ij}-\frac{1}{2}\right|$. The underlying idea lies in finding a marketed mean-preserving spread of y that is as close to the risk free ray as possible. The non-existence of any such $X\lambda^* \neq y$ would then suffice for SSD efficiency of y. Kuosmanen (2004) proposes the theoretical maximum of the test statistic as a sufficient condition[7]:

$$\theta_2^S(y) = \frac{m^2}{2} - \sum_{k=2}^{m} k d_{0k}, \tag{2.12}$$

where d_{0k} is the number of k-way ties.

Although the optimal $X\lambda^*$ from (2.11) always SSD dominates y (provided the two portfolios are distinct) it may not be SSD efficient. To see this, consider the following example. Suppose we test portfolio $y=[2,0,10]$ and both $x^1=[6,5,1]$ and $x^2=[4,4,4]$ are marketed. If $\theta_2^S(y)=3/2$, the program (2.11) may have chosen x^1 with $W_1^*=\frac{1}{2}\begin{bmatrix}1&0&1\\0&1&1\\1&1&0\end{bmatrix}$, however x^2 may have been chosen as well with $W_2^*=\frac{1}{3}\begin{bmatrix}1&1&1\\1&1&1\\1&1&1\end{bmatrix}$, since W_1^* and W_2^* give the

[7]Kuosmanen (2004) defines $\theta_2^S(y)$ as $\frac{m^2}{2}-\sum_{k=1}^{m} kd_{0k}$; however he clearly meant (2.12). Moreover, the summation $\sum_{k=1}^{m} kd_{0k}$ equals m, since it counts all m elements of y precisely once.

same value of statistic $\theta_2^S(y)$. Therefore, if (2.11) picks x^1, it dominates y but is not SSD efficient.

If the efficiency of the dominating portfolio is required, one can use the following quadratic programming extension of (2.11).

$$\begin{aligned} \theta_2^R(y) = \min\ & \lambda^\mathsf{T} X^\mathsf{T} X\lambda - y^\mathsf{T} y \\ \text{s.t. } & X\lambda = Wy \\ & W \in \Xi \\ & \lambda \in \Lambda \end{aligned} \tag{2.13}$$

Note that (2.13) minimizes the second moment of $X\lambda$ which is equivalent to minimizing the Euclidian distance from y to the risk free asset $e \cdot \mathbb{E}y$. We can prove the following

Proposition 2.1. *Suppose $\theta_2^N(y) = 0$. Portfolio y is SSD efficient with respect to Λ if and only if $\theta_2^R(y) = 0$. Moreover, λ^* from (2.13) is SSD efficient and, if $\theta_2^R(y) \neq 0$, dominates y.*

Proof. It follows from the majorization theory (see Marshall and Olkin, 1979) that if for some $W \in \Xi$, Wy is not a permutation of y, then $y^\mathsf{T} W^\mathsf{T} Wy < y^\mathsf{T} y$. Given that $X = Wy$, the objective $\lambda^\mathsf{T} X^\mathsf{T} X\lambda - y^\mathsf{T} y = y^\mathsf{T} W^\mathsf{T} Wy - y^\mathsf{T} y = y^\mathsf{T}(W^\mathsf{T} W - I_m)y \leq 0$. Therefore, $\theta_2^R(y) = 0$ implies $W^*y = Py$, for some permutation matrix $P \in \Pi$, and thus y is efficient. Similarly, if $\theta_2^R(y) < 0$, then y is dominated by W^*y, so y is SSD inefficient. The efficiency of $X\lambda^*$ follows from the fact that the existence of a strictly dominating portfolio $X\tau = WX\lambda^*$ would contradict the optimality of $X\lambda^*$ in (2.13). □

Summarizing, we can characterize the method as follows. The necessity test (2.9) is a linear program with m^2+n variables, m^2+m inequality and $2m$ equality constraints. Program (2.11) with $3m^2+n$ variables, m^2+3m equality and $3m^2$ inequality constraints is a sufficient test for SSD efficiency of y, but the optimal portfolio itself may not be SSD efficient. An alternative sufficient condition is given by Proposition 2.1 that does generate an SSD efficient dominating portfolio W^*y as a byproduct. This test is based on quadratic program (2.13) with $m^2 + n$ variables (of which n enter the objective), $3m$ linear equality and m^2 linear inequality constraints.

Based on the general theoretical result in Strassen (1965), Luedtke (2008) recently developed the majorization test (2.9) further by explicitly including the probabilities of the states (which are assumed equal in (2.9)) and suggested a branching heuristic for solving the method. His linear programming formulation, however, closely resembles (2.9), particularly in terms of the computational complexity.

2.3.3 Distribution-Based Approach

This group of methods is based on Definition 2.6 and usually employs equivalent definitions involving various modifications of the cumulative distribution function and its inverse, such as integrated (inverted) CDF, quantiles and conditional value at risk.

Dentcheva and Ruszczyński (2003) introduced the following linear program with distribution-based stochastic dominance constraints.

$$
\begin{aligned}
& \max f(\lambda) = \mathbb{E}(X\lambda) \\
\text{s.t.} \quad & \sum_{k=1}^{n} x_{ik}\lambda_k + s_{ij} \geq y_j, \quad i,j = 1,\dots,m \\
& \frac{1}{m}\sum_{i=1}^{m} s_{ij} \leq v_j, \quad j = 1,\dots,m \\
& s_{ij} \geq 0, \quad i,j = 1,\dots,m \\
& \lambda \in \Lambda
\end{aligned}
\tag{2.14}
$$

where $v_j \equiv \mathbb{E}[(y_j - y)_+] = F_Y^2(y_j)$ is the expected shortfall of y.

The constraints in (2.14) basically ensure that $\mathbb{E}[(a - X\lambda)_+] \leq \mathbb{E}[(a-y)_+], \forall a$, which by Definition 2.6 is equivalent to the SSD dominance of $X\lambda$ over y, see Dentcheva and Ruszczyński (2003, 2006b) for more details.

Rudolf and Ruszczyński (2008) elaborated on this method, suggesting two alternative implementations of (2.14): a primal cutting plane method and a dual column generation method. However, they concluded that the dual method proved to be practically prohibitive for this problem (compared to a straightforward simplex implementation of (2.14)). The primal method was shown to outperform the simplex on their data set. However it is not clear if such performance can be generalized on an arbitrary data set; the method may require a factorial number of iterations in the worst case scenario.

Just like Kuosmanen's (2004) test (2.9), program (2.14) always produces a weakly SSD efficient dominating portfolio λ^* which may not be (strongly) SSD efficient (this may happen when (2.14) has multiple solutions). To overcome this, consider the following sufficiency test statistic.

$$
\begin{aligned}
\theta^R(y) = \sum_{i=1}^{m}(y_i - \mathbb{E}(y))^2 &- \min \sum_{i=1}^{m}\left(\sum_{j=1}^{n} x_{ij}\lambda_j - \mathbb{E}(y)\right)^2 \\
\text{s.t.} \ \sum_{k=1}^{n} x_{ik}\lambda_k + s_{ij} &\geq y_j, \quad i,j = 1,\dots,m \\
\sum_{i=1}^{m} s_{ij} &\leq m v_j, \quad j = 1,\dots,m \\
s_{ij} &\geq 0, \quad i,j = 1,\dots,m \\
\lambda &\in \Lambda
\end{aligned}
\tag{2.15}
$$

Proposition 2.2. *Let $x^* = X\lambda^*$ be a solution of (2.14) for a given portfolio $y \in M_X$. Determine $\theta^R(x^*)$ by solving (2.15) and denote the optimal solution by z^*. Portfolio y is SSD inefficient if and only if*

$$
\mathbb{E}(x^*) - \mathbb{E}(y) + \theta^R(x^*) > 0 \tag{2.16}
$$

Moreover, (2.16) also implies that y is dominated by z^* which is SSD efficient.

Proof. First note, that solution z^* to (2.15) is unique, due to the strict convexity of the objective function in (2.15) and linear independence of returns. Due to the dominance restrictions imposed in (2.15), z^* is SSD efficient. Since (2.16) holds if and only if z^* and y are distinct, the result follows. □

Program (2.14) is closely related to Kuosmanen's necessary test (2.9) in terms of the information content of the result. Both methods can identify a necessary and sufficient condition for the weak SSD efficiency (Definition 2.3), but only a necessary condition for the standard SSD efficiency. The optimal reference portfolio $X\lambda^*$ dominates y and is itself weakly SSD efficient. If several dominating portfolios of equal mean are available, both methods may select a dominating portfolio that is not (strongly) SSD efficient. Moreover, the two methods are following the same principle: to maximize the mean return among all available portfolios that dominate y and hence both can be used for inefficiency gauging. The only difference is that Kuosmanen (2004) exploits a majorization-based, and Dentcheva and Ruszczyński distribution-based dominance criteria. Test (2.14) is a linear program with $m^2 + n$ variables and $2m^2 + m$ constraints which is computationally heavier than Kuosmanen's necessary test (2.9), but lighter than his sufficiency test (2.11). Combined with (2.14), test (2.15) produces an SSD efficient dominating portfolio when the subject portfolio is inefficient.

Another distribution-based test recently published in Kopa and Chovanec (2008) employs the conditional value at risk defined as

$$\mathrm{CVaR}_\alpha(z) = \mathbb{E}(z|z > \mathrm{VaR}_\alpha(z)), \tag{2.17}$$

where $\mathrm{VaR}_\alpha(z)$ is the value-at-risk of z, that is $F_Z^{-1}(\alpha)$.

The following equivalent SSD efficiency criterion holds due to Definition 2.6:

$$\mathrm{CVaR}_\alpha(-Y_1) \leq \mathrm{CVaR}_\alpha(-Y_2), \forall\alpha \in [0,1] \Longleftrightarrow Y_1 \text{ SSD dominates } Y_2. \tag{2.18}$$

Employing an equivalent formulation of CVaR derived in Rockafellar and Uryasev (2002)

$$\mathrm{CVaR}_\alpha(Y) = \min_{a\in\mathbb{R}} \left\{ a + \frac{1}{1-\alpha}\mathbb{E}\max(Y - a, 0) \right\}, \tag{2.19}$$

they propose the following linear programming test[8].

[8] The inverse SD constraints, including those based on CVaR and used in Kopa and Chovanec (2008), were developed earlier in Dentcheva and Ruszczyński (2006a). However, the linear programming test (2.20) was suggested in Kopa and Chovanec (2008).

$$
\begin{aligned}
D^*(y) = \max \sum_{k=1}^{m} D_k \qquad\qquad & & (2.20)\\
\text{s.t. } \mathrm{CVaR}_{\frac{k-1}{m}}(-y) - b_k - \frac{\sum\limits_{t=1}^{m} w_k^t}{m-k+1} \geq D_k, & \qquad k = 1, \ldots, m \\
w_k^t \geq -(X\lambda)_t - b_k, & \qquad t, k = 1, \ldots, m \\
w_k^t \geq 0, & \qquad t, k = 1, \ldots, m \\
D_k \geq 0, & \qquad k = 1, \ldots, m \\
\lambda \in \Lambda &
\end{aligned}
$$

If $D^*(y) > 0$, then y is SSD inefficient, the optimal allocation $X\lambda^*$ dominates y and $X\lambda^*$ is SSD efficient. Otherwise $D^*(y) = 0$ and y is SSD efficient.

Substituting the explicit expression for CVaR into (2.20) gives us the following formulation (we denote $y^{[k]}$ the k-th largest, and $y^{(k)}$ the k-th smallest element of y).

$$
\begin{aligned}
D^*(y) = \max \sum_{k=1}^{m} D_k \qquad\qquad & & (2.21)\\
\text{s.t. } \frac{-\sum_{i=k}^{m} y^{[i]}}{m-k+1} - b_k - \frac{\sum_{t=1}^{m} w_k^t}{m-k+1} \geq D_k, & \qquad k = 1, \ldots, m \\
w_k^t \geq -(X\lambda)_t - b_k, & \qquad t, k = 1, \ldots, m \\
w_k^t \geq 0, & \qquad t, k = 1, \ldots, m \\
D_k \geq 0, & \qquad k = 1, \ldots, m \\
\lambda \in \Lambda &
\end{aligned}
$$

Therefore, the first constraint ensures that the optimal solution $x = X\lambda^*$ satisfies

$$
\frac{-\sum_{i=k}^{m} y^{[i]}}{m-k+1} + \frac{\sum_{i=k}^{m} x^{[i]}}{m-k+1} \geq 0, \qquad k = 1, \ldots, m \qquad (2.22)
$$

and therefore

$$\sum_{i=k}^{m} x^{[i]} \geq \sum_{i=k}^{m} y^{[i]}, \text{ hence } \sum_{i=1}^{k} x^{(i)} \geq \sum_{i=1}^{k} y^{(i)}, \qquad k = 1, \dots, m \tag{2.23}$$

which guarantees dominance of x over y by Definition 2.7.

Program (2.21) comprises both necessary and sufficient condition in one linear program. However, any necessity test can be used at a pre-processing stage to identify inefficiency prior to using (2.21), for instance Post's (2.7) or (2.8). In addition, Kopa and Chovanec (2008) propose another simple test formulated as follows.

$$d^* = \max_{\lambda \in \Lambda} \sum_{k=0}^{m-1} \sum_{j=1}^{n} \lambda_j \left(\mathrm{CVaR}_{\frac{k}{m}}(-y) - \mathrm{CVaR}_{\frac{k}{m}}(-X^j) \right) \tag{2.24}$$

$$\text{s.t.} \quad \sum_{j=1}^{n} \lambda_j \left(\mathrm{CVaR}_{\frac{k}{m}}(-y) - \mathrm{CVaR}_{\frac{k}{m}}(-X^j) \right) \geq 0, \quad k = 0, \dots, m-1$$

which can be rewritten as

$$d^* = \max_{\lambda \in \Lambda} \sum_{k=0}^{m-1} \sum_{j=1}^{n} \lambda_j a_{jk} \tag{2.25}$$

$$\text{s.t.} \quad \sum_{j=1}^{n} \lambda_j a_{jk} \geq 0, \quad k = 0, \dots, m-1, \text{ with } a_{jk} = \sum_{i=1}^{k} x_j^{(i)} - \sum_{i=1}^{k} y^{(i)}$$

Kopa and Chovanec prove that if $d^* > 0$, then y is SSD inefficient. Moreover, λ^* is an SSD efficient portfolio that dominates y. Note that all a_{jk} can easily be computed a priori, and thus (2.25) is a linear program with n variables and m constraints. In contrast to Post and Kuosmanen test, it provides an SSD efficient dominating portfolio in case of inefficiency of y. Unfortunately, no conclusion can be made concerning the efficiency of y if (2.25) is infeasible.

In contrast to Kuosmanen (2004) and Dentcheva and Ruszczyński (2003, 2006b), Kopa and Chovanec (2008) use the sum of slacks of CVaRs as the objective function which results in guaranteeing that the optimal portfolio is always SSD efficient and dominates the subject portfolio when the latter is inefficient, with a similar computational complexity.

In summary, (2.21) offers an attractive linear programming algorithm which comprises the necessary and sufficient condition for SSD efficiency of a given portfolio and provides for an SSD efficient dominating portfolio. The linear program (2.21) has $m^2 + 2m + n$ variables and $2m^2 + 2m$ inequality constraints.

2.4 Extensions

Below we consider some extensions to the set of assumptions set out in the previous chapter. We refine the class of preferences and assume away short sales, among others.

2.4.1 FSD Efficiency and Optimality

Due to the important role played by the ordering of portfolio returns in both strong and weak FSD dominance, there is no easy (polynomial complexity) algorithm known to date for identifying the efficiency of even a single given portfolio. Kuosmanen (2004) proposes an MILP-based test for identifying FSD efficiency, whereas Kopa and Post (2009) offer an LP test for the FSD optimality. However, the input data for the latter test can only be obtained by solving an MILP program similar to that of Kuosmanen (2004). Optimization programs with first order stochastic dominance constraints were also studied in Dentcheva and Ruszczyński (2004) and Noyan et al. (2006). Post (2003) suggests a seemingly easier LP test for FSD optimality in section V, formula (19). He states that the following condition implies and suffices for FSD optimality of portfolio y: $\theta^*(y) = 0$, where

$$
\begin{aligned}
\theta^*(y) &= \min \theta \\
\text{s.t. } \frac{1}{m}\sum_{t=1}^{m} \beta_t (X\lambda - X_{ti}) + \theta &\geq 0, \quad i = 1, \ldots, n \\
\beta_i &\geq 1, \quad i = 1, \ldots, m-1 \\
\beta_m &= 1, \quad \lambda \in \Lambda
\end{aligned}
\tag{2.26}
$$

Although temptingly simple, this approach turns out to be erroneous. This can be seen on the following example. Consider 3 assets in 2 states: A(2, 2), B(1, 3) and C(2.5, 1.75). The constraints of program (2.26) for testing FSD efficiency of B become

$$
\begin{cases}
\frac{\beta_1(1-2)}{2} + \frac{\beta_2(3-2)}{2} + \theta \geq 0 \\
\frac{\beta_1(1-2.5)}{2} + \frac{\beta_2(3-1.75)}{2} + \theta \geq 0 \\
\frac{\beta_1(1-1)}{2} + \frac{\beta_2(3-3)}{2} + \theta \geq 0 \\
\beta_1 \geq 1, \quad \beta_2 = 1
\end{cases}
\Rightarrow
\begin{cases}
\theta \geq 0.5\beta_1 - 0.5 \\
\theta \geq 0.75\beta_1 - 0.625 \\
\theta \geq 0 \\
\beta_1 \geq 1
\end{cases}
\Rightarrow \theta^* = 0.125 > 0.
$$

Thus, (2.26) wrongly classifies B as FSD inefficient.

2.4.2 Unrestricted shortsales

The tests reviewed above assume a convex portfolio possibilities set Λ. The simplest of these in terms of computational complexity, Post (2003), explicitly assumes restricted short sales[9]. The other methods can handle any polytope Λ, but are more computationally demanding. Lizyayev (2009) suggests another method particularly efficient in the case of unrestricted short sales, based on decomposition of the matrix of returns and applying gradient optimality conditions similar to those of Dybvig and Ross (1982). The

[9] In fact (2.7) is only valid for $\Lambda = \left\{\lambda \in \mathbb{R}^n : \quad \lambda^\mathsf{T} e = 1, \lambda \geq 0\right\}$. If Λ is another polytope, X_{ti} in (2.7) should be substituted by the vertices of Λ.

method seeks to find an interior point of the set

$$\left\{\beta \in \mathbb{R}^{m-n}: \quad D\begin{bmatrix}-(X_1^\mathsf{T})^{-1}X_2^\mathsf{T}\\ I_{m-n}\end{bmatrix}\beta \leq -D\begin{bmatrix}(X_1^\mathsf{T})^{-1}e\\ 0_{m-n}\end{bmatrix}\right\} \tag{2.27}$$

where β is an $(m-n)$-parameter vector, X_1 are the first n rows of X, X_2 - the rest $(m-n)$ rows[10], D is defined via the inverse of the upper triangular m-by-m matrix U_d^m: $D = -\left(U_d^m\right)^{-1}$.

Program (2.27) can be equivalently formulated in terms of the slacks of β as follows.

$$\begin{aligned} &\max \theta \\ \text{s.t.}\ & -\Big(\left(U_d^n\right)^{-1}\left(X_1^\mathsf{T}\right)^{-1}X_2^\mathsf{T}U_d^{m-n} + A\Big)\gamma + \theta \leq -D_1\left(X_1^\mathsf{T}\right)^{-1}e, \\ & \gamma \in \mathbb{R}^{m-n}, \quad \gamma \geq 0, \quad \theta \geq 0, \end{aligned} \tag{2.28}$$

where A is the following n-by-$(m-n)$ matrix: $A = \begin{bmatrix}0 & \cdots & 0\\ \vdots & & \vdots\\ 0 & \cdots & 0\\ 1 & \cdots & 1\end{bmatrix}$.

Existence of such γ (or $\beta = U_d^{m-n}\gamma$) is a necessary condition for SSD efficiency of the subject portfolio. System (2.28) is a linear program with $m-n$ variables and $m-n$ constraints (note that as in Post (2003), γ can be normalized so that $\gamma_{m-n} = \beta_{m-n} = 1$). Although the method does not find a dominating portfolio, it always identifies if the subject portfolio is efficient and, in such cases, produces a supporting gradient (marginal utility) as a byproduct. The computational advantage of this method becomes particularly eminent when n approaches m.[11]

[10] Assuming without loss of generality that the first n rows of X are linearly independent.

[11] Note however that as n increases, the dimensionality of (2.28) becomes smaller, but one needs to invert a larger X_1 prior to solving (2.28). If X happens to be particularly ill-conditioned, one may rewrite (2.28) without decomposition as: find $z \in \mathbb{R}^m$ such that $Xz = e, z \geq 0$. This is a linear program with m variables and $2m$ constraints, and therefore remains the most efficient method for the case of unbounded Λ.

2.5 Comparison of SSD methods

In Section 2.3 we have analyzed the major SSD efficiency tests that can be represented as linear programs and therefore can be applied to relatively large real world data. We have shown that the methods differ in terms of the information content of the results and the goal of this chapter is to analyze the computational burden associated with the extra informational outcome of some tests. To make such a comparison objective, we transform each program to the standard form:

$$\min\left\{c^{\mathsf{T}}x: \quad Ax \leq b, x \geq 0\right\}. \tag{2.29}$$

A good indicator of the computational complexity of a linear program, at least when a simplex method is applied for solving it, is the average number of non-zero elements in the matrix of constraints at each simplex iteration of changing variables in the basis. This indicator is however difficult to estimate on the basis of the input data. For this reason we follow the Performance World (2009) website and give two indicators for complexity: the size of the A matrix in (2.29) and the number of non-zeros it contains. For sparse matrices (as in our case) these indicators can be taken to represent, respectively, an upper and a lower bound for the average number of non-zeros in A. The table below summarizes the information content of the outcome and the computational complexity of the methods considered in terms of those indicators[12]. All the methods assume away short sales except Lizyayev (2009), which is only applicable to unrestricted short sales.

As we can see from Table 2.1, the methods can be grouped by the information content of the results. The most informative method, Kopa and Chovanec's (2.21), is also the most computationally demanding. It is the only method that identifies an SSD efficient dominating portfolio with a linear program. Other methods require quadratic programming to identify an SSD efficient dominating portfolio. The Dentcheva and Ruszczyński (2003)

[12] The portfolio budget constraint enters every method in the same form and thus was omitted from the complexity analysis for brevitys sake.

Table 2.1: Comparison of the computational complexity of the methods.
Columns: **Variables** $\times$ **Constraints**: number of elements in constraint matrix A, after the program has been transformed to the standard form (2.29); **Non-zeros**: number of non-zero elements in this matrix; **SC**: if the method gives a sufficient condition for SSD efficiency (in this and the next two columns '$-$' stands for NO, '$\checkmark$' for YES); **NC**: if the method provides a necessary condition for SSD efficiency; **DP**: if the method produces a dominating portfolio in case of inefficiency of the subject portfolio; **EDP** indicates if the dominating portfolio in DP is (strongly) SSD efficient (here '$-$' marks those methods that do not generally produce an SSD efficient dominating portfolio, '$\checkmark$' those that do, and '$\checkmark_w$' those that produce a dominating portfolio which is only weakly (in the sense of Definition 2.3), but not strongly, efficient; '$-/\checkmark$' in the necessary test of Kopa-Chovanec means that this test provides an efficient dominating portfolio only when (2.25) is feasible, but no conclusion can be made about the efficiency of y when (2.25) is infeasible).

	Variables $\times$ Constraints	**Non-Zeros**	**SC**	**NC**	**DP**	**EDP**
Kuosmanen (2004) Necessary	$5m^3 + 5mn$	$5m^2 + mn$	$-$	$\checkmark$	$\checkmark$	$\checkmark_w$
Kuosmanen (2004) Sufficient	$6m^4 + 18m^3 + 2m^2n + 6mn$	$12m^2 + 2mn$	$\checkmark$	$-$	$\checkmark$	$\checkmark_w$
Post (2003)	$mn + n$	$mn + n$	$-$	$\checkmark$	$-$	$-$
Post (2008)	$mn + n$	$mn + n$	$-$	$\checkmark$	$\checkmark$	$-$
Dentcheva-Ruszczyński (2003)	$m^4 + m^3 + m^2n + mn$	$m^2(n + 2)$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark_w$
Kopa-Chovanec (2008)	$m^4 + 4m^3 + 3m^2 + m^2n + mn$	$m^2(n + 3) + 2m$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$
Kopa-Chovanec (Necessary)	mn	mn	$-$	$-/\checkmark$	$-/\checkmark$	$-/\checkmark$
Lizyayev (2009) Unrestr. s/s	$mn - n^2 + n$	$mn - n^2 + n$	$-$	$\checkmark$	$-$	$-$

and Kuosmanen (2004) necessary tests are identical in terms of the information content of the results. As for the computational complexity, Kuosmanen (2004) is lighter in terms of non-zeros but slightly harder regarding the size of the constraints matrix. A dominating portfolio that is weakly SSD efficient is identified, but the dominating portfolio is not always strongly SSD efficient. The test of Post (2008) is a lot lighter than the previous tests, but it loses in information content. The Kopa and Chovanec (2008) necessary test (2.25) is slightly lighter than Post (2003) and in some cases identifies an efficient dominating portfolio, although no conclusions can be made concerning the efficiency of y if (2.25) is infeasible. Both tests are applicable when short sales are restricted. The test of Lizyayev (2009), on the contrary, assumes unrestricted short sales and in this case is the lightest computationally, albeit bearing the minimal information content: just like Post (2003), it only provides a necessary condition for efficiency of the subject portfolio and a sufficient condition for its weak efficiency.

2.6 Concluding remarks

We can summarize the chapter as follows. We have taken the various methods of three different schools of thought, some of which are developing independently, without any cross-reference to or interaction with the others, placed them under a common umbrella and analyzed them in a unified methodological framework where both the information content of their results and their computational complexity were compared. We have given a principal classification into three categories based on the definition of SSD efficiency employed in each particular method, but we have also seen that methods from different categories can be grouped according to the content of the results and their computational complexity. For many large- or even medium-size data sets some of the methods may become computationally prohibitive, particularly taking bootstrapping into account when the tests have to be repeated many times on similar or even larger data sets simulated from the original distribution. We hope this chapter will assist practitioners in finding a desired tradeoff between bearable computational burden and the information content of the results required.

The methods in each school of thought are based on Definitions 2.1, 2.6 or 2.7. It is remarkable that, although those definitions are proved to be equivalent, the optimization programs corresponding to those definitions substantially differ in terms of their computational complexity, as well as the information content of their outcome.

In addition to classifying and comparing of the methods, we have also corrected some misleading arguments in the literature under consideration and suggested refinements to some of the methods.

Chapter 3

Stochastic Dominance: Convexity and Some Efficiency Tests

This chapter points out the importance of Stochastic Dominance (SD) efficient sets being convex. We review classic convexity and efficient set characterization results on SD efficiency of a given portfolio relative to a diversified set of assets and generalize them in the following aspects. First, we broaden the class of individual utilities in Rubinstein (1974) that lead to two-fund separation. Secondly, we propose a linear programming SSD test that is more efficient than that of Post (2003) and expand the SSD efficiency criteria developed by Dybvig and Ross (1982) onto the Third Order Stochastic Dominance and further to Decreasing Absolute and Increasing Relative Risk Aversion Stochastic Dominance. The efficient sets for those are finite unions of convex sets.

3.1 Introduction

Stochastic Dominance (SD) is a probabilistic concept of relation among different random variables. Unlike parametric criteria such as Mean-Variance analysis, SD accounts for the whole range of distribution function, rather than its particular characteristics such as central moments. Although SD has applications in a huge variety of areas ranging from medicine to agriculture (see, e.g., Bawa (1982), Levy (2006, 1992) for a survey and references, and Eeckhoudt et al. (2009) for recent applications), this Chapter focuses on its use in the area of finance. In financial decision-making one has to select efficient portfolios from an available portfolio possibilities set on the basis of a trade-off among their expected returns, the associated risk of having extreme losses and the potential of earning excessive gains.

We consider the expected utility framework whereby individuals select portfolios maximizing the expected value of their utility function which can capture different individual risk attitudes such as risk aversion, risk neutrality, risk seeking, or a combination of those, for different levels of wealth. The non-parametric nature of SD criteria allows us to identify efficient portfolios without having to specify the utility functions explicitly. Instead, it employs some general restrictions such as non-satiation and risk aversion. The set of all portfolios supported by some utility function in a given class is called the efficient set with respect to this utility class. It turns out that convex efficient sets have a special economic content and hence necessary and sufficient conditions leading to the convexity of efficient sets have been puzzling researchers for more than three decades. Rubinstein (1974) showed that when the preferences of all investors are similar enough, two-fund separation results. Cass and Stiglitz (1970) proved that two-fund separation holds for a more general class of utility functions, assuming that there are only two states of nature for asset returns. On the other hand, Dybvig and Ross (1982) proved that if no assumptions on investors' preferences are made other than concavity and monotonicity, the efficient set is generally non-convex. In line with this concave and monotone utility class Ross (1978) derived some assumptions on the distribution function of returns that lead to k-fund separation. Among recent researchers, Versijp (2007) reviewed Rubinstein's result in relation to stochastic dominance and asset pricing models.

In this Chapter we point out the importance of Stochastic Dominance efficient sets being convex and further review and generalize classic convexity and efficient set characterization results. We will focus on portfolio efficiency with respect to a diversified portfolio possibilities set, normally a polytope whose vertices are the assets available to investors. Further, we consider discrete distribution of returns due to its interpretability via empirically observed data, as well as tractability of the computational methods involved.

This Chapter is organized as follows: Section 3.2 provides general assumptions and problem formulation, Section 3.3 points out the importance of efficient sets being convex and reviews the associated necessary and sufficient conditions, Section 3.4 suggests some efficiency tests for a given portfolio relative to various economically meaningful classes of utility functions, and finally Section 3.5 summarizes the major results and concludes the Chapter.

3.2 Problem formulation and assumptions

Consider a single period investment decision-making problem under uncertainty in a classic expected utility framework, in which:

1. Investors select investment portfolios to maximize the expected utility of the return on their investment portfolio. Let $U = \{u : \mathbb{R} \to \mathbb{R}\}$ denote the class of von Neuman-Morgenstern utility functions and X be the m-by-n matrix of returns of n available assets in m states of the world. The probability of occurrence of state i is denoted π_i. Naturally, $0 < \pi_i \leq 1$, $i = 1, \ldots, m$, and $\sum_{i=1}^{m} \pi_i = 1$. Investors are uncertain about which of the states of nature will occur, but they know the underlying probabilities of the states with certainty.

2. Investors may diversify between available assets. Denote $\lambda \in \Lambda$ for a vector of portfolio weights. Unless otherwise specified, we assume that short sales are allowed and unrestricted. The portfolio possibilities set then becomes
$$\Lambda = \{\lambda \in \mathbb{R}^n : \lambda^\mathsf{T} e = 1\}$$
and the set of all available allocations is
$$M_X = \{x \in \mathbb{R}^m : \ x = X\lambda, \ \ \lambda \in \Lambda\}$$

3. If a riskless asset is available in the market, it will be used either as part of X (a column with equal components), or separately (in which case X will be the set of risky assets only), whichever is more convenient.

A given portfolio $\tau \in \Lambda$ is **optimal** for an investor with utility $u \in U$ if and only if

$$\mathbb{E}u(X) = \sup_{\lambda \in \Lambda} \mathbb{E}u(X\lambda) \tag{3.1}$$

where $\mathbb{E}u$ denotes the expected value of u.

If $\pi_1, \ldots, \pi_m$ are probabilities of occurrence of the states of the world, then $\mathbb{E}u(X)$ becomes

$$\sum_{i=1}^{m} \pi_i u(x^i \lambda) \tag{3.2}$$

where x^i is the i-th row of X. In practical applications full information about utility functions is not available, and (3.2) cannot be verified directly. This provides the rationale for relying on a set of general assumptions, rather than a full specification of the utility function. The Second Order Stochastic Dominance criterion (**SSD**) restricts attention to the class of strictly increasing and concave utility functions, modeling thereby non-satiable and risk-averse preferences. The Third Order SD (**TSD**) assumes that in addition to SSD utilities are positively skewed. A portfolio $\tau \in \Lambda$ is said to be **optimal** in a given utility class U if and only if there exists $u \in U$ such that τ is optimal for u in the sense of (3.1).

A portfolio $\tau \in \Lambda$ is **efficient** if it is not dominated by any other portfolio, that is

$$\forall \lambda \in \Lambda \backslash \{\tau\} \;\; \exists u \in U : \mathbb{E}u(X\lambda) < \mathbb{E}u(X\tau).$$

If both Λ and U are convex (as will be the case in this Chapter), efficiency is equivalent to optimality due to the Minimax theorem (see e.g. Post, 2003) and thus the two concepts will henceforth be used interchangeably.

An individual investor with utility $u \in U$ is facing the following portfolio allocation problem:

$$\max_{\lambda} \sum_{i=1}^{m} \pi_i u(x^i \lambda) - \nu \left(\lambda^{\mathsf{T}} e - 1\right) \tag{3.3}$$

where ν is the Lagrange multiplier. It is known (see e.g. Rockafellar, 1970) that linearity constraints do not alter convexity. So if we assume $u(x)$ to be strictly concave and twice continuously differentiable in x, it will remain concave in λ. In fact, the Hessian of u with respect to λ is $Hu = -Y^{\mathsf{T}}Y$, where Y is an m-by-n matrix defined by $y_{ij} = -x_{ij}\sqrt{-u''(x^i\lambda)}$. Therefore, Hu is always negative semidefinite, and u is concave in λ.

In light of the above-mentioned, the necessary and sufficient condition for τ to be the solution of (3.3) is that there exists $v \in \mathbb{R}$ such that for all $j = 1 \ldots n$

$$\sum_{i=1}^{m} \pi_i x_{ij} u'\left(x^i\tau\right) = \nu. \tag{3.4}$$

Note that if a risk-free asset is available, due to (3.4) there should hold: $\sum_{i=1}^{m} \pi_i u'\left(x^i\tau\right) = \frac{\nu}{r_F}$, where r_F is the risk-free return. Thus, the optimality condition (3.4) takes on the form:

$$\text{for all risky assets } j\text{:}\ \sum_{i=1}^{m} \pi_i (x_{ij} - r_F) u'\left(x^i\tau\right) = 0. \tag{3.5}$$

In financial literature, $u'(r)$ is referred to as the pricing kernel, and $\vartheta(u) \equiv \sum_{i=1}^{m} \pi_i (x_{ij} - r_F) u'\left(x^i\tau\right)$ as the vector of pricing errors: assets j corresponding to $\vartheta_j(u) < 0$ are undervalued and those corresponding to $\vartheta_j(u) > 0$ are overvalued (see Cochrane, 2005). The optimality condition (3.5) is equivalent to all pricing errors being zeros.

We could also relax the twice continuous differentiability assumption for u; (3.4) would still hold in optimality if u is substituted by ∂u – any vector from the supergradient correspondence.

3.3 Convexity

Program (3.3) represents the portfolio formation of an individual investor having a particular well-behaved utility function u. In macroeconomic settings, the aggregate investment decision of a large group of individuals, assuming all of them to be well-behaved, for instance non-satiable risk averters, is of primary importance. A reasonable theoretical model should allow us to extract information about all investors' decisions based on a small number of large composite portfolios. The largest of those, the total value-weighted aggregate portfolio, is generally referred to as the market portfolio, the efficiency of which has been a starting point of many asset pricing theories, including the capital asset pricing model (CAPM).

The simplest case when the market portfolio is efficient is two-fund separation, where any optimal portfolio is a linear combination of two assets, normally a risk-free asset and the market portfolio. In such an economy any individual investor will hold a share of the same (risky) market portfolio and will invest the rest of his/her wealth in the risk-free asset available, i.e. either borrowing or lending at the risk-free rate. Various assumptions lead to two-fund separation, such as: the mean-variance setting (when investment decision is a trade-off only between mean returns and variances of underlying portfolios, see Markowitz (1952, 1978)), homogeneity of preferences (see Rubinstein, 1974), joint normal distribution of asset returns, which is a common assumption of the traditional Capital Asset Pricing Model (CAPM, see Cochrane (2005) for an overview) and quadratic utility functions.

Despite its theoretical appeal, two-fund separation is extremely restrictive and is very unlikely to hold in practice. A straightforward generalization of the concept, preserving the market portfolio efficiency, is k-fund separation, where each efficient portfolio is a linear combination of k fixed mutual funds. Naturally, k-fund separation is of practical and theoretical interest only when $k \ll n$. Generally k-fund separation holds in complete markets (see e.g. Dybvig and Ross, 1982). Ross (1978) derives a necessary and sufficient condition for k-fund separation which however involves returns only; the result is hard to generalize on possible variation of individual preferences and is therefore not particularly informative as far as variations in utility functions are concerned.[1]

A natural further generalization of k-fund separability is the convexity of efficient sets. Indeed, the market portfolio is nothing other than a convex combination of all individual portfolios, and therefore the convexity of an efficient set suffices for the efficiency of the corresponding market portfolio. Indeed, we may assume without loss of generality that individual assets are optimal for at least one investor with a well-behaved utility function and that

[1]In fact, Theorem 3 in Ross (1978) can be seen as a refinement of the classic definition of k-fund separability, as both are given solely in terms of returns and both assume k generating factors.

therefore those assets are efficient. (If, however, we do have an asset whose returns are strictly dominated by another marketed asset or fund, we may as well discard it, as no rational investor will invest in it). Clearly, the market portfolio is now an interior point of a polyhedron whose vertices are all efficient, and if the whole efficient set is known to be convex, efficiency of the market portfolio automatically follows. In addition to being an implication of various asset pricing theories, efficiency of the market portfolio has an intuitive economic interpretation. Observing the popularity of large composite index funds (which proxy the market portfolio) among many individual and institutional investors in practice, it is natural to require that even heterogeneous investors models inconsistent with two-fund separation should imply efficiency of the market portfolio. Moreover, 2- and k-fund separation are merely particular cases of efficient sets being convex.

Conditions which lead to convexity of efficient sets have been challenging researchers for more than three decades already, as it could deliver interesting aggregation results for the models of heterogeneous investors. If an economy is close to satisfying k-fund separation, there is no need for active investment, as every investor is better off investing into k available mutual funds (with specific allocation among those funds determined individually for each investor) and saving on transaction costs associated with active trading strategies. The cases when k-fund separation does not hold but the efficient set is convex are still of theoretical interest, as one could study utility preferences that support large composite portfolios, or test implications of heterogeneous investors models, refining the utility class on the basis of observed individual allocations and composite market indices.

The convexity puzzle can be tackled from two different perspectives: returns on underlying assets and preferences of individual investors. The former would lead the reader towards arbitrage pricing theories and various factor models, while the latter remains not duly researched. Dybvig and Ross (1982) do show with a simple example that SSD efficient set is generally non-convex. However, an SSD efficient set comprises portfolios optimal for all non-satiable risk-averse preferences, many of which are known to be unrealistic. For that reason, after reviewing and providing a more constructive proof of the results in Rubinstein (1974) related to 2-fund separation, we shall summarize the result of Dybvig and Ross (1982) and give efficiency tests for some refined utility classes containing far less unrealistic preferences than all risk averters.

3.3.1 Homogeneous preferences

Rubinstein (1974) considers three heterogeneous investors models in which individual preferences are modeled according to the following utility functions:

$$u(x) \sim -\exp\left(-\frac{x}{A}\right), \qquad A > 0 \tag{3.6a}$$

$$u(x) \sim -\ln(A + x), \qquad A > 0 \tag{3.6b}$$

$$u(x) \sim \frac{(A + Bx)^{(1-b)}}{1-b}, \qquad A > 0, B > 0, b > 0, b \neq 1 \tag{3.6c}$$

Rubinstein (1974) shows that two-fund separation results if all players have the same taste parameters B and beliefs π, but may have different parameters A in (3.6a) and (3.6b), and A and B in (3.6c). He assumes availability of a risk-free asset and requires in addition that $B = \frac{1}{b}$ in (3.6c). Below we sketch a more constructive proof of two-fund separation than the original one of Rubinstein and show that varying B's across individuals will not alter the two-fund separation, provided that the agents have the same power parameter b, even if $B \neq \frac{1}{b}$, thereby generalizing the result of Rubinstein.

Since all the functions above are strictly concave and twice continuously differentiable, a sufficient and necessary condition for portfolio optimality is (3.4). Let r_F be the risk free rate, and X – all risky asset returns. It is convenient to split the portfolio into its risk-free investment α and the remaining risky part $(1 - \alpha)$. The portfolio allocation program now becomes

$$\max_{\lambda \in \Lambda} \sum_{i=1}^{m} \pi_i u\left(1 + \alpha r_F + (1 - \alpha) x^i \lambda\right) - \nu\left(\lambda^\mathsf{T} e - 1\right) \tag{3.7}$$

The optimality conditions (3.4) are now:

$$\begin{cases} \sum\limits_{i=1}^{m} \pi_i x_{ij}(1-\alpha)u'(1+\alpha r_F + (1-\alpha)x^i\tau) = \nu, \quad \forall j \in \mathcal{N} \\ \sum\limits_{i=1}^{m} \pi_i (r_F - x^i\tau)u'(1+\alpha r_F + (1-\alpha)x^i\tau) = 0 \end{cases} \tag{3.8}$$

Let us start with the exponential utility class (3.6a). Suppose a portfolio $(\alpha_1, (1-\alpha_1)\tau)$ is optimal for $u(x, A_1)$. By (3.8), this happens if and only if

$$\begin{cases} \sum\limits_{i=1}^{m} \pi_i x_{ij}(1-\alpha_1)\exp\left(-\frac{1+\alpha_1 r_F+(1-\alpha_1)x^i\tau}{A_1}\right) = \nu_1, \quad \forall j \in \mathcal{N} \\ \sum\limits_{i=1}^{m} \pi_i (r_F - x^i\tau)\exp\left(-\frac{1+\alpha_1 r_F+(1-\alpha_1)x^i\tau}{A_1}\right) = 0 \end{cases} \tag{3.9}$$

One can check by straightforward substitution to (3.9) that for any $A_2 > 0$ the optimal portfolio for the investor with utility $u(x, A_2)$ will be $(\alpha_2, (1-\alpha_2))$, with $\alpha_2 = 1 - \frac{A_2}{A_1(1-\alpha_1)}$. This proves that the efficient portfolio corresponding to $u(x, A_2)$ has the same composition of risky assets. Due to the uniqueness of the solution to (3.9), and continuity of $f(A_2) = 1 - \frac{A_2}{A_1(1-\alpha_1)}$ as a function of A_2, two-fund separation follows.

Note that we can make derivations above only if $\alpha_1 \neq 1$, that is, not all the budget is invested in the riskless asset. The portfolio $(\alpha = 1, 0)$ will not be optimal for any agent $u(x, A)$ with $A > 0$, except in cases where X happens to satisfy the second equation in (3.9) for $\alpha = 1$. However, the risk-free asset will always be asymptotically efficient as risk aversion increases.

One can similarly check by substitution that $(\alpha_2, 1-\alpha_2)$, where

$$\alpha_2 = \frac{A_1 + \alpha_1(1+r_F) - A_2(1-A_1)}{A_1 + 1 + r_F},$$

is optimal for $u(x) = -\ln(A_2 + x)$, whenever $(\alpha_1, 1-\alpha_1)$ is optimal for $u(x) = -\ln(A_1 + x)$.

Analogously, as soon as $(\alpha_1, 1-\alpha_1)$ satisfies (3.8) for $u(x) = \frac{1}{1-b}(A_1 + B_1 x)^{(1-b)}$ with Lagrangean $\nu = \nu_1$, optimality conditions (3.8) for $u(x) = \frac{1}{1-b}(A_2 + B_2 x)^{(1-b)}$

will hold with the optimal portfolio $(\alpha_2, 1-\alpha_2)$, where

$$\alpha_2 = \frac{A_2 B_1(\alpha_1 - 1) + B_2 B_1(1 + r_F) + A_1 B_2}{B_2(r_F B_1 + A_1 + B_1)},$$

and new Lagrangean multiplier

$$\nu_2 = \nu_1 \left(\frac{1-\alpha_2}{1-\alpha_1}\right)^{1-b} \left(\frac{B_2}{B_1}\right)^{-b}.$$

This proves the two-fund separation for homogeneous utility functions of the form (3.6a), (3.6b) and (3.6c).

Note, that this proof generalizes Rubinstein's result, as utility functions in (3.6c) are allowed to have different taste parameters (B's) now, provided they agree on the power parameter b. The restriction $B = \frac{1}{b}$ explicitly imposed in Rubinstein (1974) can be omitted without distorting the two-fund separation.[2] Note also that one explicit assumption behind the derivations above is that the number of underlying assets (including the riskless one) is less than or equal to the number of states: $n + 1 \leq m$.

3.4 Generalizing preferences: non-convexity and some efficiency tests

So far we have analyzed the set of utilities leading to two-fund separation. Although this is a particular case of convex SD efficient sets, it only allows for homogeneous utilities among all investors in the sense that the preferences of all investors are assumed to be parameterized by one or two single parameters, which implies that investors have very similar tastes and as a result take similar investment decisions. Therefore we would like to broaden the class of individual utility functions to allow for heterogeneity among investors. The question is whether the efficient sets for those extended utility classes remain convex. Consider first the set of all risk-averse and non-satiable investors.

[2]As pointed out by an anonymous referee, two-fund separation for utility class (3.6c) has already been proved in Ingersoll (1987, p. 146). We gratefully acknowledge the author's contribution and the referee's note.

Dybvig and Ross (1982) give a simple example of a non-convex second order Stochastic Dominance (i.e. when $U = U_2 \equiv \{u : \mathbb{R} \to \mathbb{R} : u'(x) > 0, u''(x) < 0, \forall x \in \mathbb{R}\}$) efficient set with $n = 3$ assets and $m = 4$ states. They state the following necessary and sufficient conditions for SSD efficiency of portfolio x^0.

An allocation $x^0 \in M_X$ is efficient in U_2 if and only if there exists $z^0 \in \mathbb{R}^m$ such that:

$$\text{(i) } x^\top z^0 \text{ is constant for all } x \in M_X \tag{3.10a}$$

$$\text{(ii) } x_i^0 < x_j^0 \;\Rightarrow\; \frac{z_i^0}{\pi_i} \geq \frac{z_j^0}{\pi_j}, \qquad \forall i, j \tag{3.10b}$$

$$\text{(iii) } z^0 > 0 \tag{3.10c}$$

Vector z^0 can be interpreted as a vector of marginal utility rationalizing portfolio x.

Condition (i) holds only if short sales are allowed and unrestricted. Otherwise (i) holds only for strictly interior points. In general (i) reads: $(x^0)^\top z^0 \geq x^\top z^0, \forall x \in M_X$. Conditions (ii) and (iii) reflect the existence of a strictly concave supporting utility function and the inequality sign may be changed, depending on the properties of the utility class considered. For instance, for strictly concave functions there should hold: $\frac{z_i^0}{\pi_i} \geq \frac{z_j^0}{\pi_j} \Rightarrow x_i^0 \leq x_j^0$; if the functions are in addition differentiable then $\frac{z_i^0}{\pi_i} > \frac{z_j^0}{\pi_j} \Rightarrow x_i^0 < x_j^0$.

The non-convexity example of Dybvig and Ross is both disappointing and challenging. It shows on the one hand that even relative to the set of rather well-behaved preferences the market portfolio can be inefficient. On the other hand, the result challenges us to examine more refined utility classes – after all, if the utility set is restricted to nearly-homogeneous investors as in Rubinstein (1974), not only convexity follows, but even two-fund separation. Taking this into account, below we derive efficiency tests for some higher order SD criteria. In that same section we analyze the case when short sales are allowed and unrestricted, for the following two reasons. Although by far the majority of efficiency tests published assume away short sales, the efficiency of a given portfolio in the unrestricted case implies its efficiency in the restricted case too, whereas a portfolio efficient relative to a restricted portfolio possibilities set may very well be inefficient with respect to the same set with the short sales restriction relaxed. Therefore, the unrestricted case can be seen as a generalization of the restricted short sales and has a practical advantage of not having to specify the exact boundaries for short sales. Moreover, as we shall show further in this section, some algorithms proposed below have superior properties in terms of computational complexity relative to traditional methods in the case when no short sales are imposed. Finally, some of the efficiency tests are only applicable when short sales are restricted, for instance Post (2003) test assumes the portfolio possibilities set to be a polyhedron, so the formulation of the test explicitly includes the vertices of this set.

3.4.1 SSD Efficiency

Although many SSD efficiency tests have been proposed already (see Post (2003), Dentcheva and Ruszczyński (2003), Kuosmanen (2004) and Post and Versijp (2007), among others), we shall focus on linear programming formulations only, since such methods have the lowest computational complexity, which is often a burden for real-life data sets, particularly when it comes to repeating the test many times for statistical inference and bootstrapping or high dimensionality of the data (see for instance Dentcheva and Ruszczyński, 2006b). By far the least computationally demanding SSD efficiency test known is Post (2003). In this section we derive another LP test that is even more efficient than that of Post (2003) in the case when short sales are unrestricted. In the following section we derive a TSD efficiency test, also exploiting the special structure of portfolio possibilities sets and thereby improving its computational complexity.

Consider a given portfolio x. As the ordering of states of the world is not relevant, we may assume without loss of generality that x is sorted in ascending order: $x_1 \leq x_2 \leq \cdots \leq x_m$. In order to determine if x is SSD efficient, we need to find a supporting gradient vector $z \in \mathbb{R}^m$. First note, that condition (3.10a)

$$(X\alpha)^\mathsf{T} z = C, \forall \alpha \in \Lambda$$

is equivalent to $X^\mathsf{T} z = Ce$. We are interested in the case when the market is incomplete and $m > n$. Without loss of generality we may assume that the first n rows of X are linearly independent. Partitioning X into X_1 (first n rows) and X_2 (the rest $(m-n)$ rows), we may write:

$$X^{\mathsf{T}} z = [X_1 X_2]^{\mathsf{T}} z = X_1^{\mathsf{T}} z_{1:n} + X_2^{\mathsf{T}} z_{n+1:m} = Ce.$$

Therefore, the general solution of (3.10a) can be expressed as

$$z = \begin{bmatrix} (X_1^{\mathsf{T}})^{-1}(Ce - X_2^{\mathsf{T}}\beta) \\ \beta \end{bmatrix} \tag{3.11}$$

where β is an $(m-n)$-parameter vector. Since only the ordering of elements of z matters, z can be normalized, so that $C = 1$.

Given the criteria above, the portfolio x is efficient if and only if there exists a decreasing positive vector z satisfying (3.11). If such z exists, it is also a strictly interior point to the following set:

$$\left\{ \beta \in \mathbb{R}^{m-n} : D \begin{bmatrix} -(X_1^{\mathsf{T}})^{-1} X_2^{\mathsf{T}} \\ I_{m-n} \end{bmatrix} \beta \leq -D \begin{bmatrix} (X_1^{\mathsf{T}})^{-1} e \\ 0_{m-n} \end{bmatrix} \right\}, \tag{3.12}$$

$$\text{where } D = \begin{bmatrix} -1 & 1 & 0 & 0 & 0 \\ 0 & -1 & 1 & \ddots & 0 \\ 0 & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 & -1 \end{bmatrix} \cdot \begin{bmatrix} \pi_1^{-1} \\ \pi_2^{-1} \\ \vdots \\ \pi_m^{-1} \end{bmatrix}.$$

This test can be equivalently formulated as the following linear program:

$$\max_{\beta \in \mathbb{R}^{m-n}, \theta \in \mathbb{R}} \left\{ \theta : \begin{bmatrix} -(X_1^{\mathsf{T}})^{-1} X_2^{\mathsf{T}} \\ I_{m-n} \end{bmatrix} \beta + D \begin{bmatrix} (X_1^{\mathsf{T}})^{-1} e \\ 0_{m-n} \end{bmatrix} + \theta \leq 0 \right\}. \tag{3.13}$$

Allocation x is SSD efficient if and only if (3.13) is either unbounded or $\theta^* > 0$.[3]

[3]Equivalently, x is SSD inefficient if (3.13) is infeasible. The case $\theta^* = 0$ also implies

Efficiency test (3.13) is less computationally demanding than that of Post (2003), since (3.13) has $m-n+1$ variables and m constraints, which is n variables and n^2 non-zeros in the constraints matrix less than in Post (2003). By changing variables

$$y_i = \pi_{n+i+1}^{-1}\beta_{i+1} - \pi_{n+i}^{-1}\beta_i, \quad i = 1, \ldots, m-n-1, \text{ and } y_{m-n} = -\pi_m^{-1}\beta_{m-n},$$

one can transform (3.13) to the standard form

$$\max\{c^\mathsf{T} y : Ay \leq b, y \geq 0\}$$

with an n-by-$(m-n)$ matrix of constraints. The number of non-zeros in this matrix is a good indicator of the computational complexity of a linear program (for instance, Performance World (2009) ranks linear programs based on this criterion). The test of Post (2003) in the same standard form will have an n-by-m matrix of constraints, all the elements of which are generally non-zeros. The difference of n^2 non-zero elements confirms the computational advantage of (3.13) relative to Post (2003).

The computational advantage of (3.13) becomes particularly eminent when n approaches m and for instance in the case of bootstrapping, when the efficiency test has to be run many times on multiple data samples generated from the estimated joint distribution of asset returns. However, for large values of n one needs to invert a larger X_1 prior to solving (3.13). Should X happen to be particularly ill-conditioned, one may use the following equivalent test without decomposing X:

$$\max_{z \in \mathbb{R}^m, \theta \in \mathbb{R}} \left\{\theta : X^\mathsf{T} U d = e, d \geq 0\right\}. \tag{3.14}$$

where U is an upper triangular m-by-m matrix adjusted by the probabilities of the states of nature, d is an m-vector representing the probability-adjusted

inefficiency (non-optimality) by our definition. Some authors consider portfolios corresponding to $\theta^* = 0$ efficient as well. In this case the efficiency criterion can be easily adjusted without altering the computational complexity of (3.13).

step differences of vector z, that is

$$d_j = \pi_{j+1}^{-1} z_{j+1} - \pi_j^{-1} z_j, \quad j = 1, \ldots, m-1, \text{ and } d_m = \pi_m^{-1} z_m.$$

Program (3.14) has $m+1$ variables and n equality constraints and is therefore similar to Post (2003) test in terms of computational complexity. However, the two tests are applicable in different circumstances: (3.13) and (3.14) apply when no short sales restrictions are postulated, whereas the test of Post (2003) requires the portfolio possibilities set to be bounded and to contain the subject portfolio x in its interior.

3.4.2 TSD Efficiency

A portfolio $x^0 \in M_X$ is Third Order SD (TSD) efficient if and only if there exists $u_0 \in U_3$ such that $\mathbb{E}u_0(x^0) = \sup_{x \in M_X} \mathbb{E}u_0(x)$, where

$$U_3 \equiv U_2 \cap \left\{ u : \mathbb{R} \to \mathbb{R} \quad \text{s.t. } u'''(x) > 0 \right\}, \text{ and}$$

$$U_2 \equiv \left\{ u : \mathbb{R} \to \mathbb{R} \quad \text{s.t. } u'(t) \geq 0, \text{ and } u''(t) \leq 0, \ \forall t \right\}.$$

Condition $u'''(x) > 0$ can be related to skewness-preference.

Employing concavity of the first derivative of any function in U_3, it is straightforward to formulate TSD efficiency criteria: A portfolio $x^0 \in M_X$ is efficient in U_3 if and only if there exists $z^0 \in \mathbb{R}^m$ such that:

$$\text{(i) } x^\mathsf{T} z^0 \text{ is constant for all } x \in M_X \tag{3.15a}$$

$$\text{(ii) } x_i^0 < x_j^0 \Rightarrow \frac{z_i^0}{\pi_i} \geq \frac{z_j^0}{\pi_j}, \forall i, j \tag{3.15b}$$

$$\text{(iii) } z^0 > 0 \tag{3.15c}$$

$$\text{(iv) } x_i^0 < x_j^0 < x_k^0 \Rightarrow \frac{z_j^0}{\pi_j} \leq \frac{z_k^0}{\pi_k} + \left(\frac{z_i^0}{\pi_i} - \frac{z_k^0}{\pi_k} \right) \frac{x_k^0 - x_j^0}{x_k^0 - x_i^0}, \quad \forall i, j, k \tag{3.15d}$$

The corresponding TSD efficiency test for a given portfolio also leads to a linear programming formulation. Indeed, the TSD criteria (3.15) is simply

$$\bar{D}\begin{bmatrix} -(X_1^\top)^{-1}X_2^\top \\ I_{m-n} \end{bmatrix}\beta \le -\bar{D}\begin{bmatrix} (X_1^\top)^{-1}e \\ 0_{m-n} \end{bmatrix}, \text{ where} \tag{3.16}$$

$$\bar{D} = \begin{bmatrix} a_1 & b_1 & c_1 & 0 & 0 & 0 \\ 0 & a_2 & b_2 & c_2 & \ddots & 0 \\ 0 & 0 & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & \ddots & a_{m-2} & b_{m-2} & c_{m-2} \\ 0 & \ddots & 0 & 0 & a_{m-1} & b_{m-1} \\ 0 & 0 & 0 & 0 & 0 & a_m \end{bmatrix}, \text{ with } \begin{cases} a_{i[i=1\ldots m-2]} = \frac{-1}{\pi_i(x_{i+2}-x_i)} \\ b_{i[i=1\ldots m-2]} = \frac{1}{\pi_{i+1}(x_{i+2}-x_{i+1})} \\ c_{i[i=1\ldots m-2]} = \frac{1}{\pi_{i+2}}\left(\frac{1}{x_{i+2}-x_i} - \frac{1}{x_{i+2}-x_{i+1}}\right) \\ a_{m-1} = -\pi_{m-1}^{-1}, b_{m-1} = \pi_m^{-1} \\ a_m = -\pi_m^{-1} \end{cases}$$

System (3.16) can be solved for β via the same program (3.13) that was applied to the SSD test, with redefined matrix $\bar{D}$. Therefore, the TSD efficiency test is a linear program with $m - n + 1$ variables and m constraints.

3.4.3 Stochastic Dominance for Decreasing Absolute Risk Aversion (DSD)

It is well accepted within expected utility framework that rational individuals possess decreasing absolute risk aversion (DARA).[4] Let us examine Stochastic Dominance efficiency relative to this class of utility functions. Define

$$U_d = U_2 \cap \left\{ \frac{d}{dx}\left(-\frac{u''(x)}{u'(x)}\right) < 0, \forall x \right\}.$$

An allocation $x^0 \in M_X$ is efficient in U_d (DSD efficient) if and only if there exists $u_0 \in U_d$ such that $\mathbb{E}u_0(x^0) = \sup\{\mathbb{E}u_0(x) : x \in M_X\}$.

To express properties of the risk aversion in terms of supporting vectors, we have to adapt the efficiency criterion. Let $r(x) = -\frac{u''(x)}{u'(x)}$ be the absolute risk aversion (ARA) of $u(x)$. We have: $u(x) = \exp\left(-\int r(x)dx + C\right)$, and if

[4] See e.g. Pratt (1964) for formal derivation and discussion.

$x_1 < x_2 < \ldots < x_n$, then

$$u'(x_i) = u'(x_{i-1}) \exp\left(-\int\limits_{x_{i-1}}^{x_i} r(x)dx\right).$$

Therefore, it suffices to require that $r(x) \geq 0$ and $r'(x) < 0$ for all x to ensure that $u \in U_d$. With such $r(x)$, the $\exp\left(-\int\limits_{x_{i-1}}^{x_i} r(x)dx\right)$ is bounded by

$$\exp\left(-r_{i-1}(x_i - x_{i-1})\right) \leq \exp\left(-\int\limits_{x_{i-1}}^{x_i} r(x)dx\right) \leq \exp\left(-r_i(x_i - x_{i-1})\right).$$

Thus, there should hold:

$$\exp\left(-r_{i-1}(x_i - x_{i-1})\right) \leq \frac{u'(x_i)}{u'(x_{i-1})} \leq \exp\left(-r_i(x_i - x_{i-1})\right),$$

and therefore:

$$r_m \leq \ldots \leq r_{i+1} \leq \frac{-\ln\left(\frac{u'(x_{i+1})}{u'(x_i)}\right)}{x_{i+1} - x_i} \leq r_i \leq \frac{-\ln\left(\frac{u'(x_i)}{u'(x_{i-1})}\right)}{x_i - x_{i-1}} \leq r_{i-1} \leq \ldots \leq r_1. \tag{3.17}$$

Clearly, a decreasing sequence $\{r_i\}$ in (3.17) exists if and only if

$$\frac{\ln\left(u'(x_i)\right) - \ln\left(u'(x_{i+1})\right)}{x_{i+1} - x_i} \leq \frac{\ln\left(u'(x_{i-1})\right) - \ln\left(u'(x_i)\right)}{x_i - x_{i-1}}, \text{ for all } i = 2, \ldots, m-1. \tag{3.18}$$

We are now ready to adapt the efficiency definition to the class of U_d. An allocation $x^0 \in M_X$ is efficient in U_d (DSD efficient) if and only if there exists $z^0 \in \mathbb{R}^m$ such that:

$$\text{(i) } x^\intercal z^0 \text{ is constant for all } x \in M_X \tag{3.19a}$$

$$\text{(ii) } x_i^0 < x_j^0 \Longrightarrow \frac{z_i^0}{\pi_i} \geq \frac{z_j^0}{\pi_j}, \forall i,j \tag{3.19b}$$

$$\text{(iii) } z^0 > 0 \tag{3.19c}$$

$$\text{(iv) } x_i^0 < x_j^0 < x_k^0 \Longrightarrow \frac{\ln \frac{z_j^0}{\pi_j} - \ln \frac{z_k^0}{\pi_k}}{x_k^0 - x_j^0} \leq \frac{\ln \frac{z_i^0}{\pi_i} - \ln \frac{z_j^0}{\pi_j}}{x_j^0 - x_i^0}, \quad \forall i,j,k \tag{3.19d}$$

Note that DSD efficiency implies TSD efficiency. This follows from the fact that $r'(x) = u'(x)^{-2} \left(u''(x)^2 - u'(x)u'''(x)\right)$.This is also consistent with DSD-TSD criteria (3.15) and (3.19): (3.19d) implies (3.15d), since the geometric average in (3.19d) can not exceed the arithmetic average in (3.15d).

We are now ready to formulate a test for DSD efficiency of a given portfolio which will no longer be linear, but still a convex program. Indeed, any $z = \begin{bmatrix} -\left(X_1^\intercal\right)^{-1} X_2^\intercal \\ I_{m-n} \end{bmatrix} \beta$ satisfying the DSD criteria (3.19) is also an interior point to the set[5]

$$\hat{D} \ln \left(\begin{bmatrix} -\left(X_1^\intercal\right)^{-1} \left(e - X_2^\intercal\right) \\ I_{m-n} \end{bmatrix} \div \pi \right) \leq 0, \text{ where} \tag{3.20}$$

$$\hat{D} = \begin{bmatrix} a_1 & b_1 & c_1 & 0 & 0 & 0 \\ 0 & a_2 & b_2 & c_2 & \ddots & 0 \\ 0 & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \ddots & 0 & a_{m-2} & b_{m-2} & c_{m-2} \\ 0 & 0 & 0 & 0 & a_{m-1} & b_{m-1} \end{bmatrix}, \text{ with} \begin{cases} a_{i[i=1\ldots m-2]} = \frac{-1}{x_{i+1}-x_i} \\ b_{i[i=1\ldots m-2]} = \frac{1}{x_{i+2}-x_{i+1}} - \frac{1}{x_{i+1}-x_i} \\ c_{i[i=1\ldots m-2]} = \frac{-1}{x_{i+2}-x_{i+1}} \\ a_{m-1} = -1, b_{m-1} = 1 \end{cases}$$

Even though the constraints on β are no longer linear, they are still convex, and therefore we can find strictly feasible points (or establish that they do not exist) efficiently.

[5] By $y = ln(x), x \in \mathbb{R}^m$, we mean the element-wise logarithm, that is, $y \in \mathbb{R}^m$ and $y_i = ln(x_i)$. Similarly, $z = y \div x$, $x, y \in \mathbb{R}^m$, means $z \in \mathbb{R}^m$ s.t. $z_i = y_i/x_i$.

3.4.4 SD for Decreasing Absolute and Increasing Relative Risk Aversion (DISD)

In addition to DARA, relative risk aversion is often postulated to be increasing among rational individuals (see e.g. Pratt, 1964). In this section we examine optimality conditions in the utility class U_{di} combining the two risk aversion properties:

$$U_{di} = U_d \cap \left\{ \frac{d}{dx}\left[-\frac{xu''(x)}{u'(x)} \right] \geq 0, \forall x > 0 \right\}.$$

The utility functions under consideration are therefore those having decreasing absolute (DARA) and increasing relative risk aversion (IRRA). A portfolio $x^0 \in M_X$ is said to be efficient in U_{di} (DISD efficient) if and only if there exists $u_0 \in U_{di}$ such that

$$\mathbb{E}u_0(x^0) = \sup \left\{ \mathbb{E}u(x^0) : u \in U_{di} \right\}.$$

Given the ARA values r_i and r_{i+1} (such that $r_i \geq r_{i+1}$) at nodes x_i and x_{i+1}, the IRRA requirement restricts $r(x)$ to lie above[6] $f(x) = \dfrac{r_i x_i}{x}$, for $x_i \leq x \leq x_{i+1}$, imposing thereby an extra condition:

$$r_i x_i \leq r_{i+1} x_{i+1}, \quad i = 1, \ldots, m-1. \tag{3.21}$$

Conversely, if (3.21) holds, we can always construct $f(x) = \dfrac{r_i x_i}{x}$, for $x_i \leq x \leq x_{i+1}$, such that $xr(x)$ will be non-decreasing, provided that

$$r_{i+1}x_{i+1} + r_i x_i \left(\ln \frac{r_i}{r_{i+1}} - 1 \right) = \int_{x_i}^{x_{i+1}} \max\left\{ \frac{r_i x_i}{x}, r_{i+1} \right\} dx \leq$$
$$\leq \int_{x_i}^{x_{i+1}} r(x) dx = \ln \frac{u'(x_i)}{u'(x_{i+1})} \leq \int_{x_i}^{x_{i+1}} \min\left\{ \frac{r_{i+1}x_{i+1}}{x}, r_i \right\} dx.$$

[6] $f(x)$ is a limiting case of ARA in order for RRA to remain non-decreasing in the interval $[x_{i-1}, x_i]$. It is the solution of $\{(f(x)x)' = 0, f(x_i) = r_i\}$.

Therefore

$$r_{i+1}x_{i+1} + r_i x_i \left(\ln \frac{r_i}{r_{i+1}} - 1 \right) \le \ln \frac{u'(x_i)}{u'(x_{i+1})} \le r_{i+1}x_{i+1} \left(\ln \frac{r_i}{r_{i+1}} + 1 \right) - r_i x_i. \tag{3.22}$$

This leads to the following DISD efficiency criterion: An allocation $x^0 \in M_X$ is efficient in U_{di} (DISD efficient) if and only if there exist z^0 and $r \in \mathbb{R}^m$ such that:

$$\text{(i) } x^{\mathsf{T}} z^0 \text{ is constant for all } x \in M_X \tag{3.23a}$$

$$\text{(ii) } x_i^0 < x_j^0 \Longrightarrow \frac{z_i^0}{\pi_i} \ge \frac{z_j^0}{\pi_j}, \forall i, j \tag{3.23b}$$

$$\text{(iii) } z^0 > 0 \tag{3.23c}$$

$$\text{(iv) } x_i^0 < x_j^0 < x_k^0 \Longrightarrow r_k \le \frac{-\ln \frac{x_k^0 \pi_j}{x_j^0 \pi_k}}{x_k^0 - x_j^0} \le r_j \le \frac{-\ln \frac{x_j^0 \pi_i}{x_i^0 \pi_j}}{x_j^0 - x_i^0} \le r_i \tag{3.23d}$$

$$\begin{aligned} \text{(v) } x_i^0 < x_j^0 \Longrightarrow r_j x_j^0 + r_i x_i^0 \left(\ln \frac{r_i}{r_j} - 1 \right) \le \\ \le -\ln \frac{x_j^0 \pi_i}{x_i^0 \pi_j} \le r_j x_j^0 \left(\ln \frac{r_i}{r_j} + 1 \right) - r_i x_i^0 \end{aligned} \tag{3.23e}$$

The above condition is far less convenient than those for TSD or DSD, as both r and x are now entering (3.23d) and (3.23e) in both linear and logarithmic form.

3.5 Concluding remarks

We have pointed out the importance of stochastic dominance efficient sets being convex and further summarized and extended conditions leading to convexity of efficient sets. This property has great importance both practically (passive vs. active investing strategies, efficiency of mutual funds) and theoretically (heterogeneous investors models and asset pricing) and can be analyzed from two different but interrelated approaches: the returns on underlying assets and the utilities of individual investors. Restricting distributions of returns typically leads to various factor models, where a complete class of non-satiable and risk-averse investors is assumed. Restricting the set of utilities can also affect efficient sets considerably, as can be seen, for example, in Rubinstein (1974). Unfortunately the extent to which restrictions on sets of utilities affect convexity has not been duly researched.

Based on the efficiency criteria (3.19), Dybvig and Ross (1982) derive the following characterization of SSD efficient sets E_{SSD}:

$$E_{SSD} = \bigcup_{z \in Z} \bigcap_{(i,j): \frac{z_i}{\pi_i} < \frac{z_j}{\pi_j}} \{x \in M_X : x_i > x_j\}, \tag{3.24}$$

where the union is taken over all $z \in Z$ having different orderings $\left\{\frac{z_j}{\pi_j}\right\}$. Since the dimensionality of z is m, the number of different orderings is at most $m!$. Thus, E_{SSD} is a union of a finite number of convex sets. By analogy, we can explicitly characterize TSD efficient sets:

$$E_{TSD} = \bigcup_{z \in Z} \Bigg(\bigcap_{(i,j): \frac{z_i}{\pi_i} < \frac{z_j}{\pi_j}} \{x \in M_X : x_i > x_j\} \bigcap_{(i,j,k): x_i < x_j < x_k} \left\{ x \in M_X : \left(\frac{z_j}{\pi_j} - \frac{z_k}{\pi_k}\right)(x_k - x_i) < \left(\frac{z_i}{\pi_i} - \frac{z_k}{\pi_k}\right)(x_k - x_j) \right\} \Bigg), \tag{3.25}$$

Since all restrictions on x are linear, E_{TSD} is again a finite union of convex sets. The same applies for DSD:

$$E_{DSD} = \bigcup_{z \in Z} \Bigg(\bigcap_{(i,j): \frac{z_i}{\pi_i} < \frac{z_j}{\pi_j}} \{x \in M_X : x_i > x_j\} \bigcap_{(i,j,k): x_i < x_j < x_k} \left\{ x \in M_X : \ln\left(\frac{z_j \pi_k}{z_k \pi_j}\right)(x_j - x_i) < \ln\left(\frac{z_i \pi_j}{z_j \pi_i}\right)(x_k - x_j) \right\} \Bigg), \tag{3.26}$$

The DISD efficiency characterization is slightly more complex, as r appears along with z:

$$E_{DISD} = \bigcup_{\substack{z \in Z, \\ r \in \mathbb{R}_+^m}} \Bigg(\bigcap_{(i,j): \frac{z_i}{\pi_i} < \frac{z_j}{\pi_j}} \{x \in M_X : x_i > x_j\}$$
$$\bigcap_{\substack{(i,j,k): x_i < x_j < x_k, \\ x \in M_X}} \left\{ r_k \leq \frac{-\ln \frac{x_k \pi_j}{x_j \pi_k}}{x_k - x_j} \leq r_j \leq \frac{-\ln \frac{x_j \pi_i}{x_i \pi_j}}{x_j - x_i} \leq r_i \right\}$$
$$\bigcap_{\substack{(i,j): x_i < x_j, \\ x \in M_X}} \left\{ r_j x_j + r_i x_i \left(\ln \frac{r_i}{r_j} - 1 \right) \leq -\ln \frac{x_j \pi_i}{x_i \pi_j} \leq r_j x_j \left(\ln \frac{r_i}{r_j} + 1 \right) - r_i x_i \right\} \Bigg). \tag{3.27}$$

It is not clear whether E_{DISD} is necessarily non-convex. On the other hand, it is difficult to find a general set of assumptions that would guarantee convexity of a union of convex sets, in contrast to the intersection of convex sets which is automatically convex.

With regard to the link between utility functions and convexity of efficient sets, there are only extreme cases known so far: when investors are nearly homogeneous (as in Rubinstein, 1974), in which case efficient sets are normally rays or lines, and when investors' preferences are spanned unrealistically broadly, such as the whole U_2, where efficient sets are too large (even without portfolio restrictions on short sales) and non-convex.

A possible extension of the current research could lie in searching for a reasonable set of well-behaved utility functions for which the efficient sets would be large enough and convex. In analogy with arbitrage pricing theories and factor models for returns, one could try to parameterize investors' preferences. The expo-power utility function of the form

$$u(x) = \theta - \exp(-\beta x^\alpha),$$

where $\theta > 1, \alpha \neq 0, \beta \neq 0$, and $\alpha\beta > 0$, seems to be suitable for that, as it allows all possible combinations of absolute (increasing, decreasing or constant) and relative (increasing or decreasing) risk aversion with only two key parameters (see e.g. Saha, 1993).

Another interesting approach for extending the current research would be by studying a joint restrictions on utility functions *and* the distribution of returns. Following this path, Cass and Stiglitz (1970) have proved that two-fund separation holds for a more general class of utility functions than (3.6a), (3.6b) and (3.6c), under the assumption that there are only two states of nature ("binary" returns), and each asset yields zero in one of the states and a positive value in the other. This assumption of "binary returns" is clearly oversimplified and unrealistic, but the idea of restricting utility functions and the distribution of returns jointly is extremely interesting and definitely worth researching.

In addition to the convexity analysis, we have also derived some higher order stochastic dominance efficiency tests in which we incorporate some meaningful restrictions on the set of utilities well-recognized in the expected utility framework, such as decreasing absolute and increasing relative risk aversion.

Chapter 4

Tractable Almost Stochastic Dominance

LL-Almost Stochastic Dominance (LL-ASD) is a relaxation of the Stochastic Dominance (SD) concept proposed by Leshno and Levy (2002) that explains more of realistic preferences observed in practice than SD alone does. Unfortunately, numerical applications of LL-ASD, such as identifying if a given portfolio is LL-ASD efficient or determining a marketed portfolio that LL-ASD dominates a given benchmark, are computationally prohibitive due to the structure of LL-ASD. We propose a new Almost Stochastic Dominance (ASD) concept that is computationally tractable. For instance, a marketed dominating portfolio can be identified by solving a simple linear program. Moreover, the new ASD performs well on all the intuitive examples from Leshno and Levy (2002) and Levy (2009), and in some cases leads to more realistic predictions than those of LL-ASD. We develop some properties of ASD, formulate efficient optimization programs and apply the concept to analyzing investors' preferences between bonds and stocks for the long run.

4.1 Introduction

Modeling the portfolio selection criteria of rational decision makers represents a central problem in the area of finance and asset pricing. Besides the ranking of available portfolios based on such rational rules, an important practical application lies in selecting optimal portfolios among a given portfolio possibilities set, e.g. in determining an investment portfolio that would be superior to a given benchmark for all (or most) "reasonable" investors.

The most cited model for ranking portfolios is the Mean-Variance (MV) rule proposed by Markowitz (1952) which states that one portfolio dominates another whenever it has at least the same mean and at most the same variance. An undeniable advantage of this model from the practical perspective is that the whole MV-efficient set is easy to compute, and in some cases even to express analytically. However, the MV criterion is too stringent in that it fails to capture certain obvious preferences. For instance, any rational investor would prefer portfolio X yielding \$8 and \$1 with equal probabilities $\frac{1}{2}$ to portfolio Y that yields \$4 and \$1 with the same probabilities. Nonetheless, X does not dominate Y by the MV rule.

The relation of Stochastic Dominance, introduced in Mathematics by Mann and Whitney (1947) and Lehmann (1955) and in Economics by Quirk and Saposnik (1962), Hadar and Russell (1969), Hanoch and Levy (1969), Rothschild and Stiglitz (1970, 1971), overcomes this shortcoming. A random variable X dominates a random Y by First Order Stochastic Dominance (FSD) if

$$F_X(t) \leq F_Y(t), \quad t \in \mathbb{R}.$$

It dominates Y by the Second Order Stochastic Dominance (SSD) if

$$F_X^{(2)}(t) \equiv \int_{-\infty}^{t} F_X(z)\, dz \leq \int_{-\infty}^{t} F_Y(z)\, dz \equiv F_Y^{(2)}(t), \quad t \in \mathbb{R}.$$

The SSD criterion uses the full information about the distribution rather than its particular characteristic, and therefore models "reasonable" choices more realistically than MV does. For example, portfolio X above dominates Y by SSD, but not by MV. The Stochastic Dominance rule has another intuitive equivalent formulation: X dominates Y by FSD (SSD) if and only if $\mathbb{E}[u(X)] \geq \mathbb{E}[u(Y)]$ for all increasing (increasing and concave) von Neumann-Morgenstern utility functions $u(t)$. The Stochastic Dominance criteria remain computationally tractable: various tests for efficiency of a given portfolio have been proposed, see among others Dentcheva and

Ruszczyński (2003), Post (2003), Kuosmanen (2004), Kopa and Chovanec (2008), Lizyayev (2009). The methods of Dentcheva and Ruszczyński (2003), Kuosmanen (2004), Kopa and Chovanec (2008) can also be used to identify a portfolio that SSD dominates a given benchmark. In the case of SSD this can be done by solving a linear program, which is a routine task for many available solvers that can handle relatively high dimensional data sets (e.g. GAMS, AIMMS). Dentcheva and Ruszczyński (2006a, 2010) and Rudolf and Ruszczyński (2008) proposed cutting plane formulations that are shown to solve even higher dimensional problems which may be beyond the capabilities of direct solvers. On other approaches to portfolio selection with stochastic dominance criteria, see Mansini et al. (2003), Ruszczyński and Vanderbei (2003), Miller and Ruszczyński (2008).

Although Stochastic Dominance is a tremendous step forward in modeling rational preferences relative to the MV rule, it has its own shortcomings. Leshno and Levy (2002) illustrate them with some examples and suggest a relaxation of the concept, which they name Almost Stochastic Dominance and which we, for the sake of clarity, will henceforth refer to as *LL-Almost Stochastic Dominance*, or LL-ASD.

LL-ASD seems to be an interesting relaxation of Stochastic Dominance, able to model many realistic preferences that Stochastic Dominance would rule out. However, perhaps the largest drawback of LL-ASD is its computational intractability. It is extremely difficult to identify an LL-ASD efficient (non-dominated) portfolio among a given portfolio possibilities set, such as a polyhedron whose vertices are individual assets available on the market, even on a medium-scale data set.

The purpose of this Chapter is to suggest another relaxation of Stochastic Dominance, which we shall call *ε-Almost Stochastic Dominance*, or ε-ASD, in which we allow the expected shortfall constraints in the definition of SD to deviate by a fixed small number ε. We will show that allowing for ε-deviations of expected shortfall constraints is equivalent to imposing such constraints as component-wise slacks. This allows us to generalize the major known SD algorithms (majorization and distribution-based) onto ε-Almost Stochastic Dominance and to formulate several simple linear programs for identifying the ε-ASD efficiency of a given portfolio, or determining a marketed portfolio that ε-ASD dominates a given benchmark, among an infinite set of portfolios available in the market. We also derive some properties of ε-ASD, in particular its interpretation in terms of utility preferences.

Subsequently we test the performance of the new model on various intuitive illustrative examples, some of which were suggested by Leshno and Levy (2002) and Levy (2009), and some others of which are new. The new concept performs well on those examples by correctly modeling "reasonable" preferences, just as LL-ASD does. Moreover, we show that in some intuitive cases such as these, when a highly negative outcome may occur, the proposed ε-ASD correctly models "reasonable" preferences which LL-ASD appears unable to capture.

Finally, a particularly important practical application is modeling investors' preferences between stocks and bonds in the long run, about which there is an ongoing debate in financial literature. Leshno and Levy (2002) argued that most "reasonable" investors would prefer a higher proportion of stocks as the investment horizon increases, motivating it by the decreasing ε-threshold at which stocks LL-ASD dominate bonds over time. However, Levy (2009) pointed out that due to an increasing range of possible outcomes, LL-ASD does not necessarily support the long-term preference for equities, and that, on the contrary, bonds may be preferred to stocks in the long run for standard preferences. In an attempt to address this issue, we propose employing ε-ASD to infer long-term preferences by looking at the dynamics of the ε-threshold at which stocks ε-ASD dominate bonds over time. We will derive theoretical asymptotic properties of this threshold and analyze its convergence in practice. This approach seems to perform well on all the examples given in Leshno and Levy (2002) and Levy (2009).

The Chapter is organized as follows. In Section 4.2 we review LL-ASD and define ε-Almost Stochastic Dominance. We show the equivalence of two different formulations of ε-ASD which we will need later for mathematical programming formulations. We show with some intuitive examples that ε-ASD models "reasonable" preferences correctly, even though its predictions do not always coincide with those of LL-ASD. In Section 4.3 we present several linear programming formulations for identifying a marketed portfolio that ε-ASSD dominates a given benchmark (which is impossible to do for

LL-ASSD dominance). Section 4.4 is devoted to applying ε-ASD to inferring preferences between stocks and bonds in the long run. Finally, Section 4.5 provides a discussion and concludes the Chapter.

4.2 Almost Stochastic Dominance

Leshno and Levy (2002) argue that Stochastic Dominance does not always capture intuitive choices that most "reasonable" investors are likely to make, with the following example.

Example 1. *(Leshno, Levy 2002) Investment X yields \$0.5 with probability 0.01 and \$1,000,000 with probability 0.99, and investment Y yields \$1 without uncertainty.*

Leshno and Levy argue that although X does not dominate Y in Example 1 by either FSD or SSD, most (if not all) "reasonable" investors will prefer X to Y. They conclude that the SD rule is unnecessarily stringent and propose a relaxation of the concept that they name Almost Stochastic Dominance (which we, for the sake of clarity, will henceforth refer to as LL-Almost Stochastic Dominance, or LL-ASD), as follows.

Definition 4.1. *Portfolio X LL-ε-almost dominates portfolio Y by FSD if and only if*

$$\int_{S_1} (F_Y(t) - F_X(t))\, dt \leq \varepsilon \int_S (F_Y(t) - F_X(t))\, dt, \tag{4.1}$$

where S is the combined range of outcomes of X and Y, and

$$S_1 = \{t \in S : F_Y(t) > F_X(t)\}. \tag{4.2}$$

Similarly, X LL-ε-almost dominates Y by SSD if and only if (Leshno, Levy 2002)

$$\int_{S_2} (F_Y(t) - F_X(t))\, dt \leq \varepsilon \int_S (F_Y(t) - F_X(t))\, dt, \tag{4.3}$$

where

$$S_2 = \{t \in S_1 : F_Y^{(2)}(t) > F_X^{(2)}(t)\}. \tag{4.4}$$

Clearly, although X does not dominate Y by either FSD or SSD, it does LL-Almost dominate Y by FSD for $\varepsilon \approx 5 \times 10^{-9}$.

LL-Almost Stochastic Dominance offers an interesting generalization of classic Stochastic Dominance, as demonstrated by Example 1, where the majority of "reasonable" investors will prefer X over Y. One of the drawbacks of this concept, however, is its computational intractability. One can easily check by (4.1) or (4.3) if any given portfolio LL-Almost dominates another fixed portfolio Y; however, determining a portfolio that LL-Almost dominates a given portfolio Y among an infinite portfolio possibilities set (e.g. a set of linear combinations of given assets) for a fixed $\varepsilon > 0$ would be practically impossible on a real-life data set of even medium dimensionality, due to the fractional structure and non-convexity of (4.1) and (4.3). Even identifying whether a given portfolio is LL-ASD efficient (which is a smaller problem) is computationally prohibitive.

For that reason we propose the following modified relaxation of Stochastic Dominance which, in contrast to (4.1) and (4.3), is computationally tractable and in some cases models rational decision-making even more realistically than do (4.1) and (4.3).

Definition 4.2. *A random variable X ε-almost dominates a random variable Y by SSD (or, equivalently, X ε-ASSD dominates Y) if and only if*

$$\forall t \in S : F_X^{(2)}(t) - F_Y^{(2)}(t) \leq \varepsilon. \tag{4.5}$$

We have the following equivalent representation.

Theorem 4.1. *X ε-ASSD dominates Y if and only if there exists a non-negative random variable Z such that $\mathbb{E}[Z] \leq \varepsilon$ and $X + Z$ dominates Y by SSD.*

Proof. Suppose that $Z \geq 0$, $\mathbb{E}[Z] \leq \varepsilon$, and $X + Z$ SSD-dominates Y. The function $t \mapsto (t)_+ = \max(0, t)$ is convex and positively homogeneous. Therefore, for every $\eta \in \mathbb{R}$ we have the inequalities

$$\begin{aligned}\mathbb{E}\big[(\eta - X)_+\big] &= \mathbb{E}\big[(\eta - X - Z + Z)_+\big] \\ &\leq \mathbb{E}\big[(\eta - X - Z)_+\big] + \mathbb{E}[Z] \leq \mathbb{E}\big[(\eta - Y)_+\big] + \varepsilon.\end{aligned}$$

Thus, X ε-ASSD dominates Y.

To prove the converse implication, suppose X ε-ASSD dominates Y. Let d be such that $\mathbb{E}\big[(d-X)_+\big] = \varepsilon$. Defining $Z = (d-X)_+$ we see that $X + Z = \max(d, X)$. For every $\eta \in \mathbb{R}$ we have the obvious identity:

$$\eta - \max(d, X) = (\eta - X) - Z.$$

If $\eta \leq d$, then $\big(\eta - \max(d, X)\big)_+ = 0$; it is sufficient to consider $\eta > d$. But then

$$\big(\eta - \max(X, d)\big)_+ = \begin{cases} (\eta - X)_+, & \text{when } X \geq d, \\ (\eta - X) - (d - X), & \text{when } X < d. \end{cases}$$

It turns out that

$$\mathbb{E}\Big[\big(\eta - \max(d, X)\big)_+\Big] = \mathbb{E}\big[(\eta - X)_+\big] - \mathbb{E}[Z] = \mathbb{E}\big[(\eta - X)_+\big] - \varepsilon.$$

From the definition of ε-ASSD we get

$$\mathbb{E}\big[(\eta - X)_+\big] \leq \mathbb{E}\big[(\eta - Y)_+\big] + \varepsilon.$$

Combining the last two inequalities, we conclude that

$$\mathbb{E}\Big[\big(\eta - (X + Z)\big)_+\Big] \leq \mathbb{E}\big[(\eta - Y)_+\big], \quad \eta \in \mathbb{R},$$

which is SSD of $X + Z$ over Y by Definition 2.6. □

Theorem 4.1 provides a useful interpretation of the value of ε in the ε-ASSD relation: it is the smallest value of the mean return of a random variable that needs to be added to a random variable X in order for it to dominate a given benchmark Y.

In analogy to 4.2 and 4.1, we can define ε-Almost-First Order Stochastic Dominance (ε-AFSD) in the following way.

Definition 4.3. *A random variable X ε-almost dominates a random variable Y by FSD if there exists a nonnegative random variable Z such that $\mathbb{E}[Z] \leq \varepsilon$ and $X + Z$ dominates Y by FSD.*

In our analysis we shall mainly focus on ε-ASSD, as it is related to risk averse preferences which are commonly assumed in expected utility settings (see Pratt (1964)).

Let us now examine how the newly defined concept performs on some intuitive examples. In Example 1 X ε-almost dominates Y by both LL-ASD and ε-ASD for $\varepsilon \geq 0.005$. Therefore, the newly introduced ε-ASD also supports the rational choice of X in Example 1.

Moreover, the following example illustrates a case in which the new concept models rational preferences in a different, perhaps more realistic way.

Example 2. *Consider investment A which leads to the loss of \$1,000,000 (one million) with probability $\frac{1}{2}$ or yields \$1,000,000,000 (one billion), with the same probability $\frac{1}{2}$. Investment B yields \$1,000,000 with certainty.*

Clearly, most (if not all) "reasonable" decision makers will prefer B over A, to avoid the very high risk of losing a million dollars and to receive that amount with certainty instead. As in Example 1, B does not dominate A by either first or second order stochastic dominance, indicating that a modified decision rule is needed. However, LL-Almost-Stochastic Dominance also fails to capture those preferences. Indeed, LL-Almost Stochastic Dominance still predicts that A ε-almost dominates B (by FSD and SSD) for $\varepsilon \approx 0.001998$. On the other hand, it can easily be seen that the newly defined concepts will require ε to exceed 1,000,000 in order for A to ε-almost dominate B. Therefore, the new concept of Almost Stochastic Dominance is more robust in the sense that it works well in instances related to both Example 1 and Example 2, whereas LL-ASD only handles the first example correctly, and falls short for modelling the "reasonable" choice in Example 2. In order to be fully objective it is worth noting that neither ε-ASD nor LL-ASD predict an almost-dominance of B over A for a reasonably small ε. The key difference is that LL-almost dominance implies that A is preferred over B by most "reasonable" investors which is unrealistic, whereas our new concept of ε-Almost-SD correctly models the opposite.

The key feature that makes Example 2 question the LL-Almost-SD is most "reasonable" investors' aversion to losses. Note that Example 1 contains only positive outcomes for investments X and Y. However, the concept of

Stochastic Dominance (including Almost-Stochastic Dominance) is certainly relevant for the negative outcomes as well, and Levy (2009) also uses negative stock returns in Tables 1 and 2.

Example 2 also illustrates that, even though the LL-dominating portfolios are preferred to portfolios they LL-dominate for many "reasonable" utility functions, there may still exist other "reasonable" utility functions for which dominating portfolios have much lower expected values than those of the dominated ones (take $u(x) = -e^{-x}$ in Example 2). Therefore, the value of ε in LL-ASD is difficult to interpret in probabilistic terms. Our new definition, on the contrary, ensures that the expected return of the dominated portfolio does not exceed that of dominating portfolios more than ε, for all utility functions spanning the set U_2 of concave nondecreasing functions having bounded slope.

Inspired by (Leshno and Levy, 2002, Thm. 1), below we prove an equivalent ε-ASSD formulation in terms of utility functions. Since a scaling of a utility function does not change the optimal portfolio that maximizes the expected value of that function, we may without loss of generality restrict the set U_2 to the following set:

$$\tilde{U}_2 = \{u \in U_2 : u'(t) \leq 1\}. \tag{4.6}$$

Indeed, any function $u \in U_2$ defined on a finite support, say an interval $[a, b]$, can be substituted with $\tilde{u}(t) \equiv \frac{u(t)}{u'(a)}$ which will preserve the optimal solution; in general case, $\tilde{u}(t) \equiv \frac{u(t)}{\sup_{t \in S} u'(t)}$ will fall in $\tilde{U}_2$.

Theorem 4.2. *A random variable X ε-ASSD dominates a random variable Y by SSD if and only if*

$$\mathbb{E}[u(X)] + \varepsilon \geq \mathbb{E}[u(Y)] \quad \textit{for all} \quad u \in \tilde{U}_2. \tag{4.7}$$

Proof. For a fixed $\eta \in \mathbb{R}$ consider a particular utility function

$$u_\eta(t) = -(\eta - t)_+, \quad t \in \mathbb{R}.$$

We can conclude directly from Definition 4.2 that the ε-ASSD dominance is equivalent to the relation:

$$\mathbb{E}[u_\eta(X)] + \varepsilon \geq \mathbb{E}[u_\eta(Y)] \quad \text{for all} \quad \eta \in \mathbb{R}. \tag{4.8}$$

As $u_\eta \in \tilde{U}_2$, relation (4.7) implies ε-ASSD dominance.

To prove the converse implication, consider an arbitrary $u \in \tilde{U}_2$. For every $\delta > 0$ we can find a finite collection of numbers η_k and $\alpha_k \geq 0$, $k = 1, \dots, K$, and a constant c such that $\sum_{k=1}^K \alpha_k = 1$ and the function

$$w(t) = c + \sum_{k=1}^{K} \alpha_k u_{\eta_k}(t)$$

has the following properties

$$\mathbb{E}\big[|u(X) - w(X)|\big] \leq \delta,$$
$$\mathbb{E}\big[|u(Y) - w(Y)|\big] \leq \delta.$$

This collection can be constructed by a sufficiently accurate piecewise linear approximation of the function $u(\cdot)$. Since the α_k's are nonnegative and total 1, $w \in \tilde{U}_2$. Adding inequalities (4.8) multiplied by α_k for each $u_{\eta_k}(\cdot)$ we obtain

$$\mathbb{E}[w(X)] + \varepsilon \geq \mathbb{E}[w(Y)].$$

Then

$$\mathbb{E}[u(X)] + \varepsilon + 2\delta \geq \mathbb{E}[u(Y)].$$

As $\delta > 0$ was arbitrary, inequality (4.7) follows. □

Moreover, the following proposition, similar to Leshno and Levy (2002, Prop. 3) for LL-ASD, holds for the newly introduced concepts of ε-Almost-SD.

Proposition 4.1. *X dominates Y by FSD (SSD) if and only if there exists a $d > 0$ such that X ε-almost dominates Y by FSD (SSD) for all $\varepsilon \leq d$.*

Proof. Necessity is obvious, since ε-Almost Dominance (as well as LL-Almost Dominance) is no more than a relaxation of classic Stochastic Dominance. Sufficiency follows from the fact that if X dominates Y by FSD (SSD), then there exists $z \in S$ such that $\nu = F_X(z) - F_Y(z) > 0$ $\left(\nu = F_X^{(2)}(z) - F_Y^{(2)}(z) > 0\right)$, and therefore X cannot ε-dominate Y for $\varepsilon < \nu$. □

However, the most important advantage of our modification is its computational tractability. In the following section we formulate optimization models with ε-Almost Stochastic Dominance constraints that can be used in investment modeling. These models identify a portfolio that ε-almost dominates a given benchmark, among an infinite set of portfolios available in the market. Models corresponding to ε-ASSD (4.2 and the equivalent representation due to 4.1) turn out to be linear programs that can easily be solved by standard software. Programs with ε-Almost-FSD constraints will belong to the same class as those with standard FSD constraints, to mixed integer linear programs.

4.3 Linear Programming Models for ε-ASSD

Consider m states of the world each of which occurs with probability $\pi_i, i = 1, \dots, m$. Naturally, we assume $0 < \pi_i \leq 1, \quad i = 1, \dots, m,$ and $\sum_{i=1}^m \pi_i = 1$.

Suppose there are n assets available in the market, with $n < m$. Denote the return over asset j in state i as x_{ij} and combine $x_{ij}, i = 1, \dots, m, j = 1, \dots, n$ into m-by-n matrix X. Without loss of generality we assume that all returns (the columns of X) are linearly independent, which implies that the matrix $X^\mathsf{T} X$ is positive definite. Finally, denote the return over the benchmark portfolio at state i by y_i.

Lizyayev (2010) classifies programs with second order stochastic dominance constraints into three categories: majorization, distribution-based and revealed preference-type. Below we give the majorization and distribution-based formulations for the newly defined ε-Almost SD.

The distribution-based program of Dentcheva and Ruszczyński (2003) can be formulated in two different ways, due to Theorem 4.1:

$$\max \ f(x) = \mathbb{E}(X\lambda) \tag{4.9a}$$

$$\text{s.t.} \sum_{k=1}^{n} x_{ik}\lambda_k + s_{ij} \geq y_j, \qquad i,j = 1,\ldots,m, \tag{4.9b}$$

$$\sum_{i=1}^{m} s_{ij}\pi_i \leq F_Y^{(2)}(y_j) + \varepsilon, \qquad j = 1,\ldots,m, \tag{4.9c}$$

$$s_{ij} \geq 0, \qquad i,j = 1,\ldots,m, \tag{4.9d}$$

$$\lambda \in \Lambda \tag{4.9e}$$

and

$$\max \ f(x) = \mathbb{E}(X\lambda) \tag{4.10a}$$

$$\text{s.t.} \sum_{k=1}^{n} x_{ik}\lambda_k + d_i + s_{ij} \geq y_j, \qquad i,j = 1,\ldots,m, \tag{4.10b}$$

$$\sum_{i=1}^{m} s_{ij}\pi_i \leq F_Y^{(2)}(y_j), \qquad j = 1,\ldots,m, \tag{4.10c}$$

$$\sum_{i=1}^{m} d_i\pi_i \leq \varepsilon \tag{4.10d}$$

$$s_{ij} \geq 0, \qquad i,j = 1,\ldots,m, \tag{4.10e}$$

$$d_i \geq 0, \qquad i = 1,\ldots,m, \tag{4.10f}$$

$$\lambda \in \Lambda \tag{4.10g}$$

The majorization approach (see Kuosmanen (2004), Luedtke (2008)) can be adapted to ε-ASSD in the following way.

$$\max \; f(x) = \mathbb{E}(X\lambda) \tag{4.11a}$$

$$\text{s.t.} \sum_{k=1}^{n} x_{ik}\lambda_k + d_i - \sum_{j=1}^{m} w_{ij} y_j \geq 0, \qquad i = 1, \ldots, m, \tag{4.11b}$$

$$\sum_{j=1}^{m} w_{ij} = 1, \qquad i = 1, \ldots, m, \tag{4.11c}$$

$$\sum_{i=1}^{m} w_{ij}\pi_i = \pi_j, \qquad j = 1, \ldots, m, \tag{4.11d}$$

$$\sum_{i=1}^{m} d_i \pi_i \leq \varepsilon \tag{4.11e}$$

$$w_{ij} \geq 0, \qquad i, j = 1, \ldots, m, \tag{4.11f}$$

$$d_i \geq 0, \qquad i = 1, \ldots, m, \tag{4.11g}$$

$$\lambda \in \Lambda \tag{4.11h}$$

If we restrict the w_{ij}'s in (4.11f) to be binary, we obtain a model with ε-AFSD constraints.

Linear programming problems (4.9), (4.10) and (4.11) can be used for ε-ASSD efficiency testing of a given benchmark portfolio y, as well as for identifying a marketed portfolio $\lambda^* \in \Lambda$ that dominates y. As standard linear programs, (4.9), (4.10) and (4.11) can be solved by available optimization software efficiently.

Unfortunately, simpler revealed preference-type tests, such as Post (2003), are not applicable to ε-ASSD, as those methods implicitly rely on the ordering of outcomes of y which may change with allowing the perturbation ε.

4.4 Application: Stocks in the long run

The preferences of reasonable investors for stocks versus bonds in the long run are the subject of an ongoing debate in financial literature. While many scholars and practitioners argue in favour of stocks for the long run (see among others Bernstein (1976), Markowitz (1976, 2006)), there are also those who question this conclusion (for instance, Samuelson (1969, 1989, 1994) argues that the optimal allocation of stocks remains constant over time for certain classes of utility functions). Canner et al. (1997) point out that traditional asset pricing theory fails to explain such an allocation of assets with an increasing proportion of stocks over time, emphasizing that

> "Developing portfolio models that [may help explain popular advice on portfolio allocation] and that are simple enough to implement empirically remains a challenge for future research."

Leshno and Levy (2002) approached the issue from the perspective of LL-Almost Stochastic Dominance. They argued that as the investment horizon increases, the ε-threshold at which stocks LL-almost dominate bonds, decreases, and concluded that optimal portfolios contain increasingly larger portions of stocks in the long run. Levy (2009) suggested that despite the decreasing ε-threshold, the range of marginal utilities for some particular functions decreases faster than the threshold does, due to the increasing range of possible outcomes for stock returns. He used the following example:

Example 3. *A bond yields* 4% *with certainty. A stock yields* -10% *with probability* $\frac{1}{2}$ *and* 30% *with probability* $\frac{1}{2}$.

Levy (2009) considered the following utility functions:

$$u(x) = -e^{-x}; \qquad u(x) = \log(x);$$
$$u(x) = \frac{x^{1-\gamma}}{1-\gamma},\ \gamma = 4; \qquad u(x) = \frac{(x-D)^{1-\gamma}}{1-\gamma},\ \gamma = 2.$$

Although the reasoning of Levy (2009) makes perfect sense, it might be difficult to generalize the statement to an infinite set of utility functions based on just four particular functions. For instance, taking

$$u(x) = \log(x + 30),$$

one can easily see that this function remains in the relevant set (in the notations of Levy (2009)) for all the periods considered, and so does the whole continuum of functions

$$u(x) = \log(x + a), \text{where } a \geq 30. \tag{4.12}$$

Note that all those functions are not far-fetched – they are in fact quite realistic: for any $a > 0$, (4.12) has decreasing relative and increasing absolute risk aversion, which is in line with Pratt (1964).

Levy (2009) notes that as the investment horizon increases, two opposite effects take place. On the one hand, the ε-threshold for LL-dominance of bonds over equities decreases, which enlarges the set of utility functions that favour stocks. This is usually taken as an argument in favour of stocks in the long run (as in Leshno and Levy, 2002). On the other hand, the increasing time horizon enlarges the range of possible outcomes, which in turn causes the set of utilities that favour stocks to shrink. Levy (2009) rightly points out that the latter effect is not taken into account by LL-Almost SD as far as the analysis of long-term preferences between bonds and stocks is concerned. Since Levy (2009) only pointed out the problem but did not offer a solution, we are faced with the challenge of applying the newly defined ε-Almost-SD concept to analyzing stock returns in the long run. The following simple rule of thumb has an intuitive interpretation: if the ε-threshold at which stocks ε-almost dominate bonds increases with the time horizon, stocks are not preferred to bonds in the long run, whereas an ε-threshold which is decreasing or constant over time implies that the portion of stocks in each investor's optimal portfolio increases in the long run. This criterion is in accordance with the classic Stochastic Dominance concept which places the most importance on avoiding the risk of losses, whereas LL-Almost SD implies that such risk can be compensated by upside potential in a *linear* fashion (that is: for any fixed ε, any increase in downside risk can be compensated by an increase in upside potential multiplied by a constant coefficient), which may be difficult to accept in practice.

Let us examine the similarities and differences between the two Almost SD criteria in the context of long-run preferences. Consider first the example given in Leshno and Levy (2002, sec. 2).

Example 4. *Consider cumulative return at time n over portfolios*

$$X^{(n)} = \prod_{i=1}^{n}(1+X_i) \text{ and } Y^{(n)} = \prod_{i=1}^{n}(1+Y_i), \tag{4.13}$$

where returns over each period are independent and identically distributed as follows: X_i takes value 5% with probability 0.1 and 12% with probability 0.9, and Y_i yields 7% w. p. 0.4 and 9% w. p. 0.6.

One can see that in this case ε in LL-Almost FSD decreases from $\varepsilon^{(1)} = 0.057$ to $\varepsilon^{(20)} = 5.5 \times 10^{-13}$, as the time horizon increases from $t = 1$ to 20. Similarly, the ε-Almost-FSD threshold decreases from $\varepsilon^{(1)} = 0.002$ to $\varepsilon^{(20)} = 2 \times 10^{-12}$, which means that the implications of the two criteria concur in predicting that portfolios with the lesser proportion of stocks (Y) will be increasingly preferred to more risky X in the long run.

However, the two criteria may very well differ in their predictions, which can be exemplified with the following

Example 5. *Consider cumulative product returns $X^{(n)}$ and $Y^{(n)}$ as in Example 4 with independent increments distributed as follows: X yields -20% with probability 0.4 and 25% with probability 0.6, and Y yields 3% with certainty.*

Clearly, X and Y do not dominate each other by FSD or SSD (and neither by the mean-variance rule). Applying the two Almost-SD criteria, one can check that the ε-threshold, at which X dominates Y by LL-Almost FSD, decreases from $\varepsilon^{(1)} = 0.411$ to $\varepsilon^{(50)} = 0.028$, whereas the ε-Almost-FSD threshold increases from $\varepsilon^{(1)} = 0.092$ to $\varepsilon^{(50)} = 0.733$ as the time horizon increases from $t = 1$ to 50. This is the case when the two models diverge in predictions: according to LL-Almost FSD, the portion of stocks should increase in the long run, while our ε-Almost FSD model does not predict that. The cumulative distribution functions of $X^{(50)}$ and $Y^{(50)}$ at time $t = 50$ are shown on Figure 4.1. The long-run dynamics in this example clearly differs from that of Example 4: the "negative" area between the cumulative distribution functions increases over time, and thus fewer and fewer investors will be willing to accept the increasing risk as the time goes by. The decreasing ε-threshold in LL-ASD indicates that the *linear* increase rate of the upside potential is higher than the linear rate of decrease of the downside risk. However, strictly risk-averse decision-makers require a larger-than-linear compensation in upside potential for increasing downside risk. As illustrated in Figure 4.1, even after 50 investment periods there is no clear dominance of $X^{(n)}$ over $Y^{(n)}$ in this case.

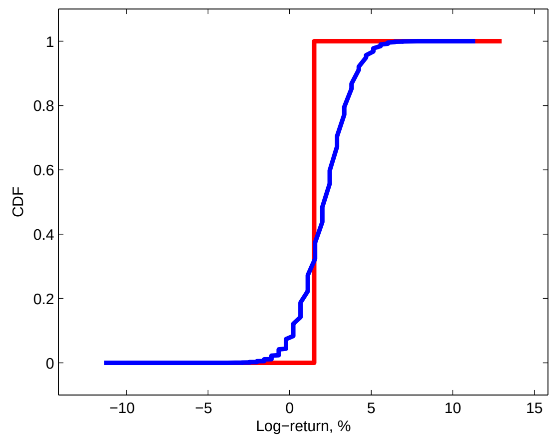


Figure 4.1: Distribution of bonds vs. stocks in Example 5; T=50.

To generalize the examples related to the long-run preferences, below we derive the exact expression for the ε-threshold. For the sake of generality we will allow an arbitrary number of outcomes for X and Y.

Consider the cumulative product returns at time n given in (4.13), where X_i and Y_i are independent and distributed as follows: each X_i takes values $x_1, x_2, \ldots, x_m$ with probabilities $p_1, p_2, \ldots, p_m$, and Y_i is $y_1, \ldots, y_M$ w. p. $q_1, \ldots, q_M$, where $\sum_{k=1}^{m} p_k = \sum_{k=1}^{M} q_k = 1$, $p_k > 0$ and $q_k > 0$.

Let $\mathbb{E}(X) = \mu_X > \mu_Y = \mathbb{E}(Y)$, and $\sigma_X > \sigma_Y$.

Clearly, $\ln(X^{(n)}) = \sum_{k=1}^{n} \ln(1 + X_i)$ and $\ln(Y^{(n)}) = \sum_{k=1}^{n} \ln(1 + Y_i)$ are asymptotically normally distributed, due to the central limit theorem. Moreover,

$$\mu_X^{(n)} = \mathbb{E}\left(X^{(n)}\right) = n \sum_{k=1}^{m} \ln(1 + x_k) p_k = n\mu_X, \text{ and}$$

$$\sigma_X^{(n)} = \sqrt{n}\sigma_X = \sqrt{n} \left(\sum_{k=1}^{m} \ln(1 + x_k)^2 p_k - \mu_X^2 \right)^{\frac{1}{2}}.$$

Analogously,

$$\mu_Y^{(n)} = \mathbb{E}(Y^{(n)}) = n\mu_Y, \text{ and } \sigma_Y^{(n)} = \sqrt{n}\sigma_Y.$$

Let $L_X \sim \mathcal{N}(\mu_X^{(n)}, \sigma_X^{(n)})$ and $L_Y \sim \mathcal{N}(\mu_Y^{(n)}, \sigma_Y^{(n)})$. Obviously, the cumulative distribution functions of L_X and L_Y, $\Phi\left(\frac{x-\mu_X^{(n)}}{\sigma_X^{(n)}}\right)$ and $\Phi\left(\frac{x-\mu_Y^{(n)}}{\sigma_Y^{(n)}}\right)$, intersect precisely once at the point

$$x^* = \frac{\mu_Y^{(n)}\sigma_X^{(n)} - \mu_X^{(n)}\sigma_Y^{(n)}}{\sigma_X^{(n)} - \sigma_Y^{(n)}} = n\frac{\mu_Y\sigma_X - \mu_X\sigma_Y}{\sigma_X - \sigma_Y}.$$

Therefore, due to the asymptotic convergence of $\ln(X^{(n)})$ to L_X and $\ln(Y^{(n)})$ to L_Y in probability, the ε-threshold at which $\ln(X^{(n)})$ ε-ASSD dominates $\ln(Y^{(n)})$, can be computed as follows:

$$\varepsilon^* = \int_{-\infty}^{x^*} \Phi\left(\frac{x-\mu_X^{(n)}}{\sigma_X^{(n)}}\right) - \Phi\left(\frac{x-\mu_Y^{(n)}}{\sigma_Y^{(n)}}\right) dx =$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{n\frac{\mu_Y\sigma_X-\mu_X\sigma_Y}{\sigma_X-\sigma_Y}} \left(\int_{-\infty}^{\frac{x-n\mu_X}{\sqrt{n}\sigma_X}} e^{-t^2/2}dt - \int_{-\infty}^{\frac{x-n\mu_Y}{\sqrt{n}\sigma_Y}} e^{-t^2/2}dt \right) dx. \quad (4.14)$$

Since $x \le x^*$ in (4.14), $\Phi\left(\frac{x-\mu_X^{(n)}}{\sigma_X^{(n)}}\right) \ge \Phi\left(\frac{x-\mu_Y^{(n)}}{\sigma_Y^{(n)}}\right)$, and therefore $\frac{x-n\mu_X}{\sqrt{n}\sigma_X} \ge \frac{x-n\mu_Y}{\sqrt{n}\sigma_Y}$. We have:

$$\varepsilon^* = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{n\frac{\mu_Y\sigma_X-\mu_X\sigma_Y}{\sigma_X-\sigma_Y}} \left(\int_{\frac{x-n\mu_Y}{\sqrt{n}\sigma_Y}}^{\frac{x-n\mu_X}{\sqrt{n}\sigma_X}} e^{-t^2/2}dt \right) dx. \quad (4.15)$$

Finally, by changing the order of integration in (4.15), we have:

$$\varepsilon^* = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{n\frac{\mu_Y\sigma_X - \mu_X\sigma_Y}{\sigma_X - \sigma_Y}} \left(\int_{\frac{x - n\mu_Y}{\sqrt{n}\sigma_Y}}^{\frac{x-n\mu_X}{\sqrt{n}\sigma_X}} e^{-t^2/2} dt \right) dx =$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{x^*-n\mu_X}{\sqrt{n}\sigma_X}} \left(\int_{n\mu_X + t\sqrt{n}\sigma_X}^{n\mu_Y + t\sqrt{n}\sigma_Y} e^{-t^2/2} dx \right) dt =$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{x^*-n\mu_X}{\sqrt{n}\sigma_X}} e^{-t^2/2} \left(n(\mu_Y - \mu_X) + t\sqrt{n}(\sigma_Y - \sigma_X) \right) dt =$$

$$= \frac{\sqrt{n}}{\sqrt{2\pi}} e^{-1/2 \frac{n(\mu_X - \mu_Y)^2}{(\sigma_Y - \sigma_X)^2}} (\sigma_X - \sigma_Y) + n\Phi\left(\frac{\sqrt{n}\,(\mu_X - \mu_Y)}{\sigma_Y - \sigma_X} \right) (\mu_Y - \mu_X). \tag{4.16}$$

Since both $axe^{-x^2} \to 0$ and $bx^2\Phi(-x) \to 0$, for any $a, b > 0$, as $x \to \infty$, it follows that $\varepsilon^* \to 0$ as $n \to \infty$, and therefore the riskier $\log X^{(n)}$ will eventually dominate $\log Y^{(n)}$, for any threshold ε, no matter how small it is. However, even though (4.16) guarantees that from some moment on the log-return over stocks will ε-ASSD dominate that of bonds, we have seen from Example 5 that this may take a long time, and the ε-threshold may very well be increasing during a rather long period of time. Figure 4.2 presents the LL-ε, ε-ASD for the exact (binomial) distribution of returns and the theoretical ε-ASD based on asymptotic normal approximation (4.16). Note that ε-ASD increases until $n = 79$.

Similarly, we can compute the threshold $\varepsilon^{(n)}$ at which $X^{(n)}$ ASSD dominates $Y^{(n)}$ by using the fact that $X^{(n)}$ and $Y^{(n)}$ are asymptotically log-normally distributed, in the following way.

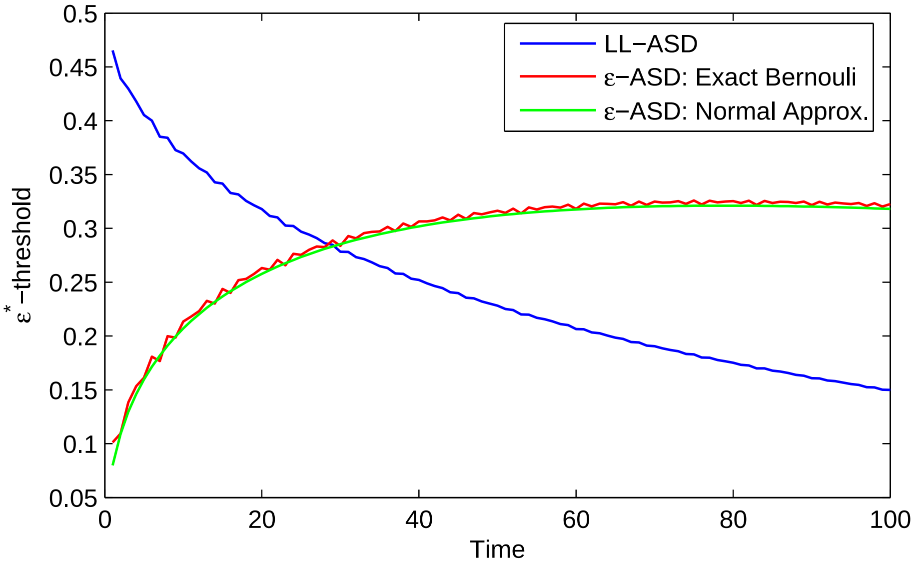


Figure 4.2: ε^*-threshold for $\log X^{(n)}$ and $\log Y^{(n)}$ from Example 5.

$$
\begin{aligned}
\varepsilon^{(n)} &= \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{\exp x^*} dx \int\limits_{\frac{\ln x - n\mu_Y}{\sqrt{n}\sigma_Y}}^{\frac{\ln x - n\mu_X}{\sqrt{n}\sigma_X}} e^{-t^2/2} dt = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{\frac{x^* - n\mu_X}{\sqrt{n}\sigma_X}} dt \int\limits_{e^{\left(n\mu_X + t\sqrt{n}\sigma_X\right)}}^{e^{\left(n\mu_Y + t\sqrt{n}\sigma_Y\right)}} e^{-t^2/2} dx = \\
&= \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{\frac{x^* - n\mu_X}{\sqrt{n}\sigma_X}} e^{-t^2/2} \left(e^{n\mu_Y + t\sqrt{n}\sigma_Y} - e^{n\mu_X + t\sqrt{n}\sigma_X} \right) dt = \\
&= \Phi\left(-\frac{\sqrt{n}\left(\sigma_Y^2 - \sigma_X\sigma_Y - \mu_X + \mu_Y\right)}{\sigma_Y - \sigma_X} \right) e^{n\left(\mu_Y + \frac{\sigma_Y^2}{2}\right)} - \\
&\qquad - \Phi\left(-\frac{\sqrt{n}\left(\sigma_X\sigma_Y - \sigma_X^2 - \mu_X + \mu_Y\right)}{\sigma_Y - \sigma_X} \right) e^{n\left(\mu_X + \frac{\sigma_X^2}{2}\right)}.
\end{aligned}
\tag{4.17}
$$

Applying L'Hôpital's rule, one can ascertain that

$$\frac{\Phi(-ax)}{e^{-bx^2}} \xrightarrow[x\to+\infty]{} \begin{cases} 0, & \text{if } b \le \frac{a^2}{2} \\ +\infty, & \text{if } b > \frac{a^2}{2}. \end{cases} \tag{4.18}$$

Denote

$$a_1 = \frac{\sigma_Y^2 - \sigma_X\sigma_Y - \mu_X + \mu_Y}{\sigma_Y - \sigma_X}, \qquad b_1 = \mu_Y + \frac{\sigma_Y^2}{2}, \tag{4.19}$$

$$a_2 = \frac{\sigma_X\sigma_Y - \sigma_X^2 - \mu_X + \mu_Y}{\sigma_Y - \sigma_X}, \qquad b_2 = \mu_X + \frac{\sigma_X^2}{2}. \tag{4.20}$$

It follows from (4.19) and (4.20) that

$$b_1 - \frac{a_1^2}{2} = b_2 - \frac{a_2^2}{2}. \tag{4.21}$$

Due to (4.18) and (4.21), $\varepsilon^{(n)}$ defined by (4.17) approaches 0 as $n \to \infty$, if $b_1 \le \frac{a_1^2}{2}$. If $b_1 > \frac{a_1^2}{2}$, $\varepsilon^{(n)}$ in (4.17) is a difference of two functions, both tending to infinity as n increases. Since $\varepsilon^{(n)} \ge 0$ for any n, $\varepsilon^{(n)}$ can either tend to $+\infty$, or to a finite limit $C \in \mathbb{R}^+$. The latter would imply that

$$e^{(b_1-b_2)x^2}\frac{\Phi(-a_1x)}{\Phi(-a_2x)} \overset{\text{by (4.21)}}{=} \frac{\Phi(-a_1x)}{e^{-\frac{1}{2}a_1^2x^2}}\frac{e^{-\frac{1}{2}a_2^2x^2}}{\Phi(-a_2x)} = \frac{a_1}{a_2}\frac{d\ln\Phi(-a_2x)}{d\ln\Phi(-a_1x)} \xrightarrow[\text{as } x\to\infty]{} 1 \tag{4.22}$$

Thus, by L'Hôpital's rule, $\frac{\ln\Phi(-a_2x)}{\ln\Phi(-a_1x)} \to \frac{a_2}{a_1}$, and hence $\Phi(-a_2x) - \Phi(-a_1x) \to e^{\frac{a_2}{a_1}}$, as $x \to \infty$. The latter contradicts the fact that $\Phi(-a_2x) - \Phi(-a_1x) \to 0$. This contradiction rules out the possibility of $\varepsilon^{(n)}$ having a finite limit. Therefore, we have just proved the following relations:

$$\lim_{n\to+\infty} \varepsilon^{(n)} = \begin{cases} 0, & \text{if } b_1 \le \frac{a_1^2}{2} \\ +\infty, & \text{if } b_1 > \frac{a_1^2}{2} \end{cases} \tag{4.23}$$

It is remarkable that although $\ln X^{(n)}$ will always eventually ASD dominate $\ln Y^{(n)}$, this does not necessarily hold for $X^{(n)}$ and $Y^{(n)}$. Table 4.1 presents the values of a_1, b_1, a_2 and b_2 defined by (4.19) – (4.20), and the difference $b - \frac{1}{2}a^2$ for Examples 1 through 5. Our ASD model concurs with

Table 4.1: Asymptotic preferences in Examples 1 through 5.

Example Nr.	a_1	b_1	a_2	b_2	$b - \frac{1}{2}a^2$	$\lim \varepsilon^{(n)}$
1	9.95	1	9.95×10^4	4.95×10^9	-48.5	0
2	1	10^6	4.995×10^8	1.25×10^{17}	10^6	$+\infty$
3	0.3	1.04	0.5	1.12	0.995	$+\infty$
4	2.78	1.08	2.79	1.11	-2.77	0
5	0.18	1.03	0.4	1.09	1.01	$+\infty$

LL-ASD in predicting the dominance of stocks over bonds in the long run for Examples 1 and 4. However, the implications of the two models deviate for Examples 2, 3 and 5, where LL-ASD still predicts the dominance of stocks over bonds and our model does not. In particular, Table 4.1 confirms the finding of Levy (2009) that any fixed utility function which for a given n favours $X^{(n)}$ to $Y^{(n)}$ in Example 3, will be likely to lose this property as n increases.

4.5 Concluding remarks

We have introduced an alternative to LL-Almost Stochastic Dominance proposed by Leshno and Levy (2002), which we call ε-Almost Stochastic Dominance. Both concepts can be seen as relaxations of standard Stochastic Dominance, due to Leshno and Levy (2002, Prop. 3) and Proposition 4.1 in this Chapter. The major advantage of the proposed ε-ASD is its computational tractability: both testing for the ε-ASD-efficiency of a given portfolio relative to an infinite portfolio possibilities set and identifying a marketed portfolio that ε-ASD-dominates a given benchmark (in the case where the benchmark portfolio is ε-ASD-inefficient) can be done by solving a simple linear programming problem such as (4.9), (4.10) or (4.11). For comparison: identifying a marketed portfolio that LL-ASD-dominates a given benchmark would require solving a non-convex optimization problem with a fractional objective function which is extremely difficult to handle on real-life data sets.

Besides being computationally tractable, ε-ASD performs well on all the intuitive examples considered in this Chapter, and in the case of Example 2, when an extreme loss may occur, ε-ASD appears to model preferences in a more realistic way than LL-ASD does.

Finally, ε-ASD can contribute to the ongoing debate about investors' preferences between stocks and bonds in the long run by looking at the dynamics of the ε-threshold at which stocks ε-ASD dominate bonds over time. This idea is similar to the analysis of Leshno and Levy (2002). However, ε-ASD is more consistent with standard Stochastic Dominance, since ε-ASD accounts for the whole area between cumulative distribution functions that violates SD. On the contrary, LL-ASD implies that any risky violation of SD can be compensated by an increase in upside potential in a *linear* fashion. This linear relation for all levels of losses may be difficult to accept in practice. Moreover, in contrast to LL-ASD, the ε-threshold in ε-ASD has a natural economic interpretation: it represents the smallest value of the mean return of a random variable that needs to be added to a portfolio in order for it to dominate a given benchmark.

Chapter 5

Stochastic Dominance and Full- and Partial Moments in Dynamic Asset Allocation.

We analyze a novel periodic asset allocation strategy based on the Second order Stochastic Dominance (SSD) efficiency and compare its performance with other strategies, such as Lower Partial Moments, Mean-Variance, Momentum, Value, Alpha, Beta and passive investing on French's 48 industry portfolios. We observe that the SSD strategy performs reasonably well in terms of realized return and indicates an out-of-sample persistence in restricting portfolio risk. Moreover, expanding the results of Grootveld and Hallerbach (1999), we find a substantial difference in efficient portfolios formed on the basis of downside risk criteria (linear lower partial moments and semideviation) and full moments (mean-variance), particularly in cases where short sales restrictions are relaxed.

5.1 Introduction

A trade-off between risk and expected return is a starting point in any sensible portfolio selection model, for researchers as well as practitioners. The mean-variance portfolio selection model introduced by Markowitz (1952), which defines risk in terms of standard deviation, has received mixed reactions ever since it was published. On the one hand, it is nearly the only model for which the whole set of efficient portfolios can be expressed in a closed form; it is definitely the easiest model to compute, even for a large number of underlying assets (see Hazetl and Norton, 1986); it has been widely used to date (Clarke et al., 2006); and it is consistent with the capital asset pricing model (see Cochrane, 2005). On the other hand, the model has many theoretical shortcomings (see Meyer (1987) and references therein). Its main drawback, as noted by Markowitz (1959), Mao (1970) and Fishburn (1977) among others, is that variance as a measure of risk penalizes downside losses just as much as it does upside gains, which is clearly counter-intuitive.

To overcome this, several downside measures of risk have been proposed, such as semi-variance (Markowitz, 1952, Mao, 1970) and Gini means (Shalit and Yitzhaki, 1984). Fishburn (1977) proposed a mean-downside risk model, substituting the variance in the Markowitz model by lower partial moments. Holthausen (1981) further refined the Fishburn model by taking weighted deviations above the target instead of the mean return, thereby combining downside risk with upside potential in his model. Kang et al. (1996) further generalized this approach in their mean-separated target deviations risk model where the separated measure of risk represents a parameterized difference between downside risk and upside potential.

In this Chapter we shall introduce another investment strategy related to restricting downside risk. This strategy is based on the concept of Second order Stochastic Dominance (see Hadar and Russell, 1969, Hanoch and Levy, 1969), further SSD, and it periodically selects a portfolio that SSD dominates a given benchmark. The method for identifying such a portfolio was introduced and studied as a one-period portfolio optimization model by Dentcheva and Ruszczyński (2003) and Kuosmanen (2004). However, to our knowledge, the method has not yet been duly investigated in a multi-period setting on a realistic data set of asset returns. In this Chapter we employ this idea for constructing a dynamic asset allocation strategy with monthly rebalancing of the investment portfolio and apply a more efficient computational algorithm proposed in Kopa and Chovanec (2008) for computing such a portfolio.

Furthermore, we compare the SSD-based strategy with other mean-risk investing strategies which employ various definitions of risk such as lower partial and full moments. Extending the study of Grootveld and Hallerbach (1999), we find a substantial difference in efficient frontiers formed based on downside risk criteria and full moments (mean-variance), particularly in cases when short sales restrictions are relaxed, but also without short sales, when expected returns are not too high. We argue that the similarity between the efficient frontiers reported in Grootveld and Hallerbach (1999) was caused by assuming no short sales and a low number of underlying assets rather than a general closeness of those frontiers.

Next to comparing various efficient frontiers of portfolios formed ex post on the basis of the historical distribution of asset returns, we study their out-of-sample performance relative to the new SD-based rebalancing, as well as some other strategies proposed in the literature such as momentum (Jegadeesh and Titman, 1993, 2001), volatility (Blitz and van Vliet, 2007), and CAPM-based betas and alphas rebalancing (inspired by Post, van Vliet, and Lansdorp, 2009). We observe that the new SD-based method performs differently from the conventional strategies (over- or underperforming those at different times and different market circumstances), but is in no way inferior to them. After performing a sensitivity and robustness check we conclude that the new method could be used as a benchmark strategy for practitioners and as a challenging research subject for scholars.

5.2 Data and Methodology

To proxy the investment universe that private and small or medium institutional investors are facing, we use monthly returns of 48 value weighted industry portfolios aggregated over different sectors running from January 1970 to December 2009. This data set is reasonably large and neither too aggregated (when no diversification strategies can be exploited) nor too dispersed (as it would be in the case of individual stocks, when statistical properties are difficult to infer in a robust way due to pronounced non-stationarity effects). The data used in this Chapter, including the market and risk-free return, can be downloaded from French (2010) but is also freely available from the authors upon request.

As any desired level of risk can be achieved in the presence of a risk-free asset without short-sale constraints by leveraging (and the riskless asset is available in our data), we will consider the 48 risky portfolios apart from the risk-free asset. One of our goals is to investigate how realistic and persistent different portfolio risk measures are. Obviously, if a strategy with a relatively stable risk exposure can be found for risky portfolios, the stability of the risk measure will not be affected by leveraging this portfolio with a risk-free asset, which will allow compensating the level of risk by adjusting the mean return. For that reason we will consider portfolios that minimize the risk according to a particular criterion. Again, we consider minimum-risk, rather than tangency portfolios in this study, because our primary goal is to examine the persistence of different risk criteria, and minimum-risk portfolios are more robust to data perturbations than tangency portfolios. In the following section, however, we will compare the realized returns of various strategies. Still, we aim to study the differences of the realized returns on moments-based strategies and their statistical properties rather than the absolute value of the returns. Minimum-risk portfolios are more likely to expose the difference in performance associated with the various definitions of portfolio risk, avoiding the extra source of numerical imprecision associated with the influence of the changing risk-free rate on the tangency portfolio.

Below we describe the investment strategies that will be analyzed in the Chapter.

5.2.1 Full vs. partial moments: MV-, LPM- and SDV-efficient portfolios.

No study of investment strategies can avoid including mean-variance efficient portfolios. Inspired by Grootveld and Hallerbach (1999), below we compare the mean-variance (MV), linear lower partial moments (LPM) and semi-deviation (SDV) efficient frontiers. All partial moments were computed with the threshold equal to the risk-free rate (averaged over the last 10 years). Figures 5.1, 5.2 and 5.3 present those frontiers (portfolios minimizing the corresponding risk subject to having the mean return larger than a given threshold) in mean-semi-deviation, mean-LPM and mean-standard deviation spaces. For each definition of risk we consider two cases: firstly when short sales are unrestricted, and when they are prohibited.

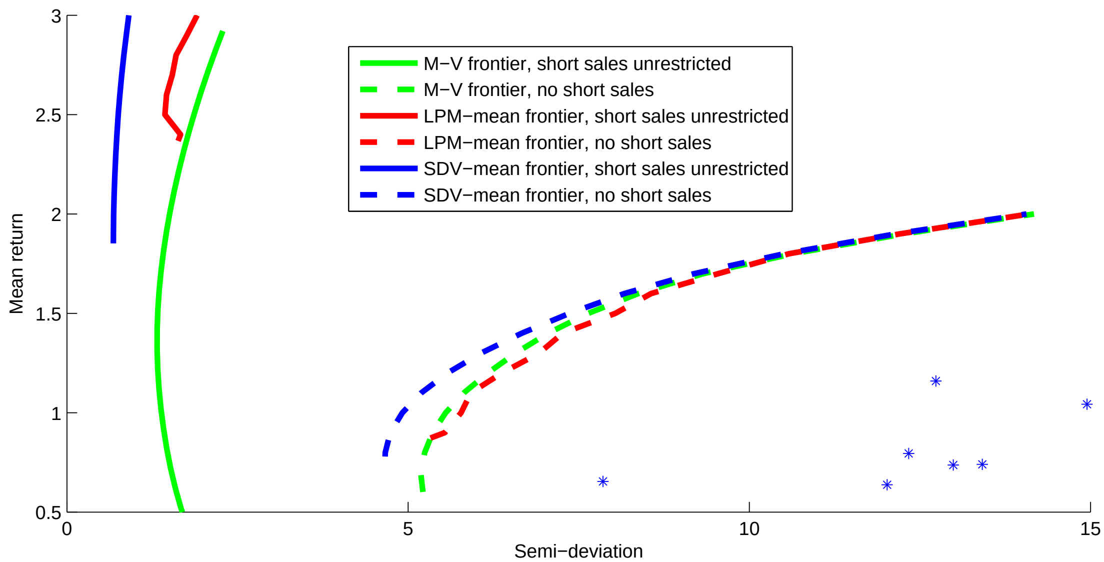


Figure 5.1: Efficient frontiers in Mean-Semideviation space.

As can be seen from Figures 5.1, 5.2 and 5.3, the difference between efficient frontiers is quite substantial, particularly when short sales are unrestricted, but also when they are restricted, for moderate mean returns. We have repeated this analysis on different subsamples and different values of the threshold, and observed similar results. The difference can clearly be seen in the Figures, and is also supported by statistical tests. For instance, the null hypothesis of semideviation- and MV-efficient long-only portfolios being equal component-wise is rejected at 1% significance level by the T-test (the implied dependence between the portfolio weights was ignored in this T-test).

This result extends and partly confirms the findings of Grootveld and Hallerbach (1999), who assumed away short sales and considered only six underlying assets. Our investment universe consists of 48 assets that are in addition more volatile and less correlated with each other than those used in Grootveld and Hallerbach (1999). We can therefore give a positive answer to the main question posed in that publication: there is a substantial difference

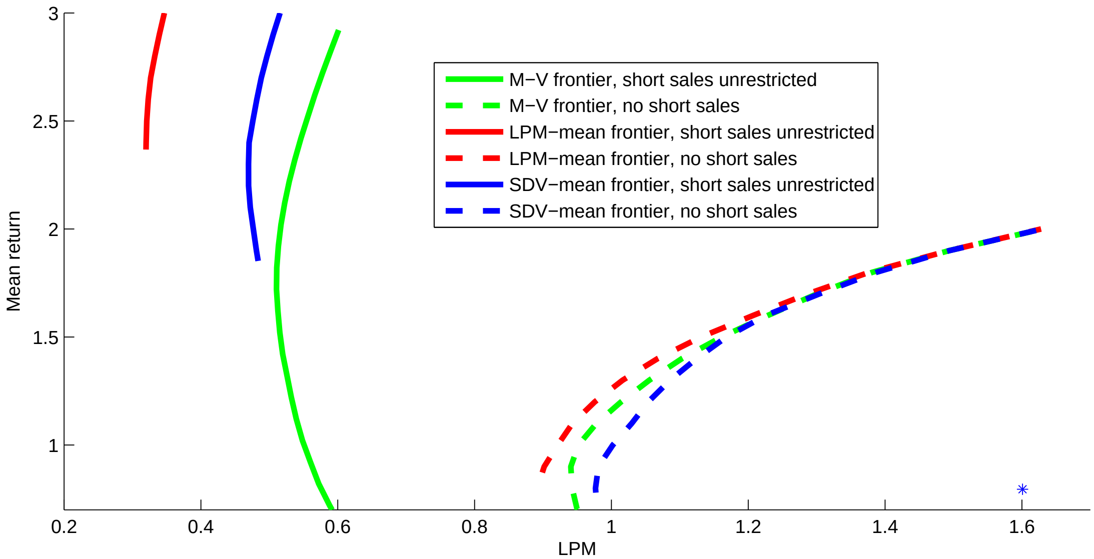


Figure 5.2: Efficient frontiers in Mean-LPM space.

between mean-variance and downside risk efficient frontiers, at least when the investment universe is large enough and short sales are (partly) allowed.

We will henceforth focus on investment strategies that minimize the corresponding risks (MV, SDV and LPM) both with and without short sales restrictions.

5.2.2 Second Order Stochastic Dominance (SSD) strategy.

The Stochastic Dominance algorithms reviewed in Chapter 2 can be seen as other portfolio selection criteria that restrict the downside risk. In the area of finance, SSD has been broadly employed for analyzing the efficiency of a given portfolio, see for example Post (2003), Kuosmanen (2004) and Post and Versijp (2007). However, we need more than an efficiency test alone for constructing a periodic rebalancing strategy based on the SSD criterion. At the end of each period we need to identify a marketed portfolio that SSD dominates a given benchmark and is SSD efficient itself. Portfolio x dominates portfolio y by SSD if all risk-averse and non-satiable investors are better off holding x than y in terms of their expected utilities. A portfo-

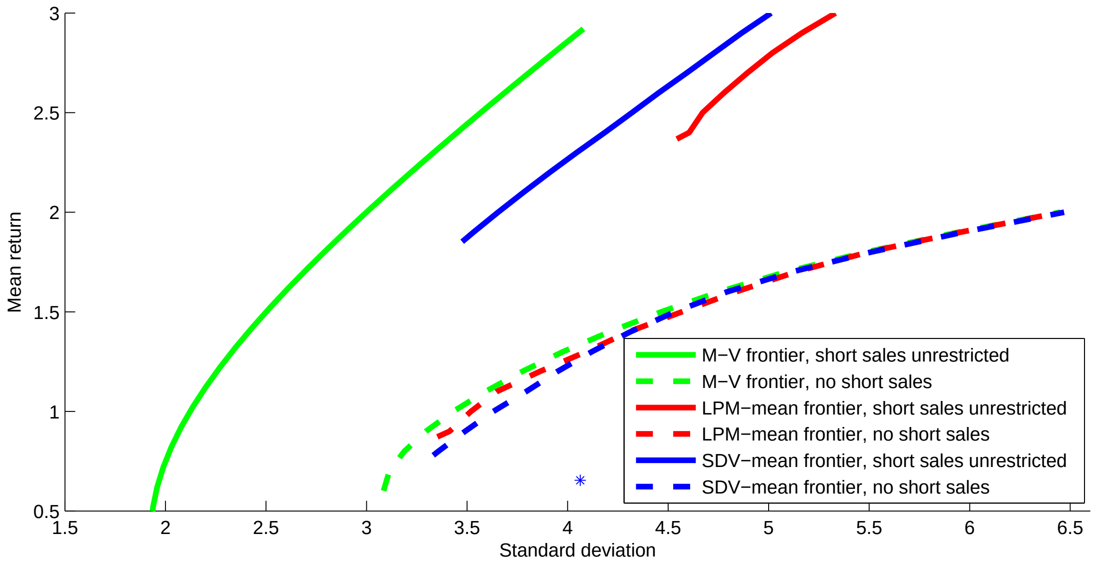


Figure 5.3: Efficient frontiers in Mean-Standard Deviation space.

lio is SSD efficient if it is not dominated by any other marketed portfolio. As shown in Chapter 2, this can be done by two-stage modified Dentcheva and Ruszczyński (2003) method, Kopa and Chovanec (2008) or Kuosmanen (2004) necessary test combined with a quadratic programming sufficiency test. Since the dimensionality of real-life data can be computationally prohibitive (see Dentcheva and Ruszczyński, 2006b), we apply the sufficiency test proposed in Kopa and Chovanec (2008) as follows. Given a benchmark portfolio y and the m-by-n state-space tableau X of asset returns (assuming X to be a full-rank matrix, which is equivalent to all asset returns being linearly independent, which in itself is a necessary condition for the absence of arbitrage), determine portfolio $\lambda \in \Lambda = \left\{\tau \in \mathbb{R}^n : \tau_i \geq 0, \quad \tau^\mathsf{T} e = 1\right\}$ by solving the following program (after Kopa and Chovanec, 2008):

$$d^* = \max_{\lambda \in \Lambda} \sum_{k=0}^{m-1} \sum_{j=1}^{n} \lambda_j a_{jk} \tag{5.1}$$

$$\text{s.t.} \quad \sum_{j=1}^{n} \lambda_j a_{jk} \geq 0, \quad k = 0, \ldots, m-1, \text{ with } a_{jk} = \sum_{i=1}^{k} x_j^{(i)} - \sum_{i=1}^{k} y^{(i)}.$$

If $d^* > 0$ in (5.1), then y is SSD inefficient and λ^* is an SSD efficient portfolio that dominates y (see (2.24) and (2.25) in Chapter 2 for more explanation). We shall use λ^* as the SSD portfolio at each portfolio rebalancing period. However, should (5.1) happen to be infeasible, no conclusion can be made regarding the efficiency of y (see Chapter 2). In this case, we will have to solve the following program (after Kuosmanen, 2004) which is more computationally demanding than (5.1):

$$\begin{aligned} \theta_2^N(y) &= \max_{\lambda, W} (X\lambda - y)^\mathsf{T} e \\ \text{s.t. } X\lambda &\geq Wy \\ W &\in \Xi \\ \lambda &\in \Lambda, \end{aligned} \tag{5.2}$$

where Ξ is the class of doubly stochastic matrices given in Definition 2.7.

Portfolio λ^* in (5.2) is weakly SSD efficient and will be used for our SSD strategy each time (5.1) is infeasible.

Naturally, we take the market portfolio as a benchmark and denote the SSD efficient portfolio that dominates it by λ_{SSD}. Dentcheva and Ruszczyński (2003) introduced a similar approach of maximizing mean return subject to stochastic dominance constraints. The difference between that and the current study is that we are applying this strategy recursively and analyzing its realized performance over multiple consecutive periods. Similarly to the LPM strategy, the SSD strategy implicitly incorporates two different criteria: firstly, restricting the feasible set to SSD dominating portfolios ensures that the downside risk will not exceed that of the benchmark; secondly, requiring the dominating portfolio to be SSD efficient can be seen as improving upon its upside potential subject to the acceptable level of downside risk.

5.2.3 Other strategies considered.

Momentum strategies. The momentum strategy, first introduced by Jegadeesh and Titman (1993) is also included in our analysis. It is based solely on return data, where portfolios are rebalanced monthly on the basis of their past performance. We follow Blitz and van Vliet (2008) in exploring a 12–1 month momentum strategy and a strategy based on returns over one previous month.

Volatility-based strategy. Another strategy we consider is based on the volatility effect reported in Blitz and van Vliet (2007), where portfolios rebalanced on the past 3-year volatility have decreasing Sharpe ratios when sorted by volatility and a negative slope in the mean-standard deviation space (in both cases values of the mean returns and standard deviations are calculated ex post for the portfolios obtained by periodic rebalancing). Similar ex post analysis performed on our data and presented in Figure 5.4 confirms the negative implied slope of the security market line, which means that this strategy can be used on our data.

Beta strategy. Inspired by Post, van Vliet, and Lansdorp (2009), we would have liked to analyze rebalancing strategies based on various downside betas proposed in that paper, namely: regular-, semivariance- (Bawa and Lindenberg, 1977), ARM- (Harlow and Rao, 1989) and downside covariance beta (Ang et al., 2006), defined for return R_i in the following way:

$$\beta_{Reg,i} = \frac{cov(R_i, R_M)}{var(R_M)}, \qquad \beta_{SV,i} = \frac{\mathbb{E}(R_i R_M | R_M \leq 0)}{\mathbb{E}(R_M^2 | R_M \leq 0)}, \tag{5.3}$$

$$\beta_{ARM,i} = \frac{cov(Z, R_M)}{var(Z)}, \qquad \beta_{DC,i} = \frac{cov(R_i, R_M | R_M \leq 0)}{var(R_M | R_M \leq 0)}, \tag{5.4}$$

with R_M being the market return and $Z \equiv R_M \mathbb{1}_{\{R_m \leq 0\}} + \mathbb{E}(R_M | R_M > 0)\mathbb{1}_{\{R_m > 0\}}$, (see Post, van Vliet, and Lansdorp (2009) for a discussion and examples). For that reason, we have constructed portfolios by sorting the industry funds in our sample by those betas on a monthly basis. Table 5.1 reports the realized return of those portfolios with and without grouping the industry funds into decile portfolios (for non-grouped 48 funds, only the returns corresponding to the highest and lowest beta are reported), along with its statistics. The results for our data set coincide with the conclusions of Post, van Vliet, and Lansdorp (2009) in that the return spread of decile portfolios formed on regular beta is lower than the spread of semivariance- and ARM beta portfolios but higher than the downside covariance spread. However, our results somewhat deviate from those of Post, van Vliet, and Lansdorp (2009) regarding the magnitude of the returns: in our case all decile beta spreads are rather too low to be employed as an investment strategy. Moreover, without the decile grouping, the regular beta spread turned out the largest (=0.57) and more than doubled the semideviation spread (=0.24). Therefore, we will only analyze the rebalancing strategy based on regular beta without grouping the funds into decile portfolios.

Alphas strategy. Assuming that empirical implications of the capital asset pricing model (see Cochrane, 2005) hold at least in part, one may employ the strategy of rebalancing the portfolio by taking short position in overpriced and long in underpriced in terms of alphas assets. We analyze this strategy as well.

Naive passive strategies. Challenged by Frankfurter et al. (1971) we also consider a naive buy-and-hold equally-weighted strategy without rebalancing and the strategy of investing in the market portfolio.

For active SSD, full- and partial moments strategies we consider two options: when short sales are allowed and unrestricted and when short sales are prohibited. Momentum, volatility, beta- and alpha-based strategies automatically require the possibility of shortselling. All active strategies are rebalanced monthly.

5.3 Realized performance.

In this section we will analyze the realized out-of-sample return and risk of the monthly rebalancing strategies described above.

5.3.1 Realized return

Figure 5.5 shows the cumulative return of the strategies described in the previous section. In all strategies we rebalance portfolios monthly. In LPM, MV, SDV long-only, as well as SSD, alpha and beta strategies, the 60 months preceding the decision time were used for estimating parameters, whereas LPM, MV and SDV efficient portfolios without shortsale restrictions imposed

Table 5.1: Performance of Beta strategies

	Regular Beta:													
	Decile portfolios											48 industry portfolios		
	Low	2	3	4	5	6	7	8	9	High	HminL	Low	High	HminL
β_{REG}	0.56	0.75	0.84	0.91	0.98	1.05	1.12	1.19	1.30	1.53	0.97	0.27	1.68	1.41
Mean return	1.20	1.06	1.16	1.24	1.06	1.23	1.26	1.21	1.26	1.19	0.00	0.92	1.49	0.57
StD	4.08	4.57	4.72	4.68	5.06	5.47	5.52	5.86	6.42	7.13		5.96	9.31	
Skewness	-0.67	-0.37	-1.06	-0.68	-0.46	-0.46	-0.52	-0.30	-0.16	-0.49		-0.16	-0.18	
Kurtosis	6.52	5.41	8.14	5.71	5.87	8.14	5.64	5.99	6.82	5.95		4.27	5.21	
	Semivariance Beta:													
	Low	2	3	4	5	6	7	8	9	High	HminL	Low	High	HminL
β_{SV}	0.45	0.65	0.79	0.88	0.98	1.06	1.13	1.23	1.35	1.63	1.17	0.12	1.81	1.69
Mean return	1.10	1.27	1.17	1.21	1.24	1.32	1.19	1.12	1.02	1.19	0.09	0.92	1.16	0.24
StD	4.22	4.58	4.56	4.81	5.37	5.47	5.36	5.77	5.95	7.28		6.37	9.56	
Skewness	-0.55	-0.64	-0.67	-0.50	-0.68	-0.52	-0.34	-0.35	-0.63	0.32		-0.86	1.11	
Kurtosis	4.81	6.33	5.55	6.07	5.90	6.69	6.32	5.61	6.62	8.13		9.24	14.27	
	ARM Beta:													
	Low	2	3	4	5	6	7	8	9	High	HminL	Low	High	HminL
β_{ARM}	0.57	0.75	0.84	0.92	1.00	1.06	1.13	1.20	1.30	1.54	0.97	0.27	1.68	1.41
Mean return	1.18	1.27	0.93	1.15	1.13	1.25	1.11	1.22	1.34	1.28	0.10	0.81	1.37	0.55
StD	4.10	4.50	4.94	4.66	4.99	5.36	5.63	5.87	6.30	7.22		6.03	9.03	
Skewness	-0.78	-0.27	-0.66	-0.66	-0.66	-0.45	-0.68	-0.44	-0.39	-0.17		-0.26	-0.28	
Kurtosis	6.62	4.82	5.83	6.80	6.16	5.63	7.09	5.66	6.21	6.37		4.63	4.92	
	Downside Covariance Beta:													
	Low	2	3	4	5	6	7	8	9	High	HminL	Low	High	HminL
β_{DC}	0.41	0.61	0.75	0.86	0.96	1.04	1.13	1.22	1.37	1.71	1.30	0.07	1.96	1.89
Mean return	1.35	0.97	1.28	1.22	1.26	1.17	1.22	1.17	1.08	1.23	-0.12	0.80	1.31	0.51
Std	4.65	4.52	4.77	4.95	5.13	5.33	5.37	5.60	6.14	6.82		6.75	8.37	
Skewness	-0.58	-0.88	-0.83	-0.63	-0.48	-0.75	-0.32	-0.19	-0.26	0.30		-0.61	0.00	
Kurtosis	5.87	7.06	7.06	6.29	6.47	6.66	5.59	5.41	5.84	7.89		9.05	6.42	

were computed based on the preceding 120 months. To make the performance of zero-net-investment strategies (such as volatility, betas or momentum) comparable with strategies where a positive net investment is present (SSD, MV, partial moments) we adjust the latter by subtracting the risk-free rate in each period. For the same reason we plot cumulative returns instead of cumulative product returns.

One can observe from Figure 5.5 that the SSD strategy is competitive relative to other strategies, over- and under-performing those in different times and different market circumstances. Table 5.2 lists the average, standard deviation, minimum and maximum values, skewness and kurtosis of the realized returns, from which we can see that SSD slightly outperformed the two momentum strategies. In addition, only the partial moment strategies with unrestricted short sales ended up having higher realized mean returns than those of SSD, but those strategies also turned out to be more volatile. Although the differences between the mean returns are not high enough to be statistically significant at 5% confidence level, we do observe them persistently while modifying the data set. Robustness and sensitivity analysis will be addressed in more detail in section 5.4.

Another interesting observation can be drawn from Table 5.3 which shows estimated pairwise correlation coefficients between realized returns. We have have observed all kinds of correlations: negative (momentum-1 and long-only mean-variance strategies), close-to-zero (momentum-1 and semideviation) and close-to-1 (the equally-weighted and market portfolios), which fact seem to indicate that the set of strategies considered is quite diverse. Note also that with short sales restricted, MV, LPM and SDV strategies are highly positively correlated, whereas if short sales are allowed, only the partial moments strategies, LPM and SDV, are highly correlated, but the correlations between them and the MV return substantially decrease. This confirms our findings in section 5.2.1 related to Grootveld and Hallerbach (1999).

5.3.2 Realized risk.

We will use two indicators for assessing the stability in the performance of each strategy. First, we employ an analogue to the root mean squared

Table 5.2: Performance statistics of asset allocation strategies.

	Mean	St. Dev.	Min	Max	Skewness	Kurtosis
SSD, net of R_f	0.80	5.11	-24.02	16.14	-0.46	5.00
Market, net of R_f	0.59	4.58	-23.14	13.58	-0.72	5.42
EW, net of R_f	0.72	4.86	-26.67	17.74	-0.68	6.78
MV, net of R_f	0.41	4.22	-16.11	17.26	-0.14	4.54
MV, l/o, net of R_f	0.54	3.51	-16.14	12.29	-0.54	5.61
SDV, net of R_f	0.98	10.22	-93.00	44.63	-1.69	23.49
SDV, l/o, net of R_f	0.44	3.78	-18.81	13.28	-0.76	6.14
LPM, net of R_f	1.14	11.18	-90.94	43.88	-1.06	16.48
LPM, l/o, net of R_f	0.52	3.73	-19.79	12.31	-0.75	5.62
Momentum 1	0.73	5.17	-20.22	23.80	0.04	4.87
Momentum 12-1	0.72	6.38	-49.06	21.91	-1.04	12.05
Regular Beta	0.57	10.24	-37.93	33.97	-0.03	4.48
Alpha	0.39	4.90	-30.56	17.88	-0.81	7.48
Volatility	0.36	11.57	-36.60	70.28	0.49	5.99

dispersion index ($RMSDI$) used in Grootveld and Hallerbach (1999). We define $RMSDI$ as the squared deviation between two optimal portfolios λ_t and λ_{t+1} at time t and $t+1$ averaged over time:

$$RMSDI(\lambda) = \frac{1}{\sqrt{(T-1)n}} \sqrt{\sum_{t=1}^{T-1} \sum_{i=1}^{n} \left(\lambda_{ti} - \lambda_{(t+1)i}\right)^2}. \tag{5.5}$$

For a given asset allocation strategy λ, $RSMDI(\lambda)$ will indicate how close optimal portfolios corresponding to this strategy are to one another at two consecutive periods. The lower the $RSMDI(\lambda)$, the more stable strategy λ is and the less effort is required for portfolio rebalancing for that strategy. The strategies resulted in the following values of $RMSDI$:

Table 5.3: Correlation coefficients between strategies (pairwise).

	SSD	R_M	EV	MV	MV_{LO}	SDV	SDV_{LO}	LPM	LPM_{LO}	M_1	M_{12-1}	β_{reg}	α	Vol
SSD	1.00	0.74	0.69	0.50	0.71	0.37	0.77	0.38	0.79	-0.06	0.14	0.30	0.23	0.11
R_M	0.74	1.00	0.96	0.53	0.79	0.27	0.80	0.29	0.84	-0.11	-0.07	0.47	0.02	0.22
EV	0.69	0.96	1.00	0.53	0.81	0.20	0.80	0.21	0.83	-0.14	-0.11	0.40	-0.08	0.26
MV	0.50	0.53	0.53	1.00	0.72	0.39	0.68	0.36	0.66	-0.11	-0.03	0.06	0.03	0.10
MV_{LO}	0.71	0.79	0.81	0.72	1.00	0.26	0.96	0.25	0.93	-0.17	-0.11	0.11	-0.03	0.12
SDV	0.37	0.27	0.20	0.39	0.26	1.00	0.31	0.96	0.32	-0.02	0.20	0.11	0.28	-0.04
SDV_{LO}	0.77	0.80	0.80	0.68	0.96	0.31	1.00	0.30	0.95	-0.15	-0.04	0.13	0.07	0.14
LPM	0.38	0.29	0.21	0.36	0.25	0.96	0.30	1.00	0.32	-0.01	0.23	0.16	0.30	-0.01
LPM_{LO}	0.79	0.84	0.83	0.66	0.93	0.32	0.95	0.32	1.00	-0.11	0.01	0.16	0.12	0.19
M_1	-0.06	-0.11	-0.14	-0.11	-0.17	-0.02	-0.15	-0.01	-0.11	1.00	0.15	0.04	0.20	-0.05
M_{12-1}	0.14	-0.07	-0.11	-0.03	-0.11	0.20	-0.04	0.23	0.01	0.15	1.00	-0.08	0.62	0.07
β_{reg}	0.30	0.47	0.40	0.06	0.11	0.11	0.13	0.16	0.16	0.04	-0.08	1.00	0.03	0.13
α	0.23	0.02	-0.08	0.03	-0.03	0.28	0.07	0.30	0.12	0.20	0.62	0.03	1.00	0.08
Vol	0.11	0.22	0.26	0.10	0.12	-0.04	0.14	-0.01	0.19	-0.05	0.07	0.13	0.08	1.00

$$\begin{array}{rlrlrl} & & RMSDI(\lambda) \text{ for} & & & \\ \text{SSD}: & 0.065, & \text{MV}: & 0.024, & \text{MV}_{LO}: & 0.011, \\ \text{SDV}: & 0.316, & \text{SDV}_{LO}: & 0.018, & \text{LPM}: & 0.309, \\ \text{LPM}_{LO}: & 0.022, & \text{M}_1: & 0.128, & \text{M}_{12-1}: & 0.067, \\ \beta_{Reg}: & 0.094, & \alpha: & 0.045, & \text{Vol}: & 0.056. \end{array} \tag{5.6}$$

We can see from (5.6) that the most stable strategy in terms of $RMSDI$ is mean-variance, which was to be expected: the variance as a conservative measure of risk leaves little room for variation in portfolio weights, compared to partial moments. The difference in $RMSDI$ between full and partial moments is especially pronounced when short sales are allowed. The most volatile strategies turned out to be LPM and SDV without short sales, which again is not surprising. A much less predictable observation is a relatively good performance of the SSD strategy. Whether this conclusion is persistent to data perturbations will be studied in the next section.

Apart from $RMSDI$, we can use a more straightforward performance indicator of realized portfolio risk, such as the average ratio of in-sample to realized risk corresponding to each strategy. For instance, for the LPM strategy this indicator leads to the ratio of ex-post to ex-ante linear partial moments with the reference point equal to the risk-free rate at the time of rebalancing. For the MV strategy it can be the ratio of portfolio variances before and after each period. We denote those ratios by r_{MV}, r_{LPM}, and r_{SDV}. By construction, r_{MV}, r_{LPM}, and r_{SDV} are greater than 1 and indicate the factor by which the realized risk exceeds the in-sample risk.

Since SSD strategy does not explicitly utilize full- or partial moments as portfolio risk, we will use the concept of ε-Almost Stochastic Dominance (ε-ASSD) recently studied in Lizyayev and Ruszczyński (2010) as the performance indicator. More precisely, we will use the average threshold at which λ_{SSD} ε-ASSD dominates the "out-of-sample" market portfolio. This threshold for portfolio Z relative to the market portfolio R_M is defined as

$$\varepsilon(Z) = \max_{\eta} \left\{ F_Z^{(2)}(\eta) - F_{R_M}^{(2)}(\eta) \right\}. \tag{5.7}$$

If Z SSD-dominates R_M, then $\varepsilon(Z) = 0$; otherwise, $\varepsilon(Z) > 0$ by construction. We will use $\varepsilon(Z)$ averaged over all periods as a stability indicator of the SSD strategy. The values of r_{MV}, r_{LPM}, r_{SDV} and ε are listed in (5.8).

$$\begin{array}{lll} r_{MV} = 1.011, & r_{LPM} = 1.238, & r_{SDV} = 11.439, \\ r_{MV_{LO}} = 1.004, & r_{LPM_{LO}} = 1.013, & r_{SDV_{LO}} = 1.047, \\ & \varepsilon_{ASSD} = 0.0046. & \end{array} \tag{5.8}$$

The figures in (5.8) concur with (5.6) in stability ranking: SDV strategy appears to be the most volatile, MV – the most stable, and LPM was ranked in between the two. Although the ε_{ASSD}-threshold reported in (5.8) for SSD strategy is not directly comparable with the realized risk ratios for full and partial moments strategies, the average value of ε_{ASSD}-threshold below 0.5% certainly indicates a very stable performance of the SSD strategy in terms of its realized risk. In the next section we will examine how robust these conclusions are with respect to data perturbations.

5.4 Robustness check: Bootstrapping.

To check whether the conclusions regarding the performance of the strategies are robust to changes in the data, we will conduct a robustness test using bootstrapping by permutations in the following way. First, 100 samples of 90 and 180 consecutive monthly observations were randomly drawn from the original data set. The 90-month samples were used for SSD, long-only moments, alpha, beta and momentum strategies, and the 180-month samples for moments strategies with shortselling. Subsequently, 60 and 120 monthly observations (not necessarily consecutive) were drawn (with replacement) from the 90- and 180-month samples, respectively.

Optimal portfolios corresponding to each strategy were computed based on those sub-samples. Further, 20 and 40 monthly observations were randomly drawn out of the remaining 30 and 60 months in each sample, to be used as the "out-of sample" part for testing the realized performance. In this manner we can ensure that the calibration sample is always three times as large as the evaluation sample and that the returns in both sub-samples have occurred within a reasonable time of each other. Moreover, this bootstrapping procedure preserves the original dependency structure among the returns.

The sub-sampling procedure was carried out $n = 1000$ times for each subsample, which brings the total number of scenarios to $100,000$. In each such scenario, the out-of-sample returns were averaged for each strategy. Table 5.4 shows the descriptive statistics for the realized out-of-sample returns. Again, the SSD strategy performed relatively well in terms of the Sharpe ratio. A general observation can be made that the long-only strategies have shown more stable performance than those with short sales, although the unrestricted MV strategy performed reasonably well, too. LPM rebalancing turned out to be less volatile than SDV.

The average ε_{ASSD}-threshold was 0.3% (min = 0, max = 38%, STD = 1.2%) and in 81.9% of the scenarios ε_{ASSD} was zero, which again indicates a stable performance of the strategy.

5.5 Concluding remarks

The major goal of this Chapter was to employ the concept of Stochastic Dominance for periodic asset allocation strategy and to investigate its performance relative to other popular strategies on a realistic data set. Having used the returns of 48 industry portfolios running from January 1970 to December 2009, we are pleased to report that the SSD strategy is rather competitive relative to other strategies, in terms of both the realized return and the realized risk. Moreover, the SSD strategy is comparable to momentum investing as far as the out-of-sample return is concerned. This observation is remarkable, given that the SSD strategy does not utilize the possibility of shortselling, unlike the momentum strategies. Besides, the 1 and 12-1 months momentum strategies have been reported to perform well in the investment management literature. The stability and relative competitiveness of the SSD strategy in

Table 5.4: Performance statistics of asset allocation strategies with bootstrapping.

	Average	St. Dev.	Sharpe	Skewness	Kurtosis
SSD:	0.499	4.674	0.107	-0.427	5.839
MV:	0.445	4.080	0.109	0.013	4.341
MV_{LO} :	0.536	3.537	0.151	-0.317	5.440
SDV:	0.149	264.680	0.001	-45.685	35405.6
SDV_{LO} :	0.470	3.745	0.126	-0.530	6.130
LPM:	0.385	17.608	0.022	-11.114	3970.2
LPM_{LO} :	0.493	3.779	0.130	-0.520	6.135
M1:	0.044	4.315	0.010	-0.043	5.048
M12-1:	0.103	4.440	0.023	-0.120	5.125
Breg:	0.241	10.321	0.023	-0.176	4.768
Alpha:	0.276	4.489	0.062	-0.257	4.609
Vol:	0.033	11.117	0.003	0.340	4.031

terms of the realized return and as a measure of portfolio risk has withstood the robustness check performed via bootstrapping of the observations.

Needless to say, we can not claim the generality of the result on each and every data set, but having drawn interesting conclusions concerning the performance of the SSD strategy on the 48 value-weighted industry portfolios from French (2010) which are used by many other scholars, we would like to encourage researchers to continue investigating the SSD strategy. And if our results are confirmed on other data sets and other robustness checks, the strategy could definitely be recommended as an alternative benchmark for investment managers in practice. Unfortunately, to our knowledge, Stochastic Dominance has not so far been duly researched in dynamic multi-period settings on a realistic data set. We believe that Stochastic Dominance has a lot of practical potential that has not been fully utilized to date. Theoretical derivations related to Stochastic Dominance may seem almost incomprehensible to practitioners, but in fact implementing the proposed strategy is not difficult at all: the corresponding algorithms reviewed in Lizyayev (2010) are standard linear and quadratic optimization programs that can be solved by numerous available optimization software packages.

Next to the Stochastic Dominance-related results, we were able to come to another conclusion with regard to the difference between portfolios formed based on full and partial moments as a measure of risk. Grootveld and Hallerbach (1999) have reported only a marginal difference between full and partial moments efficient frontiers and questioned whether this holds in general. This study leads us to the conclusion that the similarity of two frontiers are data dependent and in our comprehensive data set significant differences between the frontiers result. We have thereby expanded the results of Grootveld and Hallerbach (1999). Certainly, in some cases the difference may diminish, for example if the investment universe is rather limited. For instance, by assuming away short sales in Figures 5.1, 5.2 and 5.3 and restricting the mean return to exceed, say, 1.9, hardly any difference will be observed between the portfolios. However, this is due to the portfolio possibilities set being shrunk (there are very few portfolios that satisfy the mean return restrictions in the absence of short sales), rather than a general similarity between the full and partial moments criteria (note, that Grootveld and Hallerbach (1999) use a data set consisting of only 6 assets, 3 of which are bonds, and assume away short sales, whereas we use 48 stock returns and analyze both restricted and unrestricted short sales).

Figure 5.4: Theoretical and implied Security Market Line.

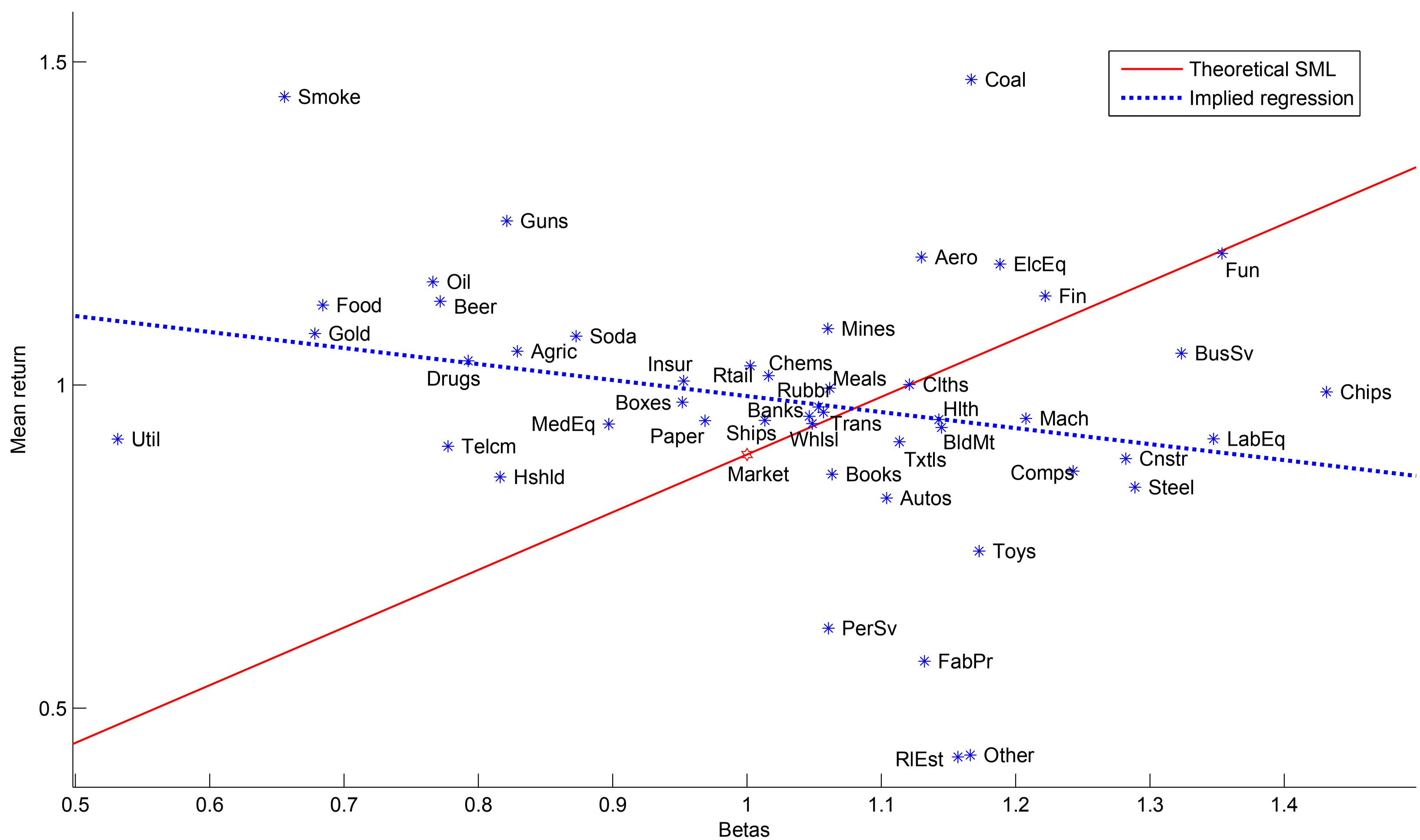
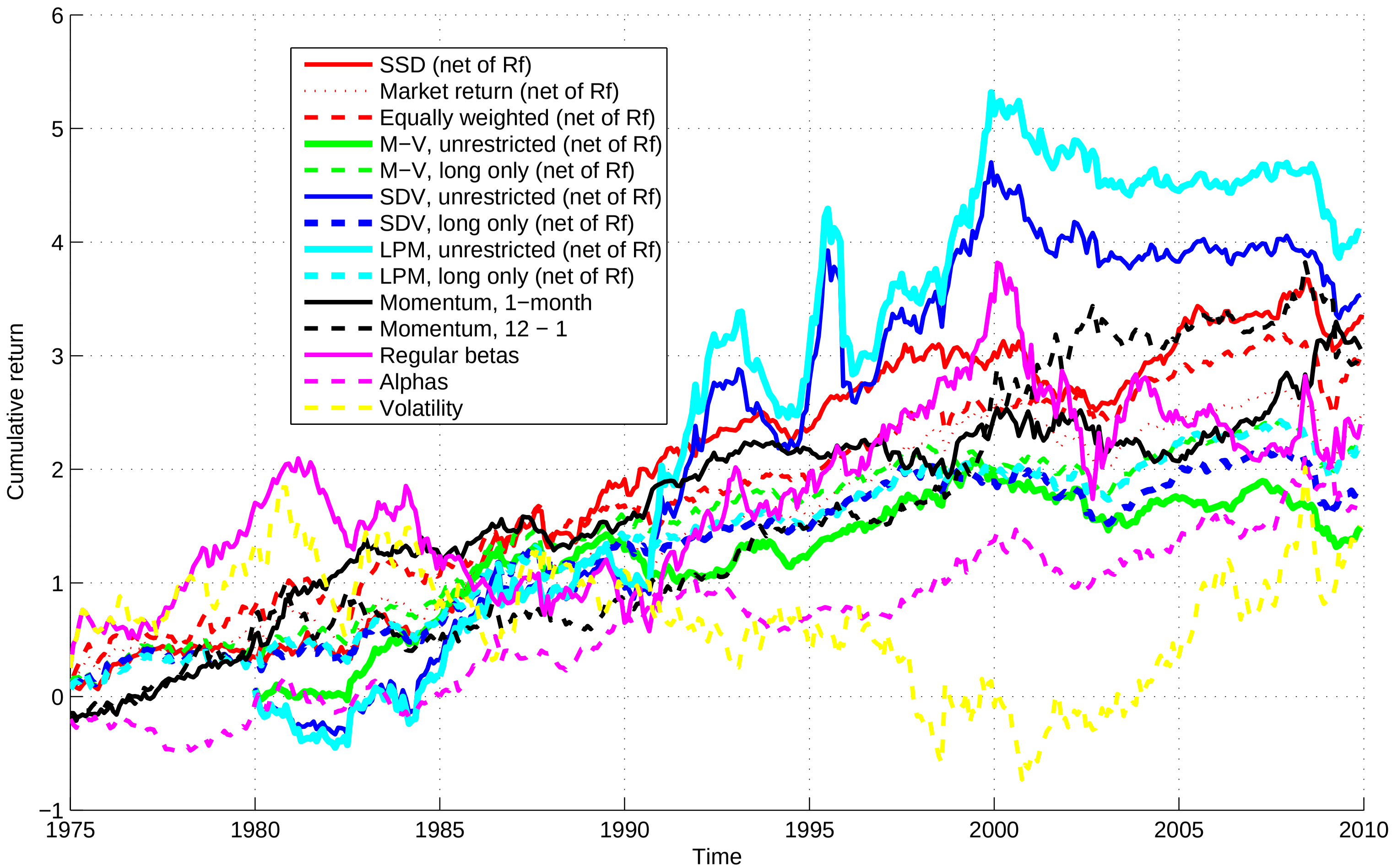


Figure 5.5: Realized monthly returns.



Chapter 6

Conclusions and suggestions for further research

Stochastic Dominance relation is a rather general concept currently being employed in, among other areas, medicine and health economics (Madden, 2009), poverty and inequality studies (Jeffrey and Eidman, 1991, Anderson, 1996) and agriculture (Davidson and Duclos, 2000). The field of financial decision-making is by no means an exception (see Annaert et al. (2009), Levy (2009), Eeckhoudt et al. (2009) and references therein). This thesis focuses on applications of SD in financial portfolio analysis and asset pricing.

In Chapter 2 we reviewed the existing algorithms related to Stochastic Dominance portfolio efficiency that account for full diversification. We classified such algorithms into three main categories: 1) majorization, 2) revealed preference and 3) distribution-based approaches. In addition to classifying and reviewing the categories, we also pointed out some misleading arguments in the existing literature and proposed some refinements.

Chapter 3 relates the concept of Stochastic Dominance to asset pricing theory by pointing out the importance of SD portfolio efficient sets' being convex. The convexity of efficient sets has an important theoretical meaning, as it implies the efficiency of the market portfolio, which in turn leads to heterogeneous investors models, and which can be tested empirically. We reviewed the related results in the literature, addressing the convexity puzzle from two different but interrelated perspectives: distribution of underlying asset returns and restrictions on the set of utility functions. We discussed the importance of finding a reasonable set of utility functions that would lead to a convex efficient set. To this end, we derived efficiency tests for the class of utility functions having increasing relative and decreasing absolute risk aversion, and decomposed the corresponding efficiency sets.

Chapter 4 is devoted to a relaxation of SD, Almost Stochastic Dominance. Almost SD was first proposed in Leshno and Levy (2002), where it was shown to exclude some unrealistic preferences from the class of monotonic and concave utility functions. However, it is extremely difficult to compute an ASD efficient portfolio relative to a diversified portfolio possibilities set in practice. For that reason, we proposed a modification of ASD that is computationally tractable and appears to perform well on all intuitive examples given in Leshno and Levy (2002), as well as some new ones. Following Leshno and Levy (2002) and Levy (2009), we applied the newly defined concept to analyzing investors' preferences between stocks and bonds for the long run.

Finally, in Chapter 5 we applied the concept of Stochastic Dominance for constructing a multi-period asset allocation strategy and analyzed its out-of-sample performance relative to other popular strategies, such as full- and lower-partial moments strategies, and value and momentum rebalancing. We observed a relatively good performance of the SSD strategy and encouraged scholars and practitioners to investigate it further. Moreover, we extended the results of Grootveld and Hallerbach (1999) by reporting a significant difference between efficient frontiers of portfolios formed on the basis of full and partial moments.

Each chapter contributes to the existing literature, but also leaves considerable room for further research. For example, the SSD asset allocation strategy is well worth further investigation, given the intriguing results of Chapter 5 and the relative simplicity of implementing the strategy computationally. Furthermore, finding a set of "well-behaved" utility functions leading to a convex corresponding efficient set would be of high theoretical value, as pointed out in Chapter 3.

Another interesting possibility for continuing the current research could lie in investigating the properties of Robust Stochastic Dominance. This concept is extremely intriguing and certainly deserves to be covered in a separate section.

Robust Stochastic Dominance

Robust programming is a new approach in the theory of optimization started in the 90's with Ben-Tal and Nemirovsky (1997), Ghaoui and Lebret (1997) and Kouvelis and Yu (1997), among others. This approach has recently gained considerable popularity in Operational Research, see Ben-Tal et al. (2009) for a survey and references. The major idea of robust optimization is to employ a so-called uncertainty region for the uncertain parameters in a constraint, and then to enforce that the constraint should hold for all parameter values within this uncertainty region. It has been shown that for several optimization problems and for several choices of the uncertainty region, such so-called robust counterpart problems can be reformulated as tractable optimization problems.

One natural way of employing the robust programming idea in the SD optimization framework is by incorporating the uncertainty about the probabilities of the states of nature, which are commonly assumed equally likely in the majority of SD literature. One way of incorporating uncertainty about the probabilities into the model would be by imposing the following constraints (as suggested in Ben-Tal et al., 2010) to the SD optimization programs:

$$p \in \left\{ q \in \mathbb{R}^m : I_\phi(p^0, q) \le \rho_\phi(N, m, \alpha, q) \right\}, \tag{6.1}$$

where $I_\phi(p,q) = \sum\limits_{i=1}^{m} p_i \phi\left(\frac{q_i}{p_i}\right)$ is a distance function (referred to as ϕ-divergence in the statistical literature) between two probability vectors p and $q \in \mathbb{R}^m$ for some convex function $\phi(t)$ for $t \ge 0$, such that $\phi(1) = 0$. The value $\rho_\phi(N, m, \alpha, q)$ in (6.1) defines a $(1-\alpha)$-confidence set around p^0, as if the null hypothesis $H_0 : p = p^0$ was tested by a goodness-of-fit test on the basis of empirical observations $q^1, \ldots, q^N$.

In the event that the exact distribution of $I_\phi(p,q)$ under H_0 is unknown, one can take a twice continuously differentiable $\phi(t)$ with $\phi''(1) \neq 0$ and use the asymptotic test statistic, as proposed by Ben-Tal et al. (2010):

$$\frac{2N}{\phi''(1)} I_\phi(p^0, q), \tag{6.2}$$

which under H_0 follows asymptotically, for $N \to \infty$, a χ^2_{m-1}-distribution with $m - 1$ degrees of freedom (see Pardo, 2006). Therefore, we can substitute $\rho_\phi(N, m, \alpha, q)$ in (6.1) with

$$\frac{\phi''(1)}{2N} \left(\gamma_\phi + \sqrt{\delta_\phi} \chi^2_{m-1,1-\alpha} \right),$$

where γ_ϕ and δ_ϕ are the bias-correcting parameters given in Pardo (2006, p. 190).

With regard to the probabilities of the states of nature, when no observations $q^1, \ldots, q^N$ are available, we can introduce a "parametric" trust region that contains the equally weighted outcome, for instance by choosing a $0 < \delta < \frac{1}{m}$, and requiring that $p \in A_\delta$, where

$$A_\delta = \left\{ p \in \mathbb{R}^m : \left| p_i - \frac{1}{m} \right| \leq \delta, \quad p^\mathsf{T} e = 1 \right\}. \tag{6.3}$$

We can subsequently require that the Stochastic Dominance constraints, for instance those of (2.14), should hold for all probability vectors π from A_δ defined by (6.3), in the following way (assuming that the states of nature are sorted according to the ordering of y):

$$\max f(\lambda) \tag{6.4a}$$

$$\text{s.t.} \quad \sum_{k=1}^{n} x_{ik} \lambda_k + s_{ij} \geq y_j, \quad i, j = 1, \ldots, m \tag{6.4b}$$

$$\sum_{i=1}^{m} \pi_i s_{ij} \leq \sum_{i=j}^{m} \pi_i (y_i - y_j), \ \forall \pi \in A_\delta, \quad j = 1, \ldots, m \tag{6.4c}$$

$$s_{ij} \geq 0, \quad i, j = 1, \ldots, m \tag{6.4d}$$

$$\lambda \in \Lambda \tag{6.4e}$$

If the objective in (6.4a) is also a function of π, for instance $\mathbb{E}(X\lambda) =$

$\sum_{i=1}^{m}(X\lambda)_i\pi_i$, we can maximize the worst-case scenario for the expected return, by defining $f(\lambda)$ in (6.4a) in the following way:

$$f(\lambda) = \min\left\{\sum_{i=1}^{m}(X\lambda)_i\pi_i : \quad \pi \in A_\delta\right\}. \tag{6.5}$$

As shown in Ben-Tal et al. (2010), the mathematical programming problem (6.4) can be reformulated as a standard convex optimization problem, for which standard optimization software is readily available.

The optimal solution to (6.4) will be robust with respect to possible (small) changes in the probability weights within the uncertainty region. Another possible modification would be to prioritize recent outcomes relative to older realizations, so that the recent outcomes have more influence on the solution, by replacing A_δ in (6.4c) and (6.5) by the following set $B_{l,u}$:

$$B_{l,u} = \left\{p \in \mathbb{R}^m : l_i \leq p_i \leq u_i, \quad p^\mathsf{T}e = 1\right\}, \tag{6.6}$$

where $\{l_i\}$ and $\{u_i\}, i = 1, \ldots, m$ are increasing sequences such that
$l_i \geq 0, \quad l_m \leq \frac{1}{m}, \quad u_i \geq \frac{1}{m}$, and $\quad u_m \leq 1$.

Thus formulated, (6.4) requires the solution (λ^*, s^*) to satisfy the constraints involving s_{ij}: (6.4b), (6.4c) and (6.4d), for all values of π from the uncertainty region imposed.

We have implicitly used $\phi(t) = |t-1|$ (or, equivalently, $I_\phi(p,q) = \sum|p_i - q_i|$) by suggesting (6.3) and (6.6), but it is certainly not the only option. Ben-Tal et al. (2010) show several other possibilities for $\phi(t)$ for which the resulting robust counterpart program can be efficiently solved by standard linear, quadratic or conic programming techniques. Those possibilities include:

$$
\begin{aligned}
&\text{Burg entropy, where} && I_\phi(p,q) = \sum p_i \log\left(\frac{p_i}{q_i}\right), \\
&\chi^2\text{-distance, where} && I_\phi(p,q) = \sum \frac{(p_i - q_i)^2}{p_i}, \\
&\text{Alpha-order entropy:} && I_\phi(p,q) = \sum p_i^\theta q_i^{1-\theta}, \\
&\text{and } \chi\text{-divergence of order } \alpha\text{:} && I_\phi(p,q) = \sum q_i \left|1 - \frac{p_i}{q_i}\right|^\alpha .
\end{aligned}
$$

Given the recent boom in developments in the Robust Optimization literature (which has also reached the Tinbergen Institute: see van Oord, Martens, and van Dijk (2009) for a robust version of the momentum strategy) and the possibility of the statistical interpretation of the trust regions for some parameters, it would definitely be of interest to study Robust Stochastic Dominance criteria as its application. Challenging research questions include: how restrictive the robustness constraints are; what the critical perturbation level may be at which an SSD efficient portfolio remains Robust-SSD efficient; and how the increasing influence of recent states in (6.6) is reflected in terms of the stability and sensitivity of the solution.

Moreover, it is technically easy to introduce a Robust Almost Stochastic Dominance (RASD), simply by adding an uncertainty region of the type (6.6) to optimization programs (4.9), (4.10) or (4.11). Investigating the economic interpretation of such criteria (e.g. studying the properties of the set of utility functions whose optimal portfolios are RASD efficient) would definitely contribute to the existing literature on Stochastic Dominance.

Nederlandse Samenvatting (Summary in Dutch)

Stochastische Dominatie (verder SD) is een relatie tussen twee kansvariabelen die in Wiskunde door Mann en Whitney (1947) en Lehmann (1955), en in Economie door Quirk en Saposnik (1962), Hadar en Russell (1969), Hanoch en Levy (1969) en Rothschild en Stiglitz (1970) is geïntroduceerd. SD maakt gebruik van de volledige informatie over het gehele domein van de CDF in het besluitvormingsproces, in tegenstelling tot de modellen op basis van gemiddelde en variantie of andere risicomaatstaven die alleen een afzonderlijk kenmerk van de CDF gebruiken als de maat van portefeuillerisico, zoals een partieel moment of de variantie. SD vindt talrijke toepassingen in o.a. de gezondheidseconomie (Madden, 2009), armoede- en ongelijkheidstudies (Jeffrey en Eidman, 1991, Anderson, 1996) en landbouwstudies (Davidson en Duclos, 2000).

Deze studie is gericht op de toepassing van SD in portefeuilleanalyse en de prijsvorming van effecten. In de portefeuilleanalyse zijn er sinds het jaar 2003 een aantal algoritmen voorgesteld om te identificeren of een bepaalde portefeuille efficiënt is (d.w.z. niet gedomineerd wordt door een andere beschikbare portefeuille). In Hoofdstuk 2 geven wij een overzicht van deze methoden en classificeren die naar drie categorieën: 1) majorizatie, 2) 'gebleken voorkeur' en 3) de distributie gebaseerde benadering. Naast het classificeren en geven van het overzicht, wijzen wij in dat hoofdstuk op enkele misleidende betogen in de bestaande publicaties en stellen verbeteringen voor sommige methoden voor.

Hoofdstuk 3 focust op de toepassing van SD in de prijsvorming van effecten door het wijzen op het belang van convexe efficiënte sets. Het blijkt dat indien een set van efficiënte portefeuilles convex is, sommige modellen van heterogene beleggers mogen worden toegepast. Wij bespreken het belang om een uitgebreide set van nutfuncties te vinden die automatisch tot een convexe efficiënte portefeuilleset zou leiden. Hiertoe ontwikkelen we een aantal SD algoritmen om efficiëntie te checken voor sommige categorieën van nutsfuncties, waaronder dalende absolute- en toenemende relatieve riscioaversie.

Hoofdstuk 4 gaat over Almost SD, een versoepeling van het SD begrip. LL-Almost SD werd geïntroduceerd door Leshno en Levy (2002) en blijkt een aantal realistische keuzen in prakijk te kunnen modelleren die buiten de mogelijkheden van de standaard SD methoden liggen. Een nadeel van LL-Almost SD is dat de rekenkundige problemen erg weerspannig zijn. Daarom introduceren we een andere versoepeling van SD die net zo goed presteert in het modeleren van de realistische keuzen en daarenboven gemakkelijk is om te berekenen. We ontwikkelen een aantal algoritmen voor de nieuwe ASD en passen dit begrip toe op het modelleren van voorkeuren tussen obligaties en aandelen op een langere termijn.

Hoofdstuk 5 is vooral empirisch. Daar passen we SSD algoritmen toe op een meer-perioden beleggingsstrategie. Daarvoor gebruiken wij 48 industriefondsen vanaf januari 1970 t/m december 2009. We vergelijken de nieuwe beleggingsstrategie met andere erkende strategieën zoals momentum, value, volle en gedeeltelijke momenten. Het blijkt dat de nieuwe SSD strategie relatief goed presteert. Daarom moedigen we wetenschappers aan om deze verder te onderzoeken.

Ten slotte bespreken we de conclusies van dit proefschrift in Hoofdstuk 6, waar we ook een aantal interessante mogelijkheden suggereren om het onderzoek voort te zetten. Een van die is een nieuw begrip van Robuste Stochastische Dominantie, die is bedoeld om een "onzekerheidregio's" te gebruiken voor sommige coëfficiënten in standaard SD algoritmen, zoals de waarschijnlijkheden van verschillende rendementen.

Bibliography

Aboudi, R. and Thon, D., Efficient algorithms for stochastic dominance tests based on financial data. *Management Science*, 40(4):508–515, April 1994.

Afriat, S., The construction of a utility function from expenditure data. *Internat. Econom. Rev.*, 8:67–77, 1967.

Anderson, G., Nonparametric tests of stochastic dominance in income distributions. *Econometrica*, 64(5):1183–1193, 1996.

Ang, A., Xing, Y. and Chen, J., Downside risk. *Review of Financial Studies*, 19(4):1191–1239, 2006.

Ang, J. S., A note on the E-S portfolio selection model. *Journal of Financial and Quantitative Analysis*, 10(5):849–857, 1975.

Annaert, J., van Osselaer, S. and Verstraete, B., Performance evaluation of portfolio insurance strategies using stochastic dominance criteria. *Journal of Banking & Finance*, 33:272–280, 2009.

Bawa, V. S., Stochastic dominance: a research bibliography. *Management Science*, 28(6):698–712, 1982.

Bawa, V. S., Lindenberg, E. B. and Rafsky, L. C., An efficient algorithm to determine stochastic dominance admissible sets. *Management Science*, 25 (7):609–622, July 1979.

Bawa, V. and Lindenberg, E., Capital market equilibrium in a mean-lower partial moment framework. *Journal of Financial Economics*, 5(2):189–200, 1977.

Ben-Tal, A., den Hertog, D., Waegenaere, A. D., Melenberg, B. and Rennen, G. Robust optimization with uncertainty regions based on phi-divergence. Center working paper, Tilburg University, 2010.

Ben-Tal, A., Ghaoui, L. E. and Nemirovsky, A. *Robust Optimization*. Princeton Univ. Press, 2009.

Ben-Tal, A. and Nemirovsky, A., Stable truss topology design via semidefinite programming. *SIAM Journal on Optimization*, 7(4):991–1016, 1997.

Bernoulli, D., Specimen theoriae novae de mensura sortis. *Commentarii academaiae scientiarium imperiales petropolitanae*, 5(I):75–192, 1738. (Translated by L. Sommer in *Econometrica*, 1954, 22, 23–36).

Bernstein, P., The time of your life. *J. of Portfolio Management*, 2(4):4–7, 1976.

Blitz, D. and van Vliet, P., The volatility effect: Lower risk without lower return. *Journal of Portfolio Management*, pages 102–113, Fall 2007.

Blitz, D. and van Vliet, P., Global tactical cross-asset allocation: Applying value and momentum across asset classes. *Journal of Portfolio Management*, 2008.

Canner, N., Mankiw, N. G. and Weil, D. N., An asset allocation puzzle. *American Economic Review*, 87(1):181–191, 1997.

Cass, D. and Stiglitz, J. E., The structure of investor preferences and asset returns, and separability in portfolio allocation: A contribution to the pure theory of mutual funds. *Journal of Economic Theory*, 2:122–160, 1970.

Clarke, R. G., de Silva, H. and Thorley, S., Minimum-variance portfolios in the U.S. equity market. *The Journal of Portfolio Management*, 33(1): 10–24, 2006.

Cochrane, J. H. *Asset Pricing*. Princeton University Press, 2nd edition, 2005.

Davidson, R. and Duclos, J.-Y., Statistical inference for stochastic dominance and for the measurement of poverty and inequality. *Econometrica*, 68(6): 1435–1464, 2000.

Dentcheva, D. and Ruszczyński, A., Optimization with stochastic dominance constraints. *SIAM J. Optim.*, 14(2):548–566, 2003.

Dentcheva, D. and Ruszczyński, A., Semi-infinite probabilistic optimization: first-order stochastic dominance constraints. *Optimization*, 53(5-6):583–601, 2004.

Dentcheva, D. and Ruszczyński, A., Inverse stochastic dominance constraints and rank dependent expected utility theory. *Mathematical Programming*, 108(2-3, Ser. B):297–311, 2006a.

Dentcheva, D. and Ruszczyński, A., Portfolio optimization with stochastic dominance constraints. *J. of Banking and Finance*, 30(2):433–451, 2006b.

Dentcheva, D. and Ruszczyński, A., Inverse cutting plane methods for optimization problems with second-order stochastic dominance constraints. *Optimization*, 59(3):323–338, 2010.

Dybvig, P. H. and Ross, S. A., Portfolio efficient sets. *Econometrica*, 50(6): 1525–1546, Nov. 1982.

Eeckhoudt, L., Schlesinger, H. and Tsetlin, I., Approximating of risks via stochastic dominance. *Journal of Economic Theory*, 144(3):994–1003, May 2009.

Fishburn, P., Convex stochastic dominance with continuous distributions functions. *Journal of Economic Theory*, 7(2):143–158, 1974.

Fishburn, P., Mean-risk analysis with risk associated with below-target returns. *American Economic Review*, pages 116–126, 1977.

Frankfurter, G., Phillips, H. and Seagle, J., Portfolio selection: The effect of uncertain means, variances, and covariances. *The Journal of Financial and Quantitative Analysis*, 6(5):1251–1262, 1971.

French, K. Online data library, Accessed May 2010. http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.

Ghaoui, L. E. and Lebret, H., Robust solution to least-squares problems with uncertain data. *SIAM Journal on Matrix Analysis and Applications*, 18: 1035–1064, 1997.

Grootveld, H. and Hallerbach, W., Variance vs downside risk: Is there really that much difference? *European Journal of Operational Research*, 114: 304–319, 1999.

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

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

Hardy, G. H., Littlewood, J. E. and Polya, G. *Inequalities.* Cambridge University Press, Cambridge, U.K., 1934.

Harlow, W. V. and Rao, R. K. S., Asset pricing in a generalized mean-lower partial moment framework: Theory and evidence. *The Journal of Financial and Quantitative Analysis*, 24(3):285–311, Sep. 1989.

Hazetl, P. B. R. and Norton, R. D. *Mathematical Programming for Economic Analysis in Agriculture.* New York: Macmillan Publishing Company, 1986.

Hogan, W. W. and Warren, J. M., Computation of the efficient boundary in the E–S portfolio selection model. *The Journal of Financial and Quantitative Analysis*, 7(4):1881–1896, September 1972.

Holthausen, D., A risk-return model with risk and return measured as deviations from a target. *American Economic Review*, 71:182–188, 1981.

Ingersoll, J. E. *Theory of Financial Decision Making.* Rowman & Littlefield Publishers, Inc., 1987. ISBN 0-8476-7359-6.

Jeffrey, S. R. and Eidman, V. R., The application of multivariate stochastic dominance criteria to agricultural economic problems. *Canadian Journal of Agricultural Economics/Revue canadienne d'agroeconomie*, 39(2):193–209, 1991.

Jegadeesh, N. and Titman, S., Returns to buying winners and selling losers: Implications for stock market efficiency. *The Journal of Finance*, 48(1): 65–91, Mar. 1993.

Jegadeesh, N. and Titman, S., Profitability of momentum strategies: An evaluation of alternative explanations. *The Journal of Finance*, 56(2): 699–720, Apr. 2001.

Kang, T., Brorsen, B. and Adam, B., A new efficiency criterion: The mean-separated target deviations model. *Journal of Economics and Business*, 48:47–66, 1996.

Kopa, M. and Chovanec, P., A second-order stochastic dominance portfolio efficiency measure. *Kybernetika (Prague)*, 44:243–258, 2008.

Kopa, M. and Post, T., A portfolio optimality test based on the First-Order Stochastic Dominance criterion. *Journal of Financial and Quantitative Analysis*, 44:1103–1124, 2009.

Kouvelis, P. and Yu, G. *Robust Discrete Optimization and its Application*. Kluwer Academic Publishers, London, 1997.

Kuosmanen, T. Stochastic dominance efficiency tests under diversification. Working Paper W-283, Helsinki School of Economics, 2001-WP.

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

Kuosmanen, T., Performance measurement and best-practice benchmarking of mutual funds: combining stochastic dominance criteria with data envelopment analysis. *Journal of Productivity Analysis*, 28(1):71–86, 2007.

Lehmann, E. L., Ordered families of distributions. *Ann. Math. Statist.*, 26: 399–419, 1955.

Leshno, M. and Levy, H., Preferred by "all" and preferred by "most" decision makers: Almost stochastic dominance. *Management Science*, 48(8):1074–1085, 2002.

Levy, H., Stochastic dominance and expected utility: Survey and analysis. *Management Science*, 38(4), 1992.

Levy, H. *Stochastic Dominance: Investment Decision Making under Uncertainty*. Springer Science + Business Media, inc., second edition, 2006.

Levy, M., Almost stochastic dominance and stocks for the long run. *European Journal of Operational Research*, 194(1):250–257, 2009.

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

Linton, O., Song, K. and Whang, Y.-J., An improved bootstrap test of stochastic dominance. *Journal of Econometrics*, 154(2):186–202, Feb. 2010.

Lizyayev, A. M. and Ruszczyński, A. Tractable almost stochastic dominance. Working paper, RUTCOR, Rutgers University, May 2010.

Lizyayev, A. M. Stochastic dominance: Convexity and some efficiency tests. Tinbergen Institute working paper series 09-112/2, Erasmus University Rotterdam, 9 December 2009.

Lizyayev, A. M. Stochastic dominance efficiency analysis of diversified portfolios: Classification, comparison and refinements. Tinbergen Institute working paper series 10-084/2, Erasmus University Rotterdam, August 2010.

Luedtke, J., New formulations for optimization under stochastic dominance constraints. *SIAM Journal on Optimization*, 19(3):1433–1450, 2008.

Madden, D., Mental stress in Ireland, 1994-2000: a stochastic dominance approach. *Health Economics*, 18(10):1202–1217, 2009.

Mann, H. B. and Whitney, D. R., On a test of whether one of two random variables is stochastically larger than the other. *Ann. Math. Statistics*, 18: 50–60, 1947.

Mansini, R., Ogryczak, W. and Speranza, M. G., LP solvable models for portfolio optimization: a classification and computational comparison. *IMA J. Manag. Math.*, 14(3):187–220 (2004), 2003.

Mao, J. C. T., Models of capital budgeting, E-V vs E-S. *Journal of Financial and Quantitative Analysis*, 4:657–675, 1970.

Markowitz, H. M., Portfolio selection. *J. Finance*, 7:77–91, 1952.

Markowitz, H. M., Investment for the long run: New evidence for an old rule. *Journal of Finance*, 31:1273–1286, 1976.

Markowitz, H. M. *Mean Variance Analysis In Portfolio Choice And Capital Markets*. Basil Blackwell, New York, 1978.

Markowitz, H. M. Samuelson and investment for the long run. In Szenberg, M., editor, *Samuelsonian Economics and the 21st Century*. Oxford University Press, 2006.

Markowitz, H. *Portfolio Selection: Effcient Diversication of Investments*. Wiley, New York, 1959.

Marshall, A. W. and Olkin, I. *Inequalities: Theory of Majorization and Its Applications*. Academic Press Inc., New York, 1979.

Meyer, J., Two-moment decision models and expected utility maximization. *American Economic Review*, 77:421–430, June 1987.

Miller, N. and Ruszczyński, A., Risk-adjusted probability measures in portfolio optimization with coherent measures of risk. *European J. Oper. Res.*, 191(1):193–206, 2008.

Noyan, N., Rudolf, G. and Ruszczyński, A., Relaxations of linear programming problems with first order stochastic dominance constraints. *Operations Research Letters*, 34, 2006.

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

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

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

Pardo, L. *Statistical Inference Based on Divergence Measures.* Chapman & Hall/CRC, Boca Raton, 2006.

Performance World, . GAMS Development Corp., Accessed October 2009. http://www.gamsworld.org/performance/plib/size.htm.

Post, T., Empirical tests for stochastic dominance efficiency. *Journal of Finance*, 58(5):1905–1931, 2003.

Post, T., On the dual test for SSD efficiency: With an application to momentum investment strategies. *European Journal of Operational Research*, 185(3):1564–1573, 2008.

Post, T., van Vliet, P. and Lansdorp, S. Sorting out downside beta. Working paper ERS-2009-006-F&A, ERIM Report series, February 2009.

Post, T. and Versijp, P., Multivariate tests for stochastic dominance efficiency of a given portfolio. *JFQA*, 42(2):489–516, June 2007.

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

Quirk, J. and Saposnik, R., Admissibility and measurable utility functions. *Rev. Econom. Stud.*, 29:140–146, 1962.

Rockafellar, R. T. *Convex Analysis.* Princeton University Press, Princeton, NJ, 1970.

Rockafellar, R. T. and Uryasev, S., Conditional value-at-risk for general loss distributions. *J. Banking Finance*, 26, 2002.

Ross, S. A., Mutual fund separation in financial theory - the separating distributions. *Journal of Economic Theory*, 17:254–286, 1978.

Rothschild, M. and Stiglitz, J., Increasing risk. I. A definition. *J. Econ. Theory*, 2:225–243, 1970.

Rothschild, M. and Stiglitz, J., Increasing risk. II. Its economic consequences. *J. Econ. Theory*, 2:66–84, 1971.

Rubinstein, M., An aggregation theorem for securities markets. *Journal of Financial Economics*, I:225–244, 1974.

Rudolf, G. and Ruszczyński, A., Optimization problems with second order stochastic dominance constraints: Duality, compact formulations, and cut generation methods. *SIAM Journal on Optimization*, 19(3):1326–1343, 2008.

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

Saha, A., Expo-power utility: A 'flexible' form for absolute and relative risk aversion. *Amer. J. Agr. Econ.*, 75:905–913, November 1993.

Samuelson, P., Lifetime portfolio selection by dynamic stochastic programming. *Review of Economics and Statistics*, 51:239–246, 1969.

Samuelson, P., The judgment of economic science on rational portfolio management. *Journal of Portfolio Management*, 16:4–12, 1989.

Samuelson, P., The long-run case for equities and how it can be oversold. *Journal of Portfolio Management*, 21:15–24, 1994.

Shalit, H. and Yitzhaki, S., Mean-gini, portfolio theory, and the pricing of risky assets. *The Journal of Finance*, 39(5):1449–1468, December 1984.

Shalit, H. and Yitzhaki, S., Marginal conditional stochastic dominance. *Management Science*, 40(5):670–684, May 1994.

Strassen, V., The existence of probability measures with given marginals. *The Annals of Mathematical Statistics*, 36(2):423–439, Apr. 1965.

van Oord, J., Martens, M. and van Dijk, H. Robust optimization of the equity momentum strategy. Discussion Paper 011/4, Tinbergen Institute, 2009. Econometrics.

Varian, H. R., Non-parametric tests of consumer behaviour. *The Review of Economic Studies*, 50(1):99–110, Jan. 1983.

Versijp, P. *Advances in the use of stochastic dominance in asset pricing*. PhD Thesis, Tinbergen Institute, Rotterdam, 2007. # 407.

Vickson, R. G., Stochastic Dominance Tests for Decreasing Absolute Risk Aversion. I: Discrete Random Variables. *Management Science*, 21(12): 1438–1446, Aug. 1975.

Vickson, R. G., Stochastic Dominance Tests for Decreasing Absolute Risk-Aversion II: General Random Variables. *Management Science*, 23(5):478–489, Jan. 1977.

The Tinbergen Institute is the Institute for Economic Research, which was founded in 1987 by the Faculties of Economics and Econometrics of the Erasmus University Rotterdam, University of Amsterdam and VU University Amsterdam. The Institute is named after the late Professor Jan Tinbergen, Dutch Nobel Prize laureate in economics in 1969. The Tinbergen Institute is located in Amsterdam and Rotterdam. The following books recently appeared in the Tinbergen Institute Research Series:

437. R. LORD, *Efficient pricing algorithms for exotic derivatives.*
438. R.P. WOLTHOFF, *Essays on Simultaneous Search Equilibrium.*
439. Y.-Y. TSENG, *Valuation of travel time reliability in passenger transport.*
440. M.C. NON, *Essays on Consumer Search and Interlocking Directorates.*
441. M. DE HAAN, *Family Background and Children's Schooling Outcomes.*
442. T. ZAVADIL, *Dynamic Econometric Analysis of Insurance Markets with Imperfect Information.*
443. I.A. MAZZA, *Essays on endogenous economic policy.*
444. R. HAIJEMA, *Solving large structured Markov Decision Problems for perishable-inventory management and traffic control.*
445. A.S.K. WONG, *Derivatives in Dynamic Markets.*
446. R. SEGERS, *Advances in Monitoring the Economy.*
447. F.M. VIEIDER, *Social Influences on Individual Decision Making Processes.*
448. L. PAN, *Poverty, Risk and Insurance: Evidence from Ethiopia and Yemen.*
449. B. TIEBEN, *The concept of equilibrium in different economic traditions: A Historical Investigation.*
450. P. HEEMEIJER, *Expectation Formation in Dynamic Market Experiments.*
451. A.S. BOOIJ, *Essays on the Measurement Sensitivity of Risk Aversion and Causal Effects in Education.*
452. M.I. LÓPEZ YURDA, *Four Essays on Applied Micro econometrics.*
453. S. MEENTS, *The Influence of Sellers and the Intermediary on Buyers' Trust in C2C Electronic Marketplaces.*
454. S. VUJIĆ, *Econometric Studies to the Economic and Social Factors of Crime.*
455. F. HEUKELOM, *Kahneman and Tversky and the Making of Behavioral Economics.*
456. G. BUDAI-BALKE, *Operations Research Models for Scheduling Railway Infrastructure Maintenance.*
457. T.R. DANIËLS, *Rationalised Panics: The Consequences of Strategic Uncertainty during Financial Crises.*
458. A. VAN DIJK, *Essays on Finite Mixture Models.*
459. C.P.B.J. VAN KLAVEREN, *The Intra-household Allocation of Time.*
460. O.E. JONKEREN, *Adaptation to Climate Change in Inland Waterway Transport.*

461. S.C. GO, *Marine Insurance in the Netherlands 1600-1870, A Comparative Institutional Approach.*
462. J. NIEMCZYK, *Consequences and Detection of Invalid Exogeneity Conditions.*
463. I. BOS, *Incomplete Cartels and Antitrust Policy: Incidence and Detection*
464. M. KRAWCZYK, *Affect and risk in social interactions and individual decision-making.*
465. T.C. LIN, *Three Essays on Empirical Asset Pricing.*
466. J.A. BOLHAAR, *Health Insurance: Selection, Incentives and Search.*
467. T. FARENHORST-YUAN, *Efficient Simulation Algorithms for Optimization of Discrete Event Based on Measure Valued Differentiation.*
468. M.I. OCHEA, *Essays on Nonlinear Evolutionary Game Dynamics.*
469. J.L.W. KIPPERSLUIS, *Understanding Socioeconomic Differences in Health An Economic Approach.*
470. A. AL-IBRAHIM, *Dynamic Delay Management at Railways: A Semi-Markovian Decision Approach.*
471. R.P. FABER, *Prices and Price Setting.*
472. J. HUANG, *Education and Social Capital: Empirical Evidences from Microeconomic Analyses.*
473. J.W. VAN DER STRAATEN, *Essays on Urban Amenities and Location Choice.*
474. K.M. LEE, *Filtering Non Linear State Space Models: Methods and Economic Applications.*
475. M.J. REINDERS, *Managing Consumer Resistance to Innovations.*
476. A. PARAKHONYAK, *Essays on Consumer Search, Dynamic Competition and Regulation.*
477. S. GUPTA, *The Study of Impact of Early Life Conditions on Later Life Events: A Look Across the Individual's Life Course.*
478. J. LIU, *Breaking the Ice between Government and Business: From IT Enabled Control Procedure Redesign to Trusted Relationship Building.*
479. D. RUSINOVA, *Economic Development and Growth in Transition Countries.*
480. H. WU, *Essays on Top Management and corporate behavior.*
481. X. LUI, *Three Essays on Real Estate Finance*
482. E.L.W. JONGEN, *Modelling the Impact of Labour Market Policies in the Netherlands*
483. M.J. SMIT, *Agglomeration and Innovations: Evidence from Dutch Microdata*
484. S.VAN BEKKUM, *What is Wrong With Pricing Errors? Essays on Value Price Divergence*
485. X. HU, *Essays on Auctions*
486. A.A. DUBOVIK, *Economic Dances for Two (and Three)*