

# MASTER FUNDS IN PORTFOLIO ANALYSIS WITH GENERAL DEVIATION MEASURES

*R. Tyrrell Rockafellar*<sup>1</sup>, *Stan Uryasev*<sup>2</sup>, *Michael Zabarankin*<sup>2</sup>

**Version: July 5, 2004**

*Correspondence should be addressed to: Stan Uryasev*

## Abstract

Generalized measures of deviation are considered as substitutes for standard deviation in a framework like that of classical portfolio theory for coping with the uncertainty inherent in achieving rates of return beyond the risk-free rate. Such measures, derived for example from conditional value-at-risk and its variants, can reflect the different attitudes of different classes of investors. They lead nonetheless to generalized one-fund theorems in which a more customized version of portfolio optimization is the aim, rather than the idea that a single “master fund” might arise from market equilibrium and serve the interests of all investors.

The results that are obtained cover discrete distributions along with continuous distributions. They are applicable therefore to portfolios involving derivatives, which create jumps in distribution functions at specific gain or loss values, well as to financial models involving finitely many scenarios. Furthermore, they deal rigorously with issues that come up at that level of generality, but have not received adequate attention, including possible lack of differentiability of the deviation expression with respect to the portfolio weights, and the potential nonuniqueness of optimal weights.

The results also address in detail the phenomenon that if the risk-free rate lies above a certain threshold, the usually envisioned master fund must be replaced by one of alternative type, representing a “net short position” instead of a “net long position” in the risky instruments. For nonsymmetric deviation measures, the second type need not just be the reverse of the first type, and there can sometimes even be an interval for the risk-free rate in which no master fund of either type exists. A notion of basic fund, in place of master fund, is brought in to get around this difficulty and serve as a single guide to optimality regardless of such circumstances.

**Keywords:** *deviation measures, risk measures, value-at-risk, conditional value-at-risk, portfolio optimization, one-fund theorem, master fund, basic fund, efficient frontier, convex analysis.*

---

<sup>1</sup>University of Washington, Department of Mathematics, Box 354350, Seattle, WA 98195-4350;  
E-mail: [rtr@math.washington.edu](mailto:rtr@math.washington.edu). Research supported by NSF grant *DMS-0104055*

<sup>2</sup>University of Florida, ISE Department, P.O. Box 116595, 303 Weil Hall, Gainesville, FL 32611-6595;  
E-mails: [uryasev@ufl.edu](mailto:uryasev@ufl.edu), [zabarank@ufl.edu](mailto:zabarank@ufl.edu); URLs: [www.ise.ufl.edu/uryasev](http://www.ise.ufl.edu/uryasev), <http://plaza.ufl.edu/zabarank>

# 1 Introduction

In classical portfolio theory, investors respond to the uncertainty of profits by selecting portfolios that minimize variance, or equivalently standard deviation, subject to achieving a specified level in expected gain [16, 10, 15]. The well known “one-fund theorem” [29, 27] stipulates that this can be accomplished in terms of a single “master fund” portfolio by means of a formula which balances the amount invested in that portfolio with the amount invested at the current risk-free rate. Nowadays, other approaches to uncertainty have gained in popularity. Portfolios are being selected on the basis of percentile characteristics such as *value-at-risk* (VaR), *conditional value-at-risk* (CVaR), or other properties proposed for use in risk assessment; cf. [2, 11, 13] and earlier alternatives such as in [6]. These measures have no pretension to being universal, however; VaR and CVaR depend, for instance, on the specification of a confidence level parameter, which could vary among investors. Instead, what is apparent in the alternative approaches currently being touted is a move toward a kind of partial *customization* of responses to risk, while still avoiding, as impractical, a reliance on specifying individual utility functions.

A question in this evolving environment is the extent to which classical facts persist when the minimization of standard deviation is replaced by the minimization of some “nonstandard deviation.” Researchers have already looked into the possibilities in special cases under various limiting assumptions (recognized explicitly or imbedded implicitly). Our goal, in contrast, is to demonstrate that important parallels with classical results exist much more broadly, despite technical hurdles, and in this way to bring out features that have not completely been analyzed, or even perceived, in the past.

We focus on the general *deviation measures* we developed axiomatically in [23]. Our idea is to substitute such a deviation measure for standard deviation in the setting of classical theory and investigate the consequences rigorously in detail. Furthermore, we aim at doing so, for the first time, in cases where the rates of return may have discrete distributions, or mixed discrete-continuous distributions (which can arise from derivatives, such as options), as well as cases where they have continuous distributions.

The deviation measures we work with are paired one-to-one, through [23], with *risk measures* in the sense of Artzner, Delbaen, Eber and Heath [5], but differ in partly relaxing their requirements while insisting on an additional property beyond theirs. A similar additional property was invoked by Ogryczak and Ruszczyński in [17, 18] for *safety measures*, which may be viewed as negatives of risk measures. Minimizing a deviation measure subject to a constraint on expected returns can anyway be different from minimizing the corresponding risk measure, since, as shown in [23], the first problem always has a solution but the second problem can sometimes fail to have a solution, due to a phenomenon of “acceptably free lunches.” We are not, however, suggesting that deviation measures are better than risk measures. Both have their place, but deviation measures fit closer to the classical picture and therefore serve better the particular purposes of this paper.

The axioms for deviation measures that we adopt from [23] entail convexity. They cover numerous choices from classical type to CVaR type, but exclude the analogous expressions of VaR type, since those lack convexity. Convexity is essential for answering most of the harder questions that confront us. Its importance for sound applications in finance has already been recognized as well in connection with the coherency concept in [5]. We do, however, try to indicate along the way the troubles that VaR type expressions would bring up.

The first of our main results says that a one-fund theorem holds regardless of the particular choice of the deviation measure, but with certain modifications. The optimal risky portfolio need not always be unique, and it might not always be expressible by a “master fund” as traditionally conceived, even when only standard deviation is involved. An alternative concept of “basic fund” is introduced to

fill the gap. The rest of our main results pin down precisely the degree to which basic funds can, or cannot, be rescaled into master funds. This turns out to require an understanding not only of an efficient frontier for risky portfolios at price 1, associated with “master funds of positive type,” but also of such a frontier for risky portfolios at price  $-1$ , associated with “master funds of negative type.” The magnitude of the risk-free rate of return plays a key role here. We prove that when it is below a certain threshold, the positive type prevails, but when it is above a certain threshold, the negative type has to be brought in. Moreover in special situations those thresholds can differ, leaving a gap filled by an interval of magnitudes of the risk-free rate for which neither type of master fund can replace a basic fund. We also explain how thresholds can be calculated by solving an auxiliary optimization problem.

It deserves emphasis that, in contrast to much of the previous work in this area, our results are obtained without relying on the existence of densities for the statistical distributions that arise, or even on the continuity of the distribution functions, which would preclude applications to discrete random variables or effects tied to derivatives. We do not take for granted, or require, the differentiability of the deviation with respect to the parameters specifying the relative weights of the instruments in the portfolio. This is not merely for the sake of technical generality. An example provided in the final section of the paper illustrates how put and call options in portfolios can lead to nondifferentiability as well as to a threshold gap for the risk-free rate. Therefore no theory, unless it faces up to such troubles, can be regarded as fully applicable to portfolios involving derivative instruments.

With standard calculus being inadequate for the problems at hand, we have had to rely instead on techniques of convex analysis [19] while adhering strictly to the principles of optimization theory.<sup>3</sup>

The need for a “negative” efficient frontier referring to “net short positions,” along with the usual “positive” one for “net long positions,” is not surprising, in view of the diversity of measures that investors may be using. In line with their different opinions about risk, some investors may find the risk-free rate high enough to warrant borrowing from the market and investing that money risk-free, while others will prefer a fund in which the “longs” outweigh the “shorts.” An interesting analogy can be found in [26, p. 507] in terms of a stock index futures contract which might even consist entirely of short positions.

The emergence of a variety of different master funds, optimal for different deviation measures, is an inescapable outcome of any theory which, like ours, attempts to cope with the current tendency toward customization in portfolio optimization. A master fund identified with respect to the wishes of one class of investors can no longer be proposed as obviously furnishing input for factor analysis of the market as a whole, because the financial markets react to the wishes of *all* investors. A master fund, in our general sense, can no longer be interpreted as associated with a sort of universal equilibrium. Whether some such master funds, individually or collectively, might nonetheless turn out to be valuable in factor analysis, is an issue outside the scope of this paper. CAPM-like covariance relations do indeed come out of the optimality conditions that characterize our master funds (as can be

---

<sup>3</sup>Lack of familiarity with the mathematics of optimization has been a handicap in some of the finance literature in this area, going all the way back to Markowitz. In his pioneering work [16], for example, he excluded short positions by constraining the portfolio weights to be nonnegative. He neglected, however, to take into account that Lagrange multipliers for those inequality constraints could come into play, in which case a closed-form solution to the optimality conditions for a master fund would be impossible. Supposing that the multipliers can be taken to be zero is equivalent (because of convexity) to supposing that, if shorting were allowed, there would anyway be no short positions at optimality. There is no support for that conviction, however, and indeed, numerical calculations are known to produce quite different answers when shorting is allowed and when it is not. Similar looseness about whether solutions to optimization problems even exist must be the reason why the magnitude of the risk-free rate was not perceived to have an effect, and the need for master funds representing net short positions went undetected. The need for allowing at least some short positions as possibilities in a master fund was emphasized by [26, pp. 500, 505].

gleaned from the optimality prescriptions in [23]), but we reserve this development, requiring further elaboration of underpinnings in convex analysis, to a follow-up paper [25].

## 2 Deviation and Risk

We start by reviewing what we mean by “deviation measures” and explaining how they are related to “risk measures” in the terminology of Artzner et al. [5]. Roughly speaking, deviation measures evaluate the degree of nonconstancy in a random variable (i.e., the extent to which outcomes may deviate from expectations), whereas risk measures evaluate overall prospective loss (from the benchmark of zero loss). Deviation measures correspond, under a basic pairing, with risk measures that have an additional “expectation boundedness” property not contemplated in [5]. This pairing is valuable for transforming known examples of risk measures into examples of deviation measures, as well as for illuminating how the valuable notion of coherency introduced in [5] translates to the context of deviations.

We consider a space  $\Omega$ , the elements  $\omega$  of which can represent future states or scenarios (perhaps just finitely many), and suppose it to be supplied with a probability measure  $P$  and the other technicalities that make it a legitimate probability space. We treat as random variables (r.v.’s) the (measurable) functions  $X$  on  $\Omega$  for which  $E[X^2] < \infty$ ; the space of such functions will be denoted, for short, by  $\mathcal{L}^2(\Omega)$ . For  $X$  in  $\mathcal{L}^2(\Omega)$ , the mean  $\mu(X)$  and variance  $\sigma^2(X)$  are well defined in particular:<sup>4</sup>

$$\begin{aligned}\mu(X) &= EX = \int_{\Omega} X(\omega) dP(\omega), \\ \sigma^2(X) &= E[X - EX]^2 = \int_{\Omega} [X(\omega) - \mu(X)]^2 dP(\omega).\end{aligned}\tag{1}$$

To assist in working with constant r.v.’s,  $X(\omega) \equiv C$ , the letter  $C$  will always denote a constant in the real numbers  $\mathbb{R}$ .

By a *deviation measure* will be meant any functional  $\mathcal{D}$  that assigns to each random variable  $X$  (understood to be in  $\mathcal{L}^2(\Omega)$  always) a value  $\mathcal{D}(X)$  in accordance with the following axioms:

- (D1)  $\mathcal{D}(X + C) = \mathcal{D}(X)$ ; equivalently,  $\mathcal{D}(X) = \mathcal{D}(X - EX)$  for all  $X$ ,
- (D2)  $\mathcal{D}(0) = 0$ , and  $\mathcal{D}(\lambda X) = \lambda \mathcal{D}(X)$  for all  $X$  and  $\lambda \geq 0$ ,
- (D3)  $\mathcal{D}(X + X') \leq \mathcal{D}(X) + \mathcal{D}(X')$  for all  $X$  and  $X'$ ,
- (D4)  $\mathcal{D}(X) > 0$  for nonconstant  $X$ , whereas  $\mathcal{D}(X) = 0$  for constant  $X$ .

A deviation measure  $\mathcal{D}$  will be called *coherent* if, in addition to these four axioms, it satisfies

- (D5)  $\mathcal{D}(X) \leq EX - \inf X$  for all  $X$ .

These axioms come from our paper [23], where the notion of a general deviation measure was first formulated at this level.<sup>5</sup> The equivalence in D1 is evident from taking  $C$  equal to  $-EX$ , and on the other hand, noting that  $[X + C] - E[X + C] = X - EX$  for any constant  $C$ . Clearly D4 puts the focus of  $\mathcal{D}$  on the uncertain part of an r.v.  $X$  and insists on this part not going undetected. The combination of D2 and D3 implies the *convexity* of  $\mathcal{D}$ , which is a key property in all contexts of optimization and makes the tools of convex analysis available to our endeavor.<sup>6</sup> We speak of the additional property in D5 as the *lower-range boundedness* of  $\mathcal{D}$ .

<sup>4</sup>Our choice of  $\mathcal{L}^2(\Omega)$  is dictated by the need for a framework in which the magnitude of  $X - EX$  can be measured by general deviations, with standard deviation as a particular case. In [5] and [7], the expectation of random variable had no required role, and the focus instead was on  $\mathcal{L}^\infty(\Omega)$ .

<sup>5</sup>In [13], a class of measures was described axiomatically in terms of D1, D2 (effectively), and the version of D4 requiring only weak inequality, but no D3. Such measures lack convexity and other crucial properties.

<sup>6</sup>In the presence of that convexity, D1 is actually implied by the seemingly much weaker condition that  $\mathcal{D}(C) = 0$  for all constants  $C$ .

Coherency will be discussed later in this section. Its chief effect mathematically is to make available, through duality, certain probabilistic interpretations of deviation that can be very helpful in analysis and in modeling as well. Such considerations also lie behind our focus on properties of  $\mathcal{D}$  with respect to all  $\mathcal{L}^2(\Omega)$ , instead of just a subspace  $\mathcal{X}$  of  $\mathcal{L}^2(\Omega)$  that generated by some particular collection of random variables.

The example of standard deviation,  $\mathcal{D}(X) = \sigma(X)$ , dominates classical portfolio theory and is *symmetric* in the sense that  $\mathcal{D}(-X) = \mathcal{D}(X)$ . Related examples of nonsymmetric deviation measures that satisfy our axioms include the standard *semideviations*  $\mathcal{D}(X) = \sigma_+(X)$  and  $\mathcal{D}(X) = \sigma_-(X)$ , where

$$\sigma_-^2(X) = E[\max\{EX - X, 0\}^2], \quad \sigma_+^2(X) = E[\max\{X - EX, 0\}^2]. \quad (2)$$

The first of these emphasizes the downside of  $X$ , while the second emphasizes the upside. A very different pair of examples, likewise oriented to downside or upside, is furnished by the *lower range* and the *upper range*,

$$\mathcal{D}(X) = EX - \inf X, \quad \mathcal{D}(X) = \sup X - EX, \quad (3)$$

where  $\inf X$  and  $\sup X$  denote the “essential” infimum and supremum of  $X(\omega)$  over  $\omega \in \Omega$  (obtained by disregarding subsets of  $\Omega$  having probability 0). For either of these, it is possible for some r.v.’s  $X$  that  $\mathcal{D}(X) = \infty$ , which is allowed by the axioms. Of course, both are sure to be finite in the case of a finite, discrete probability space  $\Omega$ .

Another class of deviation measures, of increasing interest now in applications, arises from *conditional value-at-risk*, CVaR. A brief discussion of risk measures, in contrast to deviation measures, will lay the platform for introducing this class properly.

By a *strictly expectation bounded risk measure* will be meant any functional  $\mathcal{R}$  that assigns values  $\mathcal{R}(X)$  to random variables  $X$  in such a way that

- (R1)  $\mathcal{R}(X + C) = \mathcal{R}(X) - C$  for all  $X$  and constants  $C$ ,
- (R2)  $\mathcal{R}(0) = 0$ , and  $\mathcal{R}(\lambda X) = \lambda \mathcal{R}(X)$  for all  $X$  and all  $\lambda > 0$ ,
- (R3)  $\mathcal{R}(X + X') \leq \mathcal{R}(X) + \mathcal{R}(X')$  for all  $X$  and  $X'$ ,
- (R4)  $\mathcal{R}(X) > E[-X]$  for all nonconstant  $X$ , whereas  $\mathcal{R}(X) = E[-X]$  for constant  $X$ .

A strictly expectation bounded risk measure  $\mathcal{R}$  will be called *coherent* if, in addition to these four axioms, it satisfies

- (R5)  $\mathcal{R}(X) \leq \mathcal{R}(X')$  when  $X \geq X'$ .

Axiom R4 is the property we explicitly mean by “strict expectation boundedness.” The equation part of R4 is already a consequence of R1, so the *strict* inequality for *nonconstant*  $X$  is the chief assertion. (In [23], we spoke of this simply as expectation boundedness, but now add “strict” as a safeguard against misunderstandings.)<sup>7</sup>

Artzner, Delbaen, Eber and Heath in their landmark contribution to risk theory in [5] were the first to consider risk measures from a broad perspective, but they concentrated instead on functionals  $\mathcal{R}$  satisfying R1, R2, R3 and, instead of R4, the monotonicity axiom R5. They called them *coherent risk measures*. (Actually, they posed R5 in a seemingly weaker form, namely  $\mathcal{R}(X) \leq 0$  when  $X \geq 0$ , which is equivalent to the present R5 under the other axioms. Also, they had a somewhat different version of R1, tailored to the use of an investment instrument, but the version used here was subsequently adopted by Delbaen [7].) Property R5 is natural and even critical for many purposes, and we fully believe in its importance. However, we forgo it in our basic *definition* of a strictly expectation bounded risk measure in order to capture a fundamental pairing between risk measures and deviation measures.

---

<sup>7</sup>The weak-inequality version has recently been dubbed “risk relevance” in [14].

Risk measure axiom R4, first spelled out in our own earlier work, is needed for the following result, where it emerges as the counterpart to deviation axiom D4. As background, it should be noted that the weak inequality  $\mathcal{R}(X) \geq E[-X]$ , in place of the strict one in R4, is known to hold always when  $\mathcal{R}$  is a coherent risk measure that depends only on the distribution of an r.v. and in addition has a certain convergence property, and furthermore  $X$  is an r.v. having no atoms in its distribution; cf. [8, Lemma 4.45]. That comes nowhere near to furnishing R4 in our general setting, however, so this really is a property that has to be brought in separately.

**Theorem 1** [23] (deviation vs. risk). *Deviation measures correspond one-to-one with strictly expectation bounded risk measures under the relations*

- (a)  $\mathcal{D}(X) = \mathcal{R}(X - EX)$ ,
- (b)  $\mathcal{R}(X) = E[-X] + \mathcal{D}(X)$ .

*Specifically, if  $\mathcal{R}$  is an strictly expectation bounded risk measure and  $\mathcal{D}$  is defined by (a), then  $\mathcal{D}$  is a deviation measure that yields back  $\mathcal{R}$  through (b). On the other hand, if  $\mathcal{D}$  is any deviation measure and  $\mathcal{R}$  is defined by (b), then  $\mathcal{R}$  is a risk measure that yields back  $\mathcal{D}$  through (a). In this correspondence,  $\mathcal{R}$  is coherent if and only if  $\mathcal{D}$  is coherent.*

The fact in the last part of Theorem 1 explains, of course, why we introduced D5 and used it to define coherency as a concept for deviation measures. We did not require this property in the basic definition of a deviation measure, because that would have left out standard deviation,  $\mathcal{D}(X) = \sigma(X)$ , thereby making nonsense of the terminology and excluding classical portfolio theory from our setting. The cogent arguments in favor of coherency made in [5] and elsewhere are nonetheless a prime motivation for what we are undertaking here.

In the pairing of Theorem 1, the deviation measure  $\mathcal{D}(X) = \rho\sigma(X)$  for any  $\rho \in (0, \infty)$  corresponds to  $\mathcal{R}(X) = \rho\sigma(X) - \mu(X)$ , whereas the lower-range deviation measure  $\mathcal{D}(X) = EX - \inf X$  corresponds to the *maximum loss* risk measure  $\mathcal{R}(X) = \sup[-X]$ . Coherency is present in the second example, although absent in the first.

Even more of interest here is the utilization of Theorem 1 in deriving deviation measures of CVaR type. Recall that for any  $\alpha \in (0, 1)$  the *value-at-risk* of  $X$  at level  $\alpha$  is defined by

$$\text{VaR}_\alpha(X) = -\inf\{z \mid P\{X \leq z\} > \alpha\}. \quad (4)$$

The corresponding *conditional value-at-risk* is then

$$\text{CVaR}_\alpha(X) = -[\text{expectation of } X \text{ in its lower } \alpha\text{-tail distribution}], \quad (5)$$

where the expectation is the same as the conditional expectation of  $X$  subject to  $X \leq -\text{VaR}_\alpha(X)$  when  $P\{X \mid X = -\text{VaR}_\alpha(X)\} = 0$ , but in general refers to the expectation of the r.v. whose cumulative distribution function  $F_\alpha$  is obtained from the cumulative distribution function  $F$  for  $X$  by taking  $F_\alpha(z) = F(z)/\alpha$  when  $z < -\text{VaR}_\alpha(X)$  and  $F_\alpha(z) = 1$  when  $z \geq -\text{VaR}_\alpha(X)$ .<sup>8</sup> The important thing is that a *coherent* risk-deviation pair is obtained by taking

$$\mathcal{R}(X) = \text{CVaR}_\alpha(X), \quad \mathcal{D}(X) = \text{CVaR}_\alpha(X - EX). \quad (6)$$

The functional  $\mathcal{R}(X) = \text{VaR}_\alpha(X)$ , in contrast, fails to satisfy axioms R3, R4 and R5, so correspondingly the functional  $\mathcal{D}(X) = \text{VaR}_\alpha(X - EX)$  fails to satisfy D3, D4 and D5 and is not a deviation

---

<sup>8</sup>This form of the definition of  $\text{CVaR}_\alpha(X)$  corresponds to the development of the concept in [22]; earlier, in [21], we concentrated only on the case of continuous distribution functions. Acerbi [1] independently arrived at this risk measure by an integral formula, calling it “expected shortfall,” and that term has subsequently been used also in [3, 28]. A alternative minimization expression for  $\text{CVaR}_\alpha(X)$  in [21, 22] provides a powerful approach to computations. Other background on VaR and CVaR can be found in [8].

measure in our sense, much less a coherent one. Worse, it lacks the property of *convexity* that is crucial to our developments.

Beyond “pure” measures of CVaR type that come out of the correspondence in Theorem 1, there also “mixed” CVaR measures cf. [22, 23] having a “spectral representation” as in [1].

A final example, which deserves mention because of its modeling implications and the theoretical insights it provides, is the one in which  $\Omega$  is just a finite set of future scenarios, with scenario  $\omega$  having probability  $P(\omega) > 0$ . Consider any collection  $\{P_j\}_{j \in J}$  of probability measures on  $\Omega$ ,

$$P_j(\omega) \geq 0, \quad \sum_{\omega \in \Omega} P_j(\omega) = 1, \quad (7)$$

for an index set  $J$  (finite or infinite). Denote the expectation with respect to  $P_j$  by  $E_j$ , so that  $E_j X = \sum_{\omega \in \Omega} X(\omega) P_j(\omega)$  in contrast to  $EX = \sum_{\omega \in \Omega} X(\omega) P(\omega)$ . Suppose the collection has the “richness” property that for every nonconstant  $X$  there is at least one  $j \in J$  such that  $E_j[-X] > E[-X]$ , i.e., the expected loss incurred by  $X$  would be worse if the probability measure were  $P_j$  instead of the reference probability measure  $P$ . Then a coherent risk-deviation pair is obtained by defining

$$\mathcal{R}(X) = \sup_{j \in J} E_j[-X], \quad \mathcal{D}(X) = \sup_{j \in J} \{E_j[-X] - E[-X]\}. \quad (8)$$

In this setting,  $P_j$  might be viewed as an alternative to  $P$  that an investor has selected for prudent comparisons, in case  $P$  just represents an educated guess about what the future will bring and may not be completely reliable. The deviation  $\mathcal{D}(X)$  expressed by (8) identifies the worst discrepancy that could occur between the expected losses under the specified alternatives and the nominal expected loss  $E[-X]$ .

Besides being of interest for practical purposes in portfolio optimization, deviation measures of the kind in (8) furnish a simple illustration of how differentiability can fail. Indeed, when the index set  $J$  is finite,  $\mathcal{D}$  is the pointwise maximum of a finite collection of linear functions of  $X$  and therefore is *piecewise linear* on  $\mathcal{L}^2(\Omega)$ . It fails to be differentiable on the joins between the different “pieces,” and its lower level sets  $\{X \mid \mathcal{D}(X) \leq c\}$  are *polyhedral* convex sets with ridges and flat sides.

Another way that differentiability can fail, apart from (8), will be demonstrated in the last section of the paper.

### 3 Portfolio Framework

To proceed with our effort to extend the classical results in portfolio theory for standard deviation to general deviation measure  $\mathcal{D}$ , we must provide a market setting. The market will be taken, for model purposes, to consist of instruments  $i = 0, 1, \dots, n$  having rates of return  $r_i$ . The first of these instruments, for  $i = 0$ , is risk-free; its rate of return  $r_0$  is a constant. The other instruments, for  $i = 1, \dots, m$ , are risky; their rates of return  $r_i$  are r.v.’s in  $\mathcal{L}^2(\Omega)$ . A dollar invested in instrument  $i$  brings back  $1 + r_i$ , for a gain (or profit) of  $r_i$  dollars at the end of the time period under consideration.

We will be concerned with portfolios that can be put together by investing an amount  $x_i$  in each instrument  $i$ . These amounts, which we can take to be in dollars,<sup>9</sup> can be positive, zero or negative. (A negative investment corresponds to a short position.) Such a portfolio has the present

---

<sup>9</sup>More typically, these amounts are viewed as “fractions,” even if greater than 1, but this interpretation will facilitate discussion of portfolio prices.

price  $x_0 + x_1 + \dots + x_n$  and the uncertain future value  $x_0(1 + r_0) + x_1(1 + r_1) + \dots + x_n(1 + r_n)$ . The associated gain is thus the r.v.  $X$  in  $\mathcal{L}^2(\Omega)$  described by

$$X = x_0r_0 + x_1r_1 + \dots + x_nr_n. \quad (9)$$

Here we are using “gain” in the sense that a loss is a negative gain. Costs, too, might be negative as well as positive, or zero.

To facilitate our work with these r.v.’s  $X$  while taking into account the special role of the risk-free instrument and keeping notation simple, we introduce

$$r = (r_1, \dots, r_n) \text{ (vector r.v.)}, \quad \bar{r} = (\bar{r}_1, \dots, \bar{r}_n) \text{ for } \bar{r}_i = Er_i,$$

along with the vectors

$$x = (x_1, \dots, x_n), \quad e = (1, \dots, 1).$$

The general r.v.  $X$  in (9) is then

$$X = x_0r_0 + x^\top r, \text{ with expected gain } x_0r_0 + x^\top \bar{r} \text{ and price } x_0 + x^\top e.$$

We speak of  $x = (x_1, \dots, x_n)$  itself as giving the  $x$ -portfolio for which the gain is the r.v.  $x^\top r$  in  $\mathcal{L}^2(\Omega)$ , the expected gain is  $x^\top \bar{r}$ , and the price is  $x^\top e$ .

The following assumptions on instruments  $i = 1, \dots, n$  in the model will henceforth be in effect. The rest of this section will be devoted to elucidating their immediate consequences.

### Basic Assumptions.

- (A1) No  $x$ -portfolio with  $x \neq 0$  is risk-free.
- (A2) The expected rates of return  $\bar{r}_1, \dots, \bar{r}_n$  are not all the same.
- (A3)  $\mathcal{D}(r_i) < \infty$  and  $\mathcal{D}(-r_i) < \infty$  for all  $i$ .

Assumption A1 is harmless and merely underscores our aim of letting the  $i = 0$  instrument do all the risk-free service. A notion of redundancy will help in understanding why this is true. Let us say that an instruments  $i$  is *redundant* in the model if the associated r.v.  $r_i$ , which gives the gain from investing one dollar in instrument  $i$ , can exactly be replicated by the gain r.v. of a portfolio put together from the other instruments. Note that such replication, if possible at all, would have to be achieved at price 1, or an arbitrage opportunity would exist, thereby undermining our intent of starting from a market in which prices are in equilibrium.

**Proposition 1** (elimination of redundancy). *Assumption A1 is fulfilled if and only if none of the instruments  $i$  in the model is redundant.*

**Proof.** If some  $x$ -portfolio with  $x \neq 0$  were risk-free, we could find a value  $x_0$  such that the r.v.  $X = x_0r_0 + x_1r_1 + \dots + x_nr_n$  is identically 0. One of the coefficients  $x_1, \dots, x_n$  would be nonzero; suppose for purposes of illustration that is  $x_1$ . We would then have  $r_1 = x'_0r_0 + x'_2r_2 + \dots + x'_nr_n$  for  $x'_i = -x_i/x_1$ , which would mean that the  $i = 1$  instrument is redundant.

For the converse, suppose some instrument  $i$  is redundant. If that holds for  $i = 0$ , then by the definition of redundancy there must be a nonzero  $x$ -portfolio that is risk-free. Otherwise, we can suppose for simplicity of notation that  $i = 1$  is redundant. This refers to the existence of coefficients  $x_0, x_2, \dots, x_n$  such that  $r_1 = x_0r_0 + x_2r_2 + \dots + x_nr_n$ . Then  $-r_1 + x_2r_2 + \dots + x_nr_n = -x_0r_0$ , so the  $x$ -portfolio for  $x = (-1, x_2, \dots, x_n)$  would be risk-free.  $\square$

Redundant instruments offer nothing new, so we could always eliminate them from the model one by one until nothing redundant was left. Then A1 would hold.

Another insight into A1 can be obtained through consideration of distribution functions.



**Proposition 2** (continuous distributions). *Assumption A1 is satisfied when the r.v.  $r$  is continuously distributed (i.e., the multivariate distribution function for  $r_1, \dots, r_n$  is continuous on  $\mathbb{R}^n$ ), thus guaranteeing that the gain  $x^\top r$  of any  $x$ -portfolio with  $x \neq 0$  is continuously distributed as well.*

**Proof.** The well known fact about  $x^\top r$  being continuously distributed in these circumstances precludes  $x^\top r$  from being a constant r.v., of course.  $\square$

Assumption A2 is needed to sidestep special circumstances which have little interest for us here. If it did not hold, there would be a value  $\rho$  such that  $\bar{r}_i = \rho$  for  $i = 1, \dots, n$ ; then the expected gain  $x^\top \bar{r}$  of an  $x$ -portfolio would always be  $\rho$  times its price  $x^\top e$ .

Both A1 and A2 seem to be taken for granted by many in finance, even though they are essential to the validity of commonly made assertions.<sup>10</sup> The desire to maintain mathematical rigor in our development of portfolio theory requires us to make these assumptions, and others, explicit.

**Proposition 3** (richness of price-gain combinations). *For every choice of  $(\pi, \zeta) \in \mathbb{R}^2$ , there is an  $x$ -portfolio having price  $x^\top e = \pi$  and expected gain  $x^\top \bar{r} = \zeta$ .*

**Proof.** This is the main consequence of A2. The set of pairs  $(\pi, \zeta)$  coming from portfolios in this way constitutes a subspace of  $\mathbb{R}^2$ , so if it were not all of  $\mathbb{R}^2$ , these pairs would be collinear, and we would be in the lockstep situation excluded by A2.  $\square$

Assumption A3, which will guarantee the finiteness of  $\mathcal{D}(x^\top r)$  according to the next proposition, is certainly satisfied when  $\mathcal{D}$  is a deviation measure that is finite on all of  $\mathcal{L}^2(\Omega)$ , and many measures with that property have already been indicated beyond  $\mathcal{D}(X) = \sigma(X)$ , a major example being  $\mathcal{D}(X) = \text{CVaR}_\alpha(X - EX)$ . But A3 may also be satisfied for some deviation measures that are not finite on all of  $\mathcal{L}^2(\Omega)$ . An example is  $\mathcal{D}(X) = EX - \inf X$  when the rates of return  $r_i$  are bounded. Note that we are obliged to require the finiteness of  $\mathcal{D}(r_i)$  and  $\mathcal{D}(-r_i)$  separately, because  $\mathcal{D}$  need not be symmetric.

**Proposition 4** (portfolio deviations). *The deviation function*

$$f_{\mathcal{D}}(x) = \mathcal{D}(x^\top r)$$

*is finite everywhere and convex on  $\mathbb{R}^n$  (hence also continuous), moreover with the properties that*

- (a)  $f_{\mathcal{D}}(0) = 0$ , but  $f_{\mathcal{D}}(x) > 0$  when  $x \neq 0$ ,
- (b)  $f_{\mathcal{D}}(\lambda x) = \lambda f_{\mathcal{D}}(x)$  when  $\lambda > 0$ ,
- (c)  $f_{\mathcal{D}}(x + x') \leq f_{\mathcal{D}}(x) + f_{\mathcal{D}}(x')$ ,
- (d)  $\{x \mid f_{\mathcal{D}}(x) \leq \delta\}$  is a bounded set for every  $\delta > 0$ .

**Proof.** In view of axiom D4 on  $\mathcal{D}$ , the strict inequality in (a) is equivalent to A1. Properties (b) and (c), together with the fact in (a) that  $f_{\mathcal{D}}(0) = 0$ , follow immediately from axioms D2 and D3 on  $\mathcal{D}$ . They imply in particular that  $f_{\mathcal{D}}$  is a convex function. The set of  $x$  for which  $f_{\mathcal{D}}(x) < \infty$  is then a convex subset of  $\mathbb{R}^n$ . Because of A3, that set includes the vectors,

$$(\pm 1, 0, \dots, 0), (0, \pm 1, \dots, 0), (0, 0, \dots, \pm 1),$$

which correspond to portfolios consisting of just one of the instruments  $i = 1, \dots, n$ , either in unit long position or unit short position. It must also then include all positive multiples of those vectors, through (b), as well as all sums generated from those, through (c). Thus, it has to be all of  $\mathbb{R}^n$ .

<sup>10</sup>For instance, in the text [12, p. 159], the  $n + 2$  linear equations in  $n + 2$  unknowns that describe the weights for an efficient risky portfolio in the Markowitz model are said to have a unique solution, but really that is only true when the coefficient matrix is nonsingular. The matrix in question fails to be nonsingular if A1 and A2 do not hold.

For the fact that finite convex functions on  $\mathbb{R}^n$  are continuous, see [19, Theorem 10.1]. On the principle of [19, Corollary 8.7.1] and  $f_{\mathcal{D}}$  being convex and continuous, if any set of form  $\{x \mid f_{\mathcal{D}}(x) \leq \delta\}$  is bounded, then all sets of that form must be bounded. By (a), the set  $\{x \mid f_{\mathcal{D}}(x) \leq 0\}$  is the singleton  $\{0\}$ , so (d) is correct.  $\square$

Readers familiar with the emphasis on VaR in much of finance nowadays should note Proposition 4 is not applicable to  $f_{\mathcal{D}}(x) = \mathcal{D}(x^\top r)$  when  $\mathcal{D}(x^\top r) = \text{VaR}_\alpha(x^\top r - x^\top \bar{r})$ , because the expression  $\mathcal{D}(X) = \text{VaR}(X - EX)$  does not fit our axioms for a deviation measure (obeying only D1 and D2, not D3, D4 or D5). Certainly (c) of Proposition 4 would be gone, and with it the convexity of this  $f_{\mathcal{D}}$ . Even the strict inequality in (a) would be violated in some situations, and the boundedness of the level sets in (d) would then be lost. Moreover,  $f_{\mathcal{D}}$  could no longer even be counted on to be continuous.

## 4 Basic Funds and Master Funds

The fundamental problem of optimization that we wish to examine closely with respect to the portfolio r.v.'s  $X$  in (9) is

$$\mathcal{P}(\Delta) \quad \text{minimize } \mathcal{D}(x_0 r_0 + x^\top r) \quad \text{subject to } x_0 + x^\top e = 1 \text{ and } x_0 r_0 + x^\top \bar{r} \geq r_0 + \Delta,$$

where  $\mathcal{D}(x_0 r_0 + x^\top r)$  is actually just the deviation  $f_{\mathcal{D}}(x)$  in Proposition 4, of course. The budget constraint  $x_0 + x^\top e = 1$  signifies (in our mode of interpreting the  $x_i$ 's as dollar amounts) that the price of the portfolio nets out to exactly one dollar; that is how much is to be invested initially. The gain constraint  $x_0 r_0 + x^\top \bar{r} \geq r_0 + \Delta$  requires that this unit investment should result in an expected future value of at least  $1 + r_0 + \Delta$  dollars. The parameter  $\Delta$  gives the *risk premium* — the extra amount being demanded over the gain associated with investing at the risk-free rate  $r_0$ . The gain constraint has been written as an inequality instead of an equation because there should not be any objection if some portfolio, without worsening the deviation or costing more, might have an expected gain that is more than  $r_0 + \Delta$ . It will come out below, however, that any portfolio solving problem  $\mathcal{P}(\Delta)$  must satisfy this constraint with equality, when  $\Delta > 0$ .

The unit price constraint in  $\mathcal{P}(\Delta)$  can be used to eliminate  $x_0$  by assigning it the value  $x_0 = 1 - x^\top e$ . The problem statement comes down then to:

$$\mathcal{P}_0(\Delta) \quad \text{minimize } f_{\mathcal{D}}(x) \quad \text{subject to } x^\top [\bar{r} - r_0 e] \geq \Delta.$$

Adopting this framework in terms of  $x$ -portfolios alone, we let

$$\begin{cases} d_0(\Delta) = & \text{optimal value (the infimum of the deviation) in } \mathcal{P}_0(\Delta), \\ S_0(\Delta) = & \text{optimal solution set (the minimizing vectors } x \text{) in } \mathcal{P}_0(\Delta). \end{cases} \quad (10)$$

**Proposition 5** (solution existence and homogeneity). *An optimal solution to problem  $\mathcal{P}_0(\Delta)$  is sure to exist (not necessarily uniquely), no matter what the choice of  $\Delta$ . Indeed, the optimal solution set  $S_0(\Delta)$  is always convex, closed and bounded, in addition to being nonempty. Moreover,*

$$\begin{cases} \text{for } \Delta \leq 0 : & d_0(\Delta) = 0 \text{ and } S_0(\Delta) = \{0\} \text{ (put all in the risk-free instrument),} \\ \text{for } \Delta > 0 : & d_0(\Delta) > 0, \text{ with } d_0(\Delta) = \Delta \cdot d_0(1) \text{ and } S_0(\Delta) = \{\Delta \cdot x \mid x \in S_0(1)\}. \end{cases} \quad (11)$$

*Additionally, when  $\Delta > 0$  the gain constraint is always active in  $\mathcal{P}_0(\Delta)$ , i.e., every  $x \in S_0(\Delta)$  satisfies  $x^\top [r - r_0 e] = \Delta$ .*

**Proof.** In view of Proposition 3, the constraint in problem  $\mathcal{P}_0(\Delta)$  can be satisfied regardless of the choice of  $r_0$  and  $\Delta$ . The sets of type

$$\{x \mid f_{\mathcal{D}}(x) \leq \delta, x^\top[r - r_0 e] \geq \Delta\} \quad \text{for } \delta > d_0(\Delta) \quad (12)$$

are nonempty by the definition of  $d_0(\Delta)$  as well as compact because of the continuity of  $f_{\mathcal{D}}$  and the boundedness in Proposition 4(d). Any nest of nonempty compact sets has a nonempty intersection. In this case, moreover, the sets are convex by virtue of the convexity of  $f_{\mathcal{D}}$ , so the intersection is likewise a convex set. This confirms that  $S_0(\Delta)$  is nonempty, convex and compact.

The special assertions about  $\mathcal{P}_0(\Delta)$  in the case of  $\Delta \leq 0$  are evident from Proposition 4(a). They rely also on the constraint having been stated as an inequality rather than an equation. In the case where  $\Delta > 0$ , the relationships involving  $d_0(1)$  and  $S_0(1)$  are immediate from the positive homogeneity of  $f_{\mathcal{D}}$  in Proposition 4(b), according to which anything optimal for  $\Delta = 1$  can be rescaled to be optimal for other  $\Delta$ .

The constraint in  $\mathcal{P}_0(\Delta)$  has to be active when  $\Delta > 0$ , because if  $x$  has  $x^\top[r - r_0 e] > \Delta$ , there is a factor  $\theta \in (0, 1)$  such that the vector  $x' = \theta x$  satisfies the same inequality and yet yields a deviation amount that is smaller than the one for  $x$  by the same factor. This is incompatible with  $x$  being optimal. Note that here we are invoking Proposition 4(a) once more, since this argument would fall through if the deviation in question were 0.  $\square$

Proposition 5 sets the stage for a complete understanding of the efficient frontier for portfolios that include the risk-free instrument,  $i = 0$ . It will lead quickly to the classical concept of a “master fund,” with certain extensions. However, for the sake of capturing at once what it says about the best way to invest, regardless of the technical complications that will soon come up, another “fund” concept will be helpful.

**Definition 1** (basic funds and basic deviation value). *For any  $\bar{x} \in S_0(1)$ , providing in problem  $\mathcal{P}_0(1)$  the minimum portfolio deviation for a gain of exactly 1 over the risk-free rate, the  $\bar{x}$ -portfolio will be said to furnish a basic fund. The minimum deviation amount will be called the basic deviation value and denoted by  $\bar{\delta}$ :*

$$\bar{\delta} = d_0(1) = f_{\mathcal{D}}(\bar{x}).$$

**Theorem 2** (generalized one-fund theorem in basic fund form). *Let  $\bar{x}$  furnish a basic fund, achieving the basic deviation value  $\bar{\delta}$ . Then, for any  $\Delta > 0$ , a solution to the fundamental portfolio problem  $\mathcal{P}(\Delta)$  is obtained by investing the amount  $\Delta[\bar{x}^\top e]$  in the  $\bar{x}$ -portfolio and the amount  $1 - \Delta[\bar{x}^\top e]$  in the risk-free instrument. That solution portfolio has deviation  $\Delta \cdot \bar{\delta}$ . In further detail, this prescription for investment comes down to the following three cases, depending on the price  $\bar{x}^\top e$  of the basic fund:*

(a) Positive case: price  $\bar{x}^\top e > 0$ . Invest the positive amount  $\Delta[\bar{x}^\top e]$  in the  $\bar{x}$ -fund while investing the amount  $1 - \Delta[\bar{x}^\top e]$  (possibly positive, negative or zero) in the risk-free instrument.

(b) Negative case: price  $\bar{x}^\top e < 0$ . Invest the negative amount  $\Delta[\bar{x}^\top e]$  in the  $\bar{x}$ -fund (i.e., take a short position of this magnitude in it), while investing the positive amount  $1 - \Delta[\bar{x}^\top e] > 1$  in the risk-free instrument.

(c) Threshold case: price  $\bar{x}^\top e = 0$ . Invest the original 1 dollar entirely in the risk-free instrument while assuming a position in  $\bar{x}$  of magnitude  $\Delta$ , which nets out (through longs and shorts) to price 0.

The three cases arise, in principle at least, because there is no *a priori* restriction in our optimization problems on the prices of the  $x$ -portfolios under consideration. This price is left to the optimization outcome itself. Understanding the extent to which these different cases can truly occur, or conceivably even overlap, will require serious effort in the rest of this paper. Some insights are immediately available, however.

In the positive case, one is investing in risky instruments in the classical way and making up the difference with the original 1 dollar by buying, or borrowing, a quantity of the risk-free instrument, as needed. In the negative case, the scheme effectively involves borrowing from the market in the risky instruments and putting the proceeds into the risk-free instrument. Clearly, that would not be interesting unless the risk-free return  $r_0$  is attractively high from the investor's point of view. Much more will be said about this circumstance as we go on, but observe that it does not imply the presence of some sort of arbitrage which would bring the market equilibrium into question, since it depends on the particular deviation measure  $\mathcal{D}$  being employed and not on the market as a whole. In the threshold case, despite the price 0 of taking on the  $\Delta$  position in the basic fund, that position does yield an expectation of  $\Delta$  dollars above  $r_0$ . Again, no arbitrage has to be entailed.

Theorem 2 already has many of the features associated with the classical one-fund theorem. A single fund, which has been dubbed a basic fund, serves along with the risk-free instrument to provide an optimal response to the problem of finding a deviation-minimizing portfolio, regardless of the size of the extra expected gain  $\Delta$  that an investor might demand over the risk-free rate  $r_0$ . Regardless of the price that the basic fund  $\bar{x}$  turns out to have, it completely determines the efficient frontier for portfolios including the risk-free instrument. This frontier is the upward-sloping half-line in Figure 1, where the slope is  $1/\bar{\delta}$  for the basic deviation value  $\bar{\delta}$  in Definition 1. In that way, Theorem 2 fully answers the question of how investors should act in this idealized setting.

What is missing from Theorem 2 is a connection with the traditional notion of a master fund and its geometric characterization through a picture of tangency. That will occupy us shortly, but anyway we can easily come up with a definition of what master funds must be, in relation to basic funds.

**Definition 2** (master funds).

- (a) An  $x^*$ -portfolio furnishes a master fund of positive type if  $x^{*\top}e = 1$  and  $x^* \in S_0(\Delta^*)$  for some  $\Delta^* > 0$ , or equivalently, if  $x^* = \Delta^* \cdot \bar{x}$  for a basic  $\bar{x}$ -portfolio with price  $\bar{x}^\top e > 0$ , and  $\Delta^* = 1/\bar{x}^\top e$ .
- (b) An  $x^*$ -portfolio furnishes a master fund of negative type if  $x^{*\top}e = -1$  and  $x^* \in S_0(\Delta^*)$  for some  $\Delta^* > 0$ , or equivalently, if  $x^* = \Delta^* \cdot \bar{x}$  for a basic  $\bar{x}$ -portfolio with price  $\bar{x}^\top e < 0$ , and  $\Delta^* = -1/\bar{x}^\top e$ .

A master fund of positive type thus represents a unit long position with respect to a certain weighted collection of risky instruments, whereas a master fund of negative type represents a unit short position with respect to some such collection, perhaps different.

Because the master funds in this definition are just *positively rescaled versions of basic funds*, they fit equally well into the one-fund “algorithm” of Theorem 2 for generating the efficient frontier for portfolios including the risk-free instrument — as long as there is a basic fund with nonzero price. This can be summarized as follows.

**Theorem 3** (one-fund theorem in master fund form).

- (a) Suppose a master fund of positive type exists, furnished by an  $x^*$ -portfolio that yields an expected return  $r_0 + \Delta^*$  for some  $\Delta^* > 0$ . Then, for any  $\Delta > 0$ , a solution to the fundamental portfolio problem  $\mathcal{P}(\Delta)$  is obtained by investing the positive amount  $\Delta/\Delta^*$  in the master fund and the amount  $1 - (\Delta/\Delta^*)$  (possibly positive, negative or zero) in the risk-free instrument.
- (b) Suppose a master fund of negative type exists, furnished by an  $x^*$ -portfolio that yields an expected return  $r_0 + \Delta^*$  for some  $\Delta^* > 0$ . Then, for any  $\Delta > 0$ , a solution to the fundamental portfolio problem  $\mathcal{P}(\Delta)$  is obtained by investing the negative amount  $-\Delta/\Delta^*$  in the master fund and the amount  $1 + (\Delta/\Delta^*) > 1$  in the risk-free instrument.

When no basic fund with nonzero price exists, there is nothing to be done about price rescaling, so the concept of a master fund no longer makes sense. Of course the rule for optimal investment in Theorem 2 nevertheless survives, covering the threshold case as well as the positive and negative

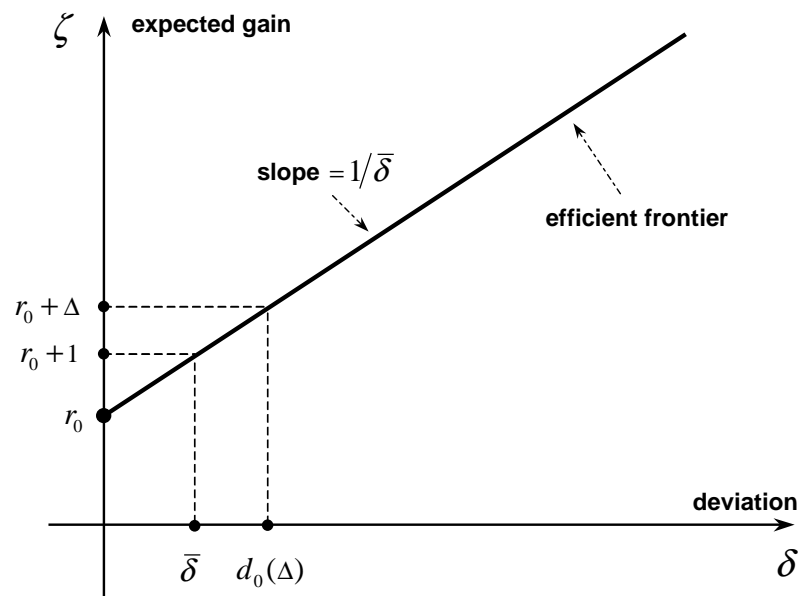


Figure 1: Efficient frontier for portfolios with the risk-free instrument.

cases. A master fund of positive type can be seen to correspond to some point on the half-line frontier in Figure 1, different in general from the point coming from the basic fund.

In the classical theory, based on  $\mathcal{D} = \sigma$ , only master funds of the positive type are contemplated, and they are immediately tied into the notion of “efficient set.” The familiar picture in Figure 2 is used to indicate that, for  $\Delta > 0$ , the optimal value in the fundamental problem  $\mathcal{P}(\Delta)$ , which is the same as  $d_0(\Delta)$ , can be obtained by interpolating along a tangent line to the efficient set that passes through the point  $(0, r_0)$ . The point of tangency corresponds to an  $x^*$ -portfolio of price 1 having expected gain  $\zeta^* = r_0 + \Delta^*$  for some  $\Delta^* > 0$ . This  $x^*$ -portfolio then solves problem  $\mathcal{P}_0(\Delta^*)$  and furnishes a master fund with positive unit price which can generate entire the half-line “frontier” in the manner already reviewed.

In our situation of nonstandard deviation, there is a need to look much more closely at this picture and recognize certain shortcomings as well as major challenges. In the first place, the “efficient set” corresponding analogously to a general deviation measure  $\mathcal{D} \neq \sigma$  may no longer be a quadratic curve like the one in Figure 2 (which is actually a hyperbola). For that reason, the one-fund theorem must contend with major complications.<sup>11</sup> Although the region marked out by the efficient set will continue to be convex in our generalized setting, its boundary may incorporate corners or straight segments. For the case of a corner, the very meaning of tangency has to be pinned down carefully. In the presence of straight segments, the tangent line could have a whole interval in common with the efficient set, and this might in fact be an infinite interval.

Another complication, which can also come up in the classical model, is the possibility that, because of the asymptotic behavior of the efficient set, the efficient set has no “tangent” line at all that passes through  $(0, r_0)$ . Indeed, apart from any troubles with asymptotic behavior, there is an unspoken difficulty in the classical picture over the fact that it typically takes for granted the *existence* of a master fund of positive type. Clearly this existence, perceived in relation to “tangency,” depends in particular on the rate  $r_0$ . It might fail if  $r_0$  were too high.

In the traditional setting with  $\mathcal{D} = \sigma$ , such a confrontation with high  $r_0$  seems mostly to have been regarded as implausible and anyway “incompatible with market equilibrium.”<sup>12</sup> This view appears to originate in traditional CAPM (capital asset pricing model) considerations and the supposition that *all* investors are effectively engaged in minimizing standard deviation. A master fund could not be of negative type, for the reason that if all investors wanted to take a net short position as represented by a certain portfolio, so as to obtain money to invest at the risk-free rate, something must be wrong with the risk-free rate — an implicit market instability.

None of that really applies to our setting, however, because we are only exploring portfolio optimization for a subclass of investors, those who choose the particular deviation measure  $\mathcal{D}$  we are focussing on. Other investors, with different measures  $\mathcal{D}$ , can be expected to come to different conclusions about their portfolio choices. Some may end up with net short positions, while others may not. From that angle, there is no hint of conflict with market equilibrium in thinking about a master fund of negative type possibly emerging from a particular choice of  $\mathcal{D}$  at some level of  $r_0$ .

We are compelled, therefore, in our framework of a diversity of deviation measures  $\mathcal{D}$ , to face up to all price possibilities for  $x$ -portfolios as potential solutions to problem  $\mathcal{P}_0(\Delta)$ . This will lead us to study how such solutions may depend on  $r_0$  as a parameter.

<sup>11</sup>Without the curve being quadratic, there is no hope at all, by the way, of generalizing the classical “two-fund” theorem [29], which asserts the existence of two portfolios from which all efficient portfolios can be constructed as linear combinations. That result is intrinsically “quadratic” in its mathematical underpinnings.

<sup>12</sup>Huang and Litzenberger in [9] do allow for it by passing to a tangent line to the lower part of the curve in Figure 2 when  $r_0$  is high and interpret it as prescribing a short position in the classical master fund. But they leave out the details, which anyway would have to be tied to the special nature of the standard deviation measure, including its symmetry.

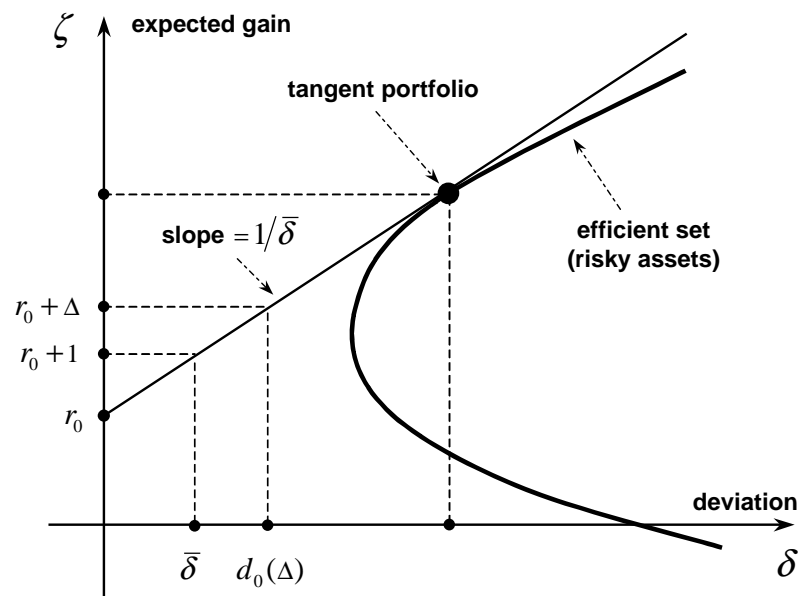


Figure 2: Classical efficient set: expected gain versus standard deviation.

Before getting into that parametric analysis, we can record a key fact about duality in  $\mathcal{P}_0(\Delta)$ . This problem is, after all, a convex programming problem, in which the convex function  $f_{\mathcal{D}}$  is minimized subject to a single linear constraint. The Lagrangian function is

$$L_{\Delta}(x, \lambda) = f_{\mathcal{D}}(x) + \lambda(\Delta - x^{\top}[\bar{r} - r_0e]), \quad \text{for } \lambda \geq 0, \quad (13)$$

and the problem *dual* to  $\mathcal{P}_0(\Delta)$  consists therefore of maximizing the function  $g_{\mathcal{D}}(\lambda) = \inf_x L_{\Delta}(x, \lambda)$  subject to  $\lambda \geq 0$  (cf. the general theory in [19] or [20]). Because of the positive homogeneity in Proposition 4(b), however, we have  $g_{\mathcal{D}}(\lambda) = \lambda\Delta$  when  $f_{\mathcal{D}}(x) - \lambda x^{\top}[\bar{r} - r_0e] \geq 0$  for all  $x \in \mathbb{R}^n$ , but  $g_{\mathcal{D}}(\lambda) = -\infty$  otherwise. The problem dual to  $\mathcal{P}_0(\Delta)$  therefore takes the form:

$$\text{maximize } \lambda\Delta \text{ with respect to } \lambda \text{ satisfying } f_{\mathcal{D}}(x) \geq \lambda x^{\top}[\bar{r} - r_0e] \text{ for all } x \in \mathbb{R}^n. \quad (14)$$

Of course, we really only need to understand the case of  $\Delta = 1$ , since everything else can be obtained from that through rescaling. By applying known duality results about the relationship between a convex programming problem and its dual, we get the following conclusion about that case.

**Proposition 6** (duality for basic deviation value). *The basic deviation value  $\bar{\delta}$  has the dual characterization of being the highest  $\lambda$  such that*

$$f_{\mathcal{D}}(x) \geq \lambda x^{\top}[\bar{r} - r_0e] \text{ for all } x \in \mathbb{R}^n. \quad (15)$$

**Proof.** We are dealing with a convex programming problem in which the objective function  $f_{\mathcal{D}}$  has bounded level sets, the property in Proposition 4(d), and on the other hand the Slater constraint qualification holds (it is possible to satisfy inequality constraints, here just one, with strict inequality). In that case the dual problem has an optimal solution, as does the primal problem, and the optimal values in the two problems (the min value in the primal problem and the max value in the dual problem) coincide; cf. [20].  $\square$

In Proposition 6, by the way, we have another fact that would fall by the wayside if we tried to work with  $\mathcal{D}(X) = \text{VaR}(X - EX)$ , which does not conform to our axioms. Results analogous to Theorems 2 and 3 could still be obtained, but there could be no serious follow-up of the kind we are about to get into, because of the absence of convexity.<sup>13</sup>

## 5 Efficient Sets and Frontiers

In our endeavor to understand how the classical picture in Figure 2 might have to be modified and expanded, we cannot limit our attention to  $x$ -portfolios with price  $x^{\top}e > 0$ , for reasons already explained. It is essential to look at price  $x^{\top}e \leq 0$  as well. Moreover, we have to adjust to the fact that the deviation measure  $\mathcal{D}$  under scrutiny might not be symmetric. If we have an  $x$ -portfolio representing a “net long position,” in the sense that  $x^{\top}e > 0$ , and we wish to pass to the associated  $\tilde{x}$ -portfolio with  $\tilde{x} = -x$ , representing a “net short position” because  $\tilde{x}^{\top}e = -x^{\top}e < 0$ , we cannot count on having  $\mathcal{D}(\tilde{x}^{\top}r) = \mathcal{D}(x^{\top}r)$ . Switches between “long” and “short” could have significant effects on risk perception.

---

<sup>13</sup>In order to proceed further with VaR, Alexander and Baptista in [4] limited themselves to the case where the returns of the risky assets are jointly normally distributed. Under this special assumption, however, the VaR setting actually reduces to the CVaR setting, which does fit our framework.



Out of these considerations, we are obliged to investigate an auxiliary optimization problem with respect to the instruments  $i = 1, \dots, n$ . In this problem,  $\pi$  and  $\zeta$  are parameters denoting targeted price and expected gain, and we seek to solve:

$$\mathcal{P}(\pi, \zeta) \quad \text{minimize } f_{\mathcal{D}}(x) = \mathcal{D}(x^{\top} r) \quad \text{subject to } x^{\top} e = \pi \text{ and } x^{\top} \bar{r} = \zeta.$$

We wish to investigate it without any preconditions on the signs of  $\pi$  or  $\zeta$ .

The gain constraint in  $\mathcal{P}(\pi, \zeta)$  has been written as an equation this time because of the chiefly technical role that the analysis will play and the simpler geometry afforded by having an equation instead of an inequality. We let

$$\begin{cases} d(\pi, \zeta) = \text{optimal value (the infimum of the deviation) in } \mathcal{P}(\pi, \zeta), \\ S(\pi, \zeta) = \text{optimal solution set (the minimizing vectors } x \text{) in } \mathcal{P}(\pi, \zeta). \end{cases} \quad (16)$$

**Proposition 7** (parametric framework for price and expected gain). *An optimal solution to problem  $\mathcal{P}(\pi, \zeta)$  is sure to exist (not necessarily uniquely), no matter what the choice of  $\pi$  and  $\zeta$ . Indeed, the solution set  $S(\pi, \zeta)$  in  $\mathbb{R}^n$  is always convex, closed and bounded, with*

$$S(0, 0) = \{0\} \text{ and } S(\lambda\pi, \lambda\zeta) = \{\lambda x \mid x \in S(\pi, \zeta)\} \text{ when } \lambda > 0, \quad (17)$$

while the function  $d$  on  $\mathbb{R}^2$  giving the minimum deviation is finite everywhere and convex (hence also continuous), moreover with the properties that

- (a)  $d(0, 0) = 0$ , but  $d(\pi, \zeta) > 0$  when  $(\pi, \zeta) \neq (0, 0)$ ,
- (b)  $d(\lambda\pi, \lambda\zeta) = \lambda d(\pi, \zeta)$  when  $\lambda > 0$ ,
- (c)  $d(\pi_1 + \pi_2, \zeta_1 + \zeta_2) \leq d(\pi_1, \zeta_1) + d(\pi_2, \zeta_2)$ ,
- (d)  $\{(\pi, \zeta) \mid d(\pi, \zeta) \leq \delta\}$  is a bounded set for every  $\delta > 0$ .

**Proof.** Our assumption A2 guarantees through Proposition 3 that the constraints in  $\mathcal{P}(\pi, \zeta)$  can be satisfied, regardless of how  $\pi$  and  $\zeta$  are chosen. The finiteness of  $d(\pi, \zeta)$  and nonemptiness of  $S(\pi, \zeta)$  follow then from the properties of  $f_{\mathcal{D}}$  in Proposition 4 (much as in the proof of Proposition 5). As the set of solutions to a convex programming problem,  $S(\pi, \zeta)$  is convex and closed. By virtue of the boundedness of the level sets of  $f_{\mathcal{D}}$  in Proposition 4(d),  $S(\pi, \zeta)$  is bounded. Properties (a), (b), (c), and (d) of the function  $d$  follow from the corresponding properties of  $f_{\mathcal{D}}$  in Proposition 4. In particular, the set in (d) is the image of the compact set in Proposition 4(d) under the (continuous) linear transformation  $x \mapsto (x^{\top} e, x^{\top} \bar{r})$ , and that guarantees it is compact as well.  $\square$

The relevance of problem  $\mathcal{P}(\pi, \zeta)$  for our goal of analyzing the price of a solution to problem  $\mathcal{P}_0(\Delta)$  comes from the following observation.

**Proposition 8** (reduced optimization perspective). *When  $\Delta > 0$ , problem  $\mathcal{P}_0(\Delta)$  is equivalent to the problem*

$$\mathcal{P}'_0(\Delta) \quad \text{minimize } d(\pi, \zeta) \text{ subject to } \pi \text{ and } \zeta \text{ satisfying } \zeta - r_0\pi = \Delta,$$

in the sense that optimal values in both problems are the same, and the solutions to  $\mathcal{P}_0(\Delta)$  are the vectors  $x$  in  $\mathbb{R}^n$  such that the pair  $(\pi, \zeta) = (x^{\top} e, x^{\top} \bar{r})$  solves  $\mathcal{P}'_0(\Delta)$ .

In particular,  $\bar{x}$  furnishes a basic portfolio if and only if the pair  $(\bar{x}^{\top} e, \bar{x}^{\top} \bar{r})$  minimizes  $d(\pi, \zeta)$  along the line  $\zeta = r_0\pi + 1$ .

**Proof.** This is elementary in view of the nonemptiness of the solution sets  $S(\pi, \zeta)$  established in Proposition 7, but notice that the single linear constraint, which was an inequality in  $\mathcal{P}_0(\Delta)$ , has been

written now as an equation. Proposition 5 has made this possible by establishing that, when  $\Delta > 0$ , the inequality must be tight at optimality.  $\square$

According to Proposition 8, the pairs  $(x^\top e, x^\top \bar{r})$  giving the price and expected gain associated with the solutions  $x$  (or solution, if unique) to problem  $\mathcal{P}_0(\Delta)$  are the pairs  $(\pi, \zeta)$  that furnish the minimum of the function  $d$  along the line in  $\mathbb{R}^2$  described by the equation  $\zeta = r_0\pi + \Delta$ . Due to positive homogeneity in (11) of Proposition 5, of course, we can concentrate on the case where  $\Delta = 1$ , which corresponds to looking for a basic portfolio. In that case, depicted in Figure 3, the line in question is the one with slope  $r_0$  that passes through the point  $(0, 1)$ .

The “curves” shown in Figure 3 are given by the equations  $d(\pi, \zeta) = \delta$  for various  $\delta > 0$  and reflect the properties in Proposition 7. They are the boundaries of certain compact, convex sets which are merely rescaled versions of each other, generated by expanding or contracting the one for  $\delta = d(0, 1)$ . In the classical case of standard deviation, the curves would be ellipses, but in general they might have corners and straight segments.

The issue of whether the  $x$ -portfolios for  $x \in S_0(1)$  have price  $x^\top e > 0$ ,  $x^\top e = 0$  or  $x^\top e < 0$ , comes down to whether, in minimizing  $d$  along the line in Figure 3, the points  $(\pi, \zeta)$  that are obtained at the minimum have  $\pi > 0$ ,  $\pi = 0$  or  $\pi < 0$ . (We have to speak in general of “ $x$ -portfolios” and “points  $(\pi, \zeta)$ ” because uniqueness of optimal solutions is not assured here, in general.) It is immediately clear that this must depend largely on the magnitude of the risk-free rate  $r_0$  and cannot be resolved merely on the basis of any of the assumptions that have been made, so far, on the rates of return of the instruments  $i = 1, \dots, n$  in our model.

For the  $r_0$  that is illustrated, the slanted line in Figure 3 cuts into the open set  $\{(\pi, \zeta) \mid d(\pi, \zeta) < d(0, 1)\}$  toward the right, and one therefore has  $\pi > 0$  at optimality. But for higher and higher levels of  $r_0$ , a stage will eventually be reached where the line henceforth cuts into this set instead toward the left, in which case  $\pi < 0$  at optimality. A formal analysis of these circumstances, aimed at characterizing the threshold value, or values, of  $r_0$  where the line does not cut into the set at all, will have to be undertaken.

Observe that in the case shown in Figure 3 there is not just one  $r_0$  for which line does not cut into the set in question, but indeed a whole interval of such values. Geometrically, this corresponds to the boundary of the set having a “corner point” at  $(0, 1)$ . Although that can be regarded as an exceptional situation, it cannot be ruled out. Our result on threshold behavior (Theorem 5 in the next section) must therefore, in general, allow for threshold gap corresponding to an *interval* of  $r_0$  values for which one has  $\pi = 0$  at optimality.

For now, the essential thing to recognize is the need to study *two* efficient sets, if price behavior in the one-fund theorem is to be understood over the whole range of possible  $r_0$  values. There has to be an efficient set corresponding to  $x$ -portfolios furnishing “unit long positions” (price = 1), but also one for  $x$ -portfolios furnishing “unit short positions” (price =  $-1$ ). Because the deviation measure  $\mathcal{D}$  might not be symmetric, neither of these efficient sets can be expected to be derivable in a simple way from the other.

Dictates of simplicity in dealing with the geometry of efficiency and its relationship to the one-fund theorem and properties of the function  $d$  cause us to adopt a convention different from the one in Figure 2, where deviation is on the horizontal axis and expected gain on the vertical axis. Instead, we will have deviation on the vertical axis and expected gain on the horizontal axis. Of course, a flip across the  $45^\circ$  line between the two axes can be used to convert our convention to the classical one, when desired.

**Definition 2** (efficient sets and frontiers, positive and negative). *By the positive efficient set and the*

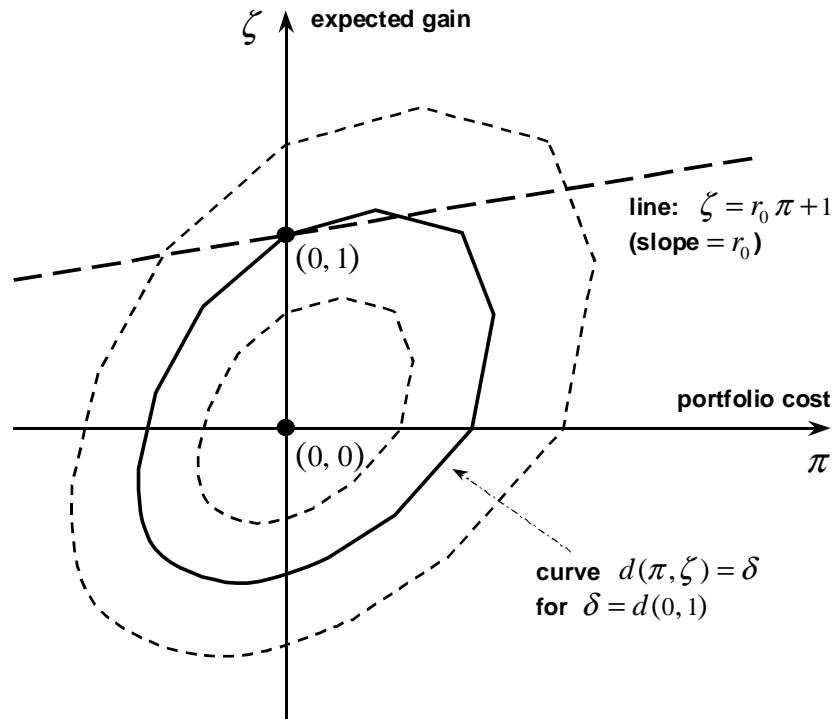


Figure 3: Reduced optimization perspective.

negative efficient set will be meant the boundaries  $G^+$  and  $G^-$ , respectively, of the feasibility sets

$$\begin{aligned} F^+ &= \{(\zeta, \delta) \mid \exists x \text{ with } x^\top e = 1, x^\top \bar{r} = \zeta, f_{\mathcal{D}}(x) \leq \delta\}, \\ F^- &= \{(-\zeta, -\delta) \mid \exists x \text{ with } x^\top e = -1, x^\top \bar{r} = \zeta, f_{\mathcal{D}}(x) \leq \delta\}. \end{aligned} \quad (18)$$

By the positive efficient frontier will be meant the part of  $G^+$  consisting of all  $(\zeta, \delta) \in F^+$  for which there is no  $(\zeta', \delta) \in F^+$  with  $\zeta' > \zeta$ . Likewise, by the negative efficient frontier will be meant the part of  $G^-$  consisting of all  $(-\zeta, -\delta) \in F^-$  for which there is no  $(\zeta', \delta) \in F^-$  with  $\zeta' > \zeta$ .

The virtue of passing to  $(-\zeta, -\delta)$  in the definition of  $F^-$  will emerge in the results below as a way of getting the most out of a single geometric picture in which both of the efficient sets, or frontiers, can be seen, namely the picture in Figure 4, which will be explained in due course, after Theorem 4. Note that, through a notational switch between  $x$  and  $-x$ , one can express  $F^-$  equivalently by

$$F^- = \{(\zeta, -\delta) \mid \exists x \text{ with } x^\top e = 1, x^\top \bar{r} = \zeta, f_{\mathcal{D}}(-x) \leq \delta\}. \quad (19)$$

Thus, in cases where the deviation measure  $\mathcal{D}$  is symmetric, so that  $f_{\mathcal{D}}(-x) = f_{\mathcal{D}}(x)$ , the set  $F^-$  would merely be the reflection of the set  $F^+$  across the  $\zeta$ -axis, and there would be less of an imperative for considering it separately.<sup>14</sup>

As depicted in Figure 4, the positive efficient frontier is the “right” boundary of  $F^+$ , in contrast to  $G^+$  being the whole boundary. In the same way, the negative efficient frontier is the “left” boundary of  $F^-$ , in contrast to  $G^-$  being the whole boundary. Only these partial boundaries will really have a role in what follows, but it is convenient mathematically to work with  $G^+$  and  $G^-$  themselves. For convenience of comparisons, Figure 5 poses all these sets in the reversed coordinate system that is customary in finance. There the positive efficient frontier becomes an “upper” boundary and the negative efficient frontier a “lower” boundary.

**Proposition 9** (efficient sets as function graphs). *The positive efficient set  $G^+$  is the graph of the convex function*

$$d^+(\zeta) = d(1, \zeta) = \text{min deviation for price 1 and expected gain } \zeta, \quad (20)$$

whereas the negative efficient set  $G^-$  is the graph of the concave function

$$d^-(\zeta) = -d(-1, -\zeta) = -\text{min deviation for price } -1 \text{ and expected gain } -\zeta. \quad (21)$$

Indeed,  $F^+$  is the closed, convex set consisting of the pairs  $(\zeta, \delta)$  for which  $\delta \geq d^+(\zeta)$ , whereas  $F^-$  is the closed, convex set consisting of the pairs  $(\zeta, \delta)$  for which  $\delta \leq d^-(\zeta)$ . Furthermore, the asymptotic slope of  $G^+$  on the right is the same as the asymptotic slope of  $G^-$  on the left,

$$\lim_{\zeta \rightarrow \infty} \frac{d^+(\zeta)}{\zeta} = \lim_{\zeta \rightarrow -\infty} \frac{d^-(\zeta)}{\zeta} = d(0, 1) > 0. \quad (22)$$

**Proof.** The convexity of  $d^+$  and concavity of  $d^-$  are evident from the convexity of  $d$  in Proposition 7. Like  $d$ , these functions are finite and continuous, in particular. Those properties, along with the fact in Proposition 7 that the minimum deviations in (20) and (21) are sure to be attained, immediately yield the descriptions claimed for  $F^+$  and  $F^-$ .

<sup>14</sup>It would be possible then, in the manner of [9] for the classical case of standard deviation, to identify master funds of negative type with the reverses of master funds with positive type. But the main trend now in this subject is instead to get away from symmetry and treat the downside differently from the upside.

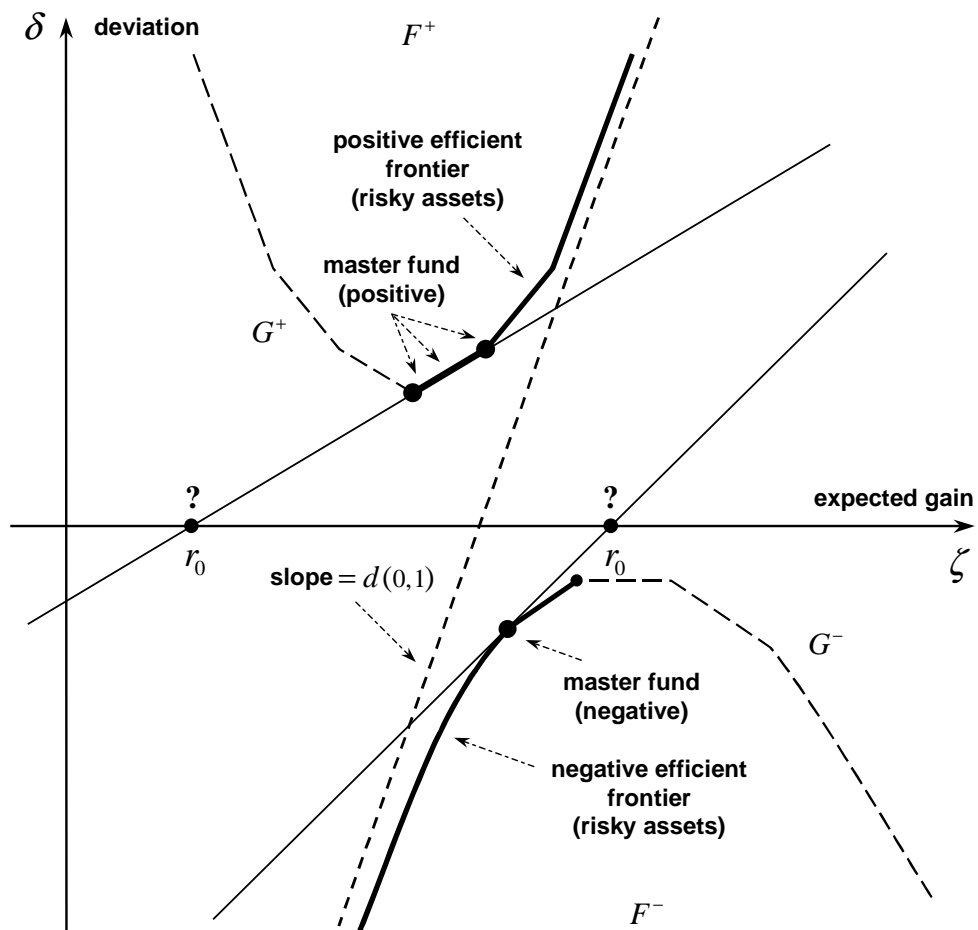


Figure 4: Efficient sets and frontiers for portfolios of risky instruments in coordinates standard for graphs of functions.

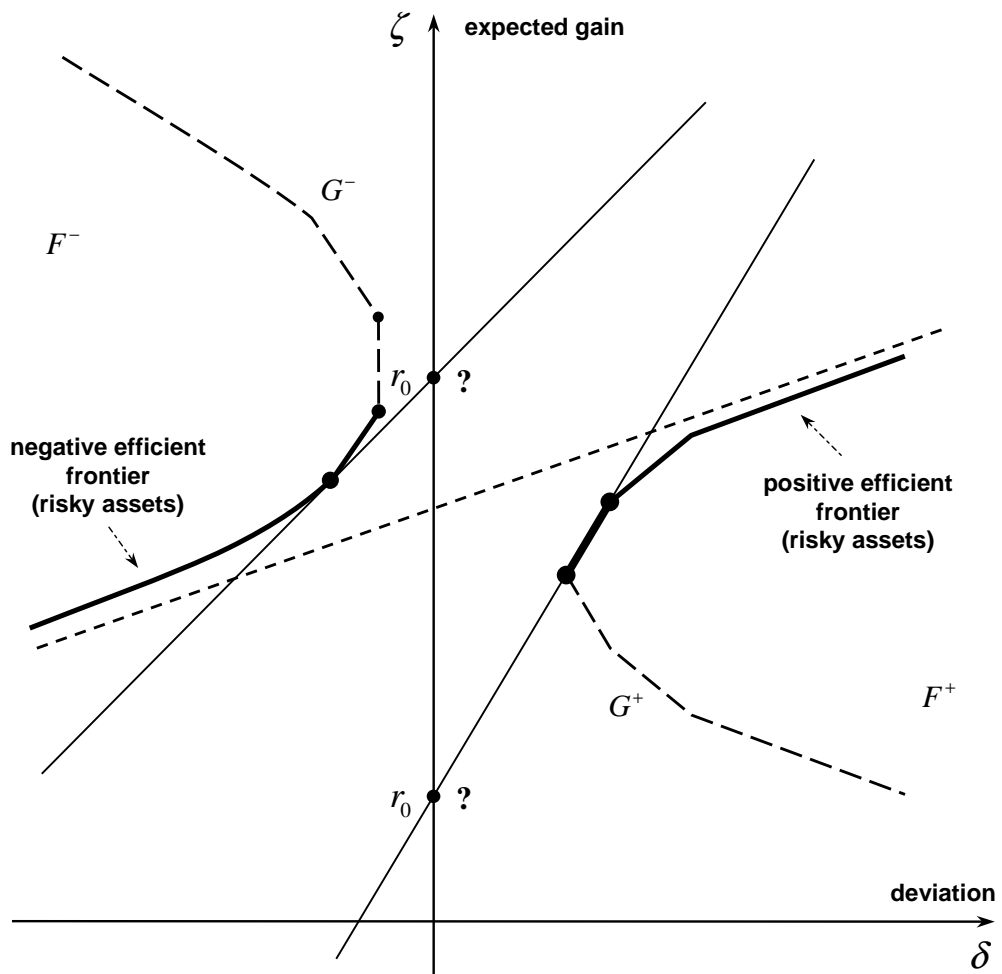


Figure 5: Efficient sets and frontiers for portfolios of risky instruments in coordinates traditional for finance.

The relations in (22) specialize to  $d$  a well known property in convex analysis: the asymptotic slope of a finite convex function along any half-line depends only on the direction of the half-line, not on its starting point (see [19, Theorem 8.5]). Here, we are looking at such slopes along half-lines in the  $\pi$ - $\zeta$  space of Figure 3 that are parallel to the  $\zeta$ -axis. On the positive  $\zeta$ -axis itself,  $d$  is linear because of the positive homogeneity in Proposition 7(b): we have

$$d(0, \zeta) = \zeta d(0, 1) \text{ for all } \zeta > 0,$$

so  $d(0, 1)$  is the limit of  $d(0, \zeta)/\zeta$  as  $\zeta \rightarrow \infty$ . Then  $d(0, 1)$  is also the limit of  $d(1, \zeta)/\zeta$  as  $\zeta \rightarrow \infty$ , as well as the limit of  $d(-1, \zeta)/\zeta$  as  $\zeta \rightarrow \infty$ . By the definitions of  $d^+$  and  $d^-$ , these two limits are, respectively, the limit of  $d^+(\zeta)/\zeta$  as  $\zeta \rightarrow \infty$  and the limit of  $d^-(-\zeta)/(-\zeta)$  as  $-\zeta \rightarrow -\infty$ .  $\square$

Incidentally, the asymptotic slope of  $G^+$  on the right agrees likewise with the asymptotic slope of  $G^-$  on the left, but this fact will not play any role here. (The portions of  $G^+$  on the left and  $G^-$  on the right would drop out if we wrote the  $\zeta$  constraint in  $\mathcal{P}(\pi, \zeta)$  as an inequality instead of an equation.)

We are in position now to answer the question of how the tangency relationship associated with the classical one-fund theorem, as in Figure 2, can be extended to our framework of general deviation measures, as a complement to the one-fund results in Theorem 2.

**Theorem 4** (efficiency characterization of master funds). *The basic deviation value  $\bar{\delta} = d_0(1)$  is the highest of the slopes of all the lines in  $\mathbb{R}^2$  through  $(r_0, 0)$  that lie between the curves  $G^+$  and  $G^-$  (perhaps touching them, but not crossing them). In referring to the line through  $(r_0, 0)$  with slope  $\bar{\delta}$  as the “ $r_0$ -line”, the following conclusions can be drawn.*

- (a) *If the  $r_0$ -line touches  $G^+$  at a point  $(\zeta^*, \delta^*)$ , then any  $x^* \in S(1, \zeta^*)$  furnishes a master fund of positive type: it has price  $x^{*\top} e = 1$  and belongs to the optimal solution set  $S_0(\Delta^*)$  for  $\Delta^* = \zeta^* - r_0 > 0$ .*
- (b) *If the  $r_0$ -line touches  $G^-$  at a point  $(-\zeta^*, -\delta^*)$ , then any  $x^* \in S(1, \zeta^*)$  furnishes a master fund of negative type: it has price  $x^{*\top} e = -1$  and belongs to the optimal solution set  $S_0(\Delta^*)$  for  $\Delta^* = \zeta^* + r_0 > 0$ .*
- (c) *The maximum value that  $\bar{\delta}$  can have with respect to different values of  $r_0$  is the common asymptotic slope value  $d(0, 1)$  for  $G^+$  on the right and  $G^-$  on the left.*

**Proof.** Our strategy is to derive this from the dual characterization of  $\bar{\delta}$  in Proposition 6. That characterization translates in terms of the definition of  $d$  into having

$$\bar{\delta} = \max\{\lambda \mid \lambda[\zeta - r_0\pi] \leq d(\pi, \zeta) \text{ for all } (\pi, \zeta) \in \mathbb{R}^n\}. \quad (23)$$

For the inequality condition inside this description to hold, it only has to hold when  $\pi > 0$  or  $\pi < 0$ , since it must then hold automatically for  $\pi = 0$  by the continuity of  $d$ . Indeed, because of the positive homogeneity of  $d$  in Proposition 7(b), it merely has to hold for  $\pi = 1$  and for  $\pi = -1$ , in order for this conclusion to be reached.

In the case of  $\pi = 1$ , the inequality  $\lambda[\zeta - r_0\pi] \leq d(\pi, \zeta)$  comes down to  $\lambda(\zeta - r_0) \leq d^+(\zeta)$ . Having this hold for all  $\zeta \in \mathbb{R}$  means that the line in  $(\zeta, \delta)$ -space through  $(r_0, 0)$  with slope  $\lambda$  does not cross above the graph of  $d^+$ .

In the case of  $\pi = -1$ , the inequality  $\lambda[\zeta - r_0\pi] \leq d(\pi, \zeta)$  comes down to  $\lambda[\zeta + r_0] \leq -d^-(-\zeta)$ . With a switch of notation between  $\zeta$  and  $-\zeta$ , this becomes  $\lambda[\zeta - r_0] \geq d^-(\zeta)$ . Having that hold for all  $\zeta \in \mathbb{R}$  means that the line in question does not cross below the graph of  $d^-$ . Thus, the characterization of  $d_0$  in (23) reduces to the graphical characterization claimed in the theorem.

Now (a) and (b) are obvious from the representations of  $G^+$  and  $G^-$  in Proposition 9. On the other hand, (c) follows from the monotonicity in the curvatures of  $G^+$  and  $G^-$  (i.e., of the left and right

derivatives of the functions  $d^+$  and  $d^-$ , due to their convexity and concavity), as expressed through Proposition 8 and the claim there about asymptotic slopes.  $\square$

Figure 4 illustrates the general situation described in Theorem 4. Two possible values of the risk-free rate  $r_0$  are indicated, corresponding to alternatives (a) and (b) of the theorem. Note the possibility of more than one point of “tangency,” and on the other hand, the possibility of a range of  $r_0$  values all yielding the same point of tangency. Alternative (c) of Theorem 4 corresponds to  $r_0$ -line having the same slope as the diagonal line shown in dashes. Typically there might be only one such line, associated with a unique threshold rate, but sometimes there could be a family of parallel lines corresponding to an interval of  $r_0$  rates. This will be the subject of the next section, and eventually, Figure 6.

The same basic relationships underlie the reversed-coordinate picture in Figure 5, of course, but there they cannot be described so simply in terms of slopes. A line having slope  $\bar{d}$  in Figure 4 turns into a line having slope  $1/\bar{d}$  in Figure 5, so for instance, the optimal value in problem  $\mathcal{P}_0(1)$  emerges instead as the *reciprocal* of the *lowest* of the slopes of all the lines through  $(0, r_0)$  that lie between the two efficient sets. The awkwardness of this kind of statement, insisting on reciprocals, is another of the reasons why we have chosen to give priority to the presentation in Figure 4.

## 6 Threshold Determination for the Risk-Free Rate

The task immediately ahead of us is the analysis of the transitional behavior between the cases in Theorem 4. For that, we will make use of the Lagrange multipliers associated with the problem  $\mathcal{P}(\pi, \zeta)$ , specifically in the case of  $(\pi, \zeta) = (0, 1)$ . The Lagrangian for  $\mathcal{P}(\pi, \zeta)$  is the function

$$L_{(\pi, \zeta)}(x, \rho, \eta) = f_{\mathcal{D}}(x) + \rho[\pi - x^{\top} e] + \eta[\zeta - x^{\top} \bar{r}]. \quad (24)$$

We say that  $(\rho, \eta)$  is a *Lagrange multiplier vector* for  $\mathcal{P}(\pi, \zeta)$  when

$$\inf_x L_{(\pi, \zeta)}(x, \eta, \rho) = \text{optimal value } d(\pi, \zeta) \text{ in } \mathcal{P}(\pi, \zeta). \quad (25)$$

This definition, the standard one for a convex programming problem like  $\mathcal{P}(\pi, \zeta)$  (cf. [19]), gets around the fact that the objective function  $f_{\mathcal{D}}(x)$  might not be differentiable everywhere with respect to  $x$ . We let

$$M(\pi, \zeta) = \text{set of Lagrange multiplier vectors } (\rho, \eta) \text{ in } \mathcal{P}(\pi, \zeta). \quad (26)$$

Under our assumptions, the Lagrange multiplier set  $M(\pi, \zeta)$  is always nonempty, convex and bounded. This is true from the general theory of convex programming problems because the optimal solution set to  $\mathcal{P}(\pi, \zeta)$  is always nonempty and bounded, and the optimal value  $d(\pi, \zeta)$  is always finite; cf. [20], [19]. Moreover, the multiplier vectors for  $\mathcal{P}(\pi, \zeta)$  are known from that theory to be the “subgradients” of the (optimal-value) function  $d$  at the point  $(\pi, \zeta)$ . Accordingly, they furnish the formula

$$d'(\pi, \zeta; \pi', \zeta') = \max_{\rho, \eta} \{ \pi' \rho + \zeta' \eta \mid (\rho, \eta) \in M(\pi, \zeta) \}, \quad (27)$$

where the left side denotes the one-sided directional derivative of  $d$  at  $(\pi, \zeta)$  with respect to a vector  $(\pi', \zeta')$  and is defined by

$$d'(\pi, \zeta; \pi', \zeta') = \lim_{\varepsilon \rightarrow 0^+} \frac{d(\pi + \varepsilon \pi', \zeta + \varepsilon \zeta') - d(\pi, \zeta)}{\varepsilon} \quad (28)$$

Such derivatives exist because  $d$  is convex.



**Theorem 5** (rate thresholds and price behavior).

(a) Threshold values  $\hat{r}_0^+$  and  $\hat{r}_0^-$  exist which satisfy  $\hat{r}_0^- \leq \hat{r}_0^+$  and have the following effect on basic funds, and therefore on the existence of master funds:

$$\left\{ \begin{array}{l} \text{every } \bar{x} \text{ furnishing a basic fund has price } \bar{x}^\top e > 0 \text{ when } r_0 < \hat{r}_0^-, \\ \text{every } \bar{x} \text{ furnishing a basic fund has price } \bar{x}^\top e < 0 \text{ when } r_0 > \hat{r}_0^+, \\ \text{every } \bar{x} \text{ furnishing a basic fund has price } \bar{x}^\top e = 0 \text{ when } \hat{r}_0^- < r_0 < \hat{r}_0^+, \end{array} \right.$$

where the third case falls away if actually  $\hat{r}_0^- = \hat{r}_0^+$  so as to yield a single threshold rate  $\hat{r}_0$ .

(b) In the borderline case of  $r_0 = \hat{r}_0^-$ , at least one  $\bar{x}$  furnishing a basic fund has price  $\bar{x}^\top e = 0$ , but there could be other basic fund vectors  $\bar{x}$  having price  $\bar{x}^\top e > 0$ . Similarly, in the case of  $r_0 = \hat{r}_0^+$ , at least one  $\bar{x}$  furnishing a basic fund has price  $\bar{x}^\top e = 0$ , but there could be other basic fund vectors  $\bar{x}$  having price  $\bar{x}^\top e < 0$ .

(c) The threshold rates  $\hat{r}_0^+$  and  $\hat{r}_0^-$  can be determined from the Lagrange multiplier set  $M(0, 1)$  for problem  $\mathcal{P}(0, 1)$ . That set consists of the pairs  $(\rho, \eta)$  satisfying  $\eta = d(0, 1)$  and  $\rho^- \leq \rho \leq \rho^+$  for a certain interval  $[\rho^-, \rho^+]$ , and in those terms one has

$$\hat{r}_0^- = \frac{-\rho^+}{d(0, 1)}, \quad \hat{r}_0^+ = \frac{-\rho^-}{d(0, 1)}.$$

**Proof.** Let  $\varphi(\pi) = d(\pi, r_0\pi + 1)$ , this being a finite, convex function on  $\mathbb{R}$  (by Proposition 7). Because of the scaling relation in (11) of Proposition 5, we need only look at the case where  $\Delta = 1$ . As seen in Proposition 9, the prices  $x^\top e$  of the vectors  $x \in S_0(1)$  are the values of  $\pi$  that minimize  $\varphi$  over  $\mathbb{R}$ . Such values form a nonempty, closed, bounded interval in  $\mathbb{R}$ , inasmuch as  $S_0(1)$  is a nonempty, closed, bounded, convex subset of  $\mathbb{R}^n$  (cf. Proposition 7); this interval may well collapse to just one  $\pi$  value, of course. The issue is the extent to which the values of  $\pi$  that minimize  $\varphi$  may be positive, negative or zero.

As a finite, convex function on  $\mathbb{R}$ ,  $\varphi$  has right and left derivatives  $\varphi'_+(\pi)$  and  $\varphi'_-(\pi)$  which are nondecreasing as functions of  $\pi$ , with  $\varphi'_-(\pi) \leq \varphi'_+(\pi)$ . The minimum of  $\varphi$  is attained at  $\pi$  if and only if  $\varphi'_-(\pi) \leq 0 \leq \varphi'_+(\pi)$ .

We can test this condition at  $\pi = 0$ . If  $\varphi'_+(0) < 0$ , the minimum of  $\varphi$  can only be attained at some  $\pi > 0$ , whereas if  $\varphi'_-(0) > 0$ , it can only be attained at some  $\pi < 0$ . If  $\varphi'_-(0) < 0 < \varphi'_+(0)$ , it can only be attained at  $\pi = 0$ . When  $\varphi'_+(0) = 0$ , the minimum is attained at  $\pi = 0$ , but it is conceivable that  $\varphi$  might be constant over some interval  $[0, \varepsilon]$  (with  $\varepsilon > 0$ ), and the minimum would also be attained then by the positive values of  $\pi$  in that interval. Likewise, when  $\varphi'_-(0) = 0$ , the minimum is attained at  $\pi = 0$ , but it is conceivable that  $\varphi$  might be constant over some interval  $[-\varepsilon, 0]$ , and the minimum would also be attained then by the negative values of  $\pi$  in that interval.

The crucial left and right derivatives  $\varphi'_+(0)$  and  $\varphi'_-(0)$  are obtainable from the one-sided directional derivatives of  $d$ :

$$\varphi'_+(0) = d'(0, 1; 1, r_0), \quad \varphi'_-(0) = -d'(0, 1; -1, -r_0).$$

The Lagrange multiplier characterization of the directional derivatives of  $d$  in (27) tells us then that

$$\varphi'_+(0) = \max_{\rho, \eta} \{ \rho + r_0\eta \mid (\rho, \eta) \in M(0, 1) \}, \quad \varphi'_-(0) = \min_{\rho, \eta} \{ \rho + r_0\eta \mid (\rho, \eta) \in M(0, 1) \}.$$

We note next that, because  $d(0, \lambda) = \lambda d(0, 1)$  when  $\lambda > 0$  by the positive homogeneity in Proposition 7(b), we have  $d(0, 1) = d'(0, 1; 0, 1)$  and  $-d(0, 1) = d'(0, 1; 0, -1)$ , and consequently by (27) that

$$d(0, 1) = \max_{\rho, \eta} \{ 0 \cdot \rho + 1 \cdot \eta \mid (\rho, \eta) \in M(0, 1) \} = \min_{\rho, \eta} \{ 0 \cdot \rho + 1 \cdot \eta \mid (\rho, \eta) \in M(0, 1) \},$$

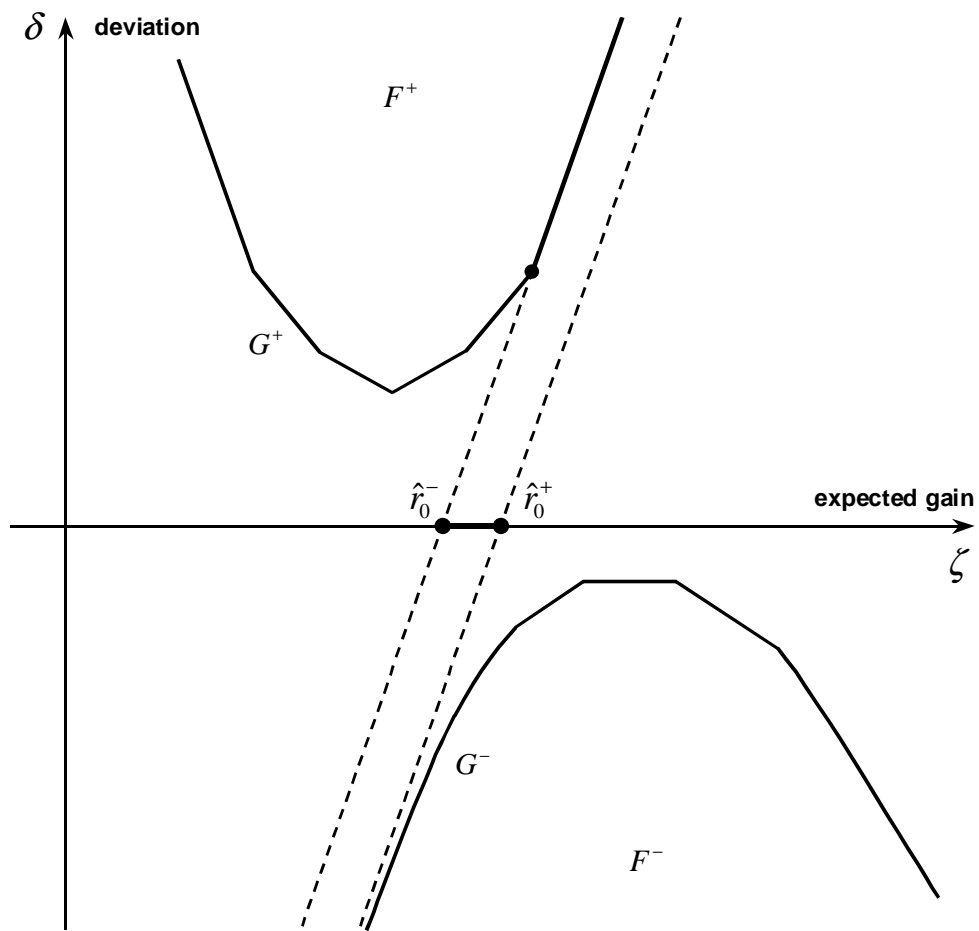


Figure 6: Threshold values for the risk-free rate.

so  $d(0, 1)$  coincides with both the maximum and minimum values of the  $\eta$ -coordinate for points in  $M(0, 1)$ . This implies that  $M(0, 1)$  lies within the horizontal line consisting of the pairs  $(\rho, \eta)$  such that  $\eta = d(0, 1)$ . Because  $M(0, 1)$  is a compact, convex set, it must actually be a closed segment of that line (possibly reduced to a single point): the corresponding  $\rho$  values must comprise an interval  $[\rho^-, \rho^+]$ . Thus,  $M(0, 1)$  has the special form claimed in the theorem. The formulas we already have for  $\varphi'_+(0)$  and  $\varphi'_-(0)$  tell us then that

$$\varphi'_+(0) = \rho^+ + r_0 d(0, 1), \quad \varphi'_-(0) = \rho^- + r_0 d(0, 1).$$

Finally, we can put this together with the criterion already developed for the location of the values of  $\pi$  that minimize  $\varphi$ . Clearly  $\varphi'_+(0) < 0$  if and only if  $r_0 < -\rho^+/d(0, 1)$ , whereas  $\varphi'_-(0) > 0$  if and only if  $r_0 > -\rho^-/d(0, 1)$ , and the proof is thus finished.  $\square$

**Corollary** (optimal portfolios at transition). *The values of  $r_0$  for which the maximum in part (c) of Theorem 4 is attained are those in the interval  $[\hat{r}_0^-, \hat{r}_0^+]$ . For such  $r_0$ , the optimal solution set  $S_0(\Delta)$  to  $\mathcal{P}_0(\Delta)$ , for any  $\Delta > 0$ , contains an  $x$  having price  $x^\top e = 0$ .*

Theorem 5 provides important information about master funds, in particular.

**Theorem 6** (existence of master funds). *The threshold values  $\hat{r}_0^+$  and  $\hat{r}_0^-$  in Theorem 5 have the property that*

$$\left\{ \begin{array}{l} \text{when } r_0 < \hat{r}_0^-, \text{ there is a master fund of positive type but none of negative type,} \\ \text{when } r_0 > \hat{r}_0^+, \text{ there is a master fund of negative type but none of positive type,} \\ \text{when } \hat{r}_0^- < r_0 < \hat{r}_0^+, \text{ there is neither a master fund of positive type nor one of negative type.} \end{array} \right.$$

*In the borderline case of  $r_0 = \hat{r}_0^-$ , there might be a master fund of positive type, whereas in the borderline case of  $r_0 = \hat{r}_0^+$ , there might be a master fund of negative type. (When  $\hat{r}_0^- = \hat{r}_0^+$ , it is not excluded that master funds of both types exist simultaneously.)*

**Proof.** When  $r_0 < \hat{r}_0^-$ , there exists by Proposition 5 and Theorem 5(a) an  $x \in S_0(\Delta)$  having  $x^\top e > 0$ . By setting  $x^* = x/x^\top e$  and  $\Delta^* = \Delta/x^\top e$ , we get  $x^{*\top} e = 1$  and have  $x^* \in S_0(\Delta^*)$  (again by Proposition 5). This  $x^*$  meets the prescription in Definition 1 for furnishing a master fund of positive type.

Similarly, when  $r_0 > \hat{r}_0^+$ , there exists by Proposition 5 and Theorem 5(a) an  $x \in S_0(\Delta)$  having  $x^\top e < 0$ . Then, by setting  $x^* = x/|x^\top e|$  and  $\Delta^* = \Delta/x^\top e$ , we get  $x^{*\top} e = -1$  and have  $x^* \in S_0(\Delta^*)$ . This  $x^*$  furnishes a master fund of negative type.

On the other hand, Theorem 5(a) makes clear that a master fund of positive type cannot exist when  $r_0 > \hat{r}_0^-$ , but might exist (by the argument just given) when  $r_0 = \hat{r}_0^-$ . Likewise, a master fund of negative type cannot exist when  $r_0 < \hat{r}_0^+$ , but might exist when  $r_0 = \hat{r}_0^+$ .  $\square$

The interpretation coming from Theorems 5 and 6 is that when the risk-free rate  $r_0$  is high enough (specifically, above the threshold rate  $\hat{r}_0^+$ ), it is advantageous, for investors whose attitudes toward risk are captured by the particular deviation measure  $\mathcal{D}$  under investigation, to take a net “short position” in the market (an  $x$ -portfolio with negative price) and invest at the risk-free rate all the money that is obtained that way.

The relation between threshold behavior and the efficient set geometry in Figure 4 is indicated in Figure 6. Ordinarily, it can be expected that  $\hat{r}_0^- = \hat{r}_0^+$ , in which case the unified threshold value may be denoted simply by  $\hat{r}_0$ . The explanation is that these values are determined in the proof of Theorem 5 from right and left derivatives of the convex function  $\varphi(\pi) = d(\pi, r_0\pi + 1)$ , and such derivatives have to coincide almost everywhere. When  $\hat{r}_0^- = \hat{r}_0^+$ , there is only one line that fits between the two efficient sets.

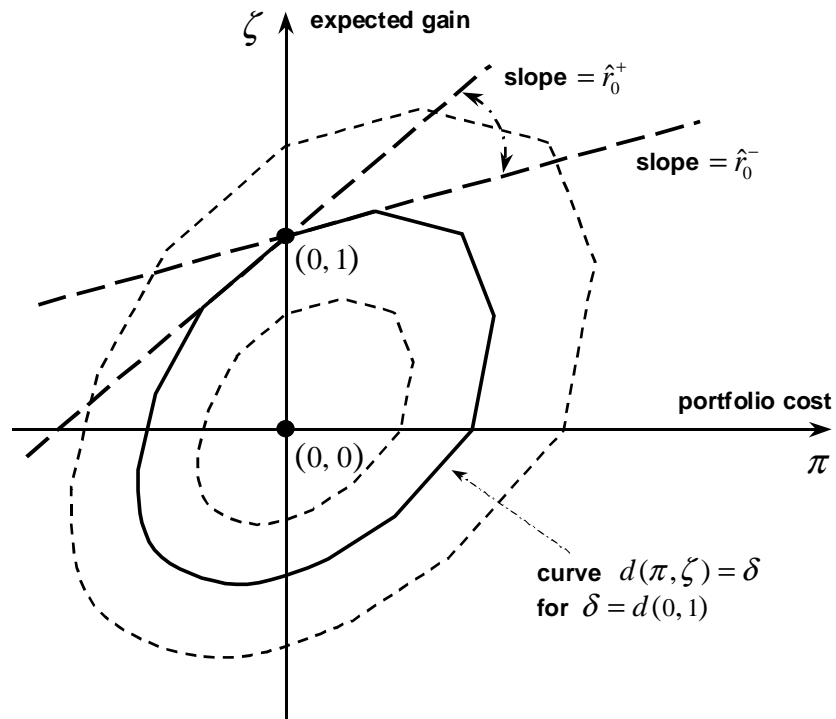


Figure 7: Threshold interval tied to corner behavior.

The reason why cases with  $\hat{r}_0^- < \hat{r}_0^+$  can truly occur is seen, from this vantage point, to be tied to the fact that the right and left derivatives of  $\varphi$  may differ in some places, due to the function  $d$  not being differentiable. But it can also be understood from the remarks made earlier about the geometry in Figure 3. When the curve through  $(0, 1)$  in Figure 3 has a corner there, one has a range of slopes  $r_0$  corresponding to cases for which the minimum in the reduced format of Proposition 7 occurs with  $\pi = 0$ . This range of slopes is marked by the two threshold values  $\hat{r}_0^-$  and  $\hat{r}_0^+$ , as shown in Figure 7.

Allowance for corner points really does have to be made, because of deviation measures such as the one in (8), for instance. There  $f_{\mathcal{D}}(x)$  is piecewise linear and the curves in Figure 7 are polygonal.

It should be noted that, although Theorem 6 conveys the circumstances in which master funds of one type or the other are sure to *exist*, it says nothing about when they might be *unique*. That is an entirely separate issue. Uniqueness could fail on two grounds. The first is the possibility of more than one point of tangency where the  $r_0$ -line meets the frontier, as seen in Figure 4. The second arises when more than one portfolio can yield the same point on the frontier.

It may be anticipated that these phenomena are “rare,” but they cannot readily be eliminated, a priori, in the absence of a suitable strict convexity property of  $\mathcal{D}$ . However, we have demonstrated in [24, Example 4] that the required version of strict convexity is unavailable, in general, for coherent deviation measures such as lower semideviation, lower range, and CVaR. The same can be seen for mixed CVaR and mean absolute deviation.

In contrast to these observations, and the facts in Theorem 6, both the existence and uniqueness of a master portfolio seem to be taken for granted in much of the literature on portfolio optimization. The belief is widespread, moreover, that a master portfolio of positive type always suffices, regardless of the magnitude of the risk-free rate  $r_0$ . Our hope is that the rigorous methodology pursued in this paper will help to dispel such misconceptions.

## 7 Option Example with Nondifferentiability and a Threshold Gap

We present now an example of CVaR type in which the deviation function  $f_{\mathcal{D}}$  fails to be differentiable in some places and this moreover creates a threshold gap where  $\hat{r}_0^- < \hat{r}_0^+$  as in Figures 6 and 7. The trouble will be seen to arise in this case through the utilization of options, although it can also arise in other ways and with non-CVaR deviation measures, especially when the scenario space  $\Omega$  is discrete, e.g. as in (8).

Before getting into the specific details, we explain the idea behind our example and why it works. To keep matters at their simplest, we consider only two risky instruments, having returns  $r_1$  and  $r_2$ , so that portfolios correspond to points  $x = (x_1, x_2)$  in  $\mathbb{R}^2$  and have returns  $x_1 r_1 + x_2 r_2$ . We aim at identifying circumstances in which the convex function

$$f_{\mathcal{D}}(x_1, x_2) = \text{CVaR}_{\alpha}([x_1 r_1 + x_2 r_2] - E[x_1 r_1 + x_2 r_2]) = \text{CVaR}_{\alpha}(x_1 r_1 + x_2 r_2) + x_1 \bar{r}_1 + x_2 \bar{r}_2 \quad (29)$$

for a value  $\alpha \in (0, 1)$  (yet to be specified) is not differentiable along the line  $x_1 + x_2 = 0$  at points where  $x_1 > 0$  and  $x_2 < 0$ . In view of the positive homogeneity property in Proposition 4(b), only one such point needs to be considered. The nondifferentiability can thus be confirmed by showing that the partial derivative of  $f_{\mathcal{D}}(x_1, x_2)$  with respect to  $x_1$  does not exist at  $(1, -1)$ , or in other words, that the convex function of  $s \in \mathbb{R}$  defined by

$$f(s) = \text{CVaR}_{\alpha}(r_{[s]}) \quad \text{for } r_{[s]} = (1 + s)r_1 - r_2 = r_{[0]} + s r_1 \quad (30)$$

is not differentiable at  $s = 0$ . Through its convexity,  $f$  necessarily has right and left derivatives

$$f'_+(0) = \lim_{\varepsilon \rightarrow 0^+} \frac{f(\varepsilon) - f(0)}{\varepsilon}, \quad f'_-(0) = \lim_{\varepsilon \rightarrow 0^+} \frac{f(0) - f(-\varepsilon)}{\varepsilon}, \quad (31)$$

which satisfy  $f'_-(0) \leq f'_+(0)$ , so our task is to set the example up to get  $f'_-(0) < f'_+(0)$ .

Once this nondifferentiability has been confirmed, the existence also of a threshold gap will automatically follow, for reasons explained next. The nondifferentiability of  $f_{\mathcal{D}}$  means geometrically that each of the convex level sets given by an inequality of the form  $f_{\mathcal{D}}(x_1, x_2) \leq c$  (for a constant  $c > 0$ ) has a “corner point” where its boundary meets the line  $x_1 + x_2$  in the fourth quadrant of  $\mathbb{R}^2$ . The key then is the observation that the function  $d$  can be obtained from  $f_{\mathcal{D}}$  merely by a linear change of coordinates in  $\mathbb{R}^2$ , from  $(x_1, x_2)$  coordinates to  $(\pi, \zeta)$  coordinates, in which the line  $x_1 + x_2 = 0$  turns into the line  $\pi = 0$ . Indeed, from (16) we see that  $d(\pi, \zeta)$  is obtained by solving the equations  $x_1 + x_2 = \pi$  and  $x_1\bar{r}_1 + x_2\bar{r}_2 = \zeta$  for  $x_1$  and  $x_2$  in terms of  $\pi$  and  $\zeta$  and substituting those expressions into  $f_{\mathcal{D}}(x_1, x_2)$ . The demonstrated corner points of the convex level sets where  $f_{\mathcal{D}}(x_1, x_2) \leq c$  along  $x_1 + x_2 = 0$  turn into corner points of the convex level sets where  $d(\pi, \zeta) \leq c$  that lie on the vertical axis as in Figure 7 and thereby signal a threshold gap.

Now come the details. We start with a stock whose rate of return is governed by a cumulative distribution function  $F$  arising from a positive density function on  $(-1, \infty)$ ; then  $F$  is differentiable and increasing on  $[-1, \infty)$ , although identically 0 on  $(-\infty, -1]$ . We take  $\Omega$  to be  $[-1, \infty)$  supplied with the probability measure associated with  $F$ . Instrument 1 is the stock in question, so  $r_1(\omega) = \omega$ . Instrument 2 is an “insured” version of this stock which always returns at least  $\omega_1$  but never more than  $\omega_2$ , where  $-1 < \omega_1 < 0 < \omega_2$ ; then

$$r_2(\omega) = \begin{cases} \omega_1 & \text{if } \omega \leq \omega_1, \\ \omega & \text{if } \omega \in [\omega_1, \omega_2], \\ \omega_2 & \text{if } \omega \geq \omega_2, \end{cases} \quad r_{[s]}(\omega) = \begin{cases} (1+s)\omega - \omega_1 & \text{if } \omega \leq \omega_1, \\ s\omega & \text{if } \omega \in [\omega_1, \omega_2], \\ (1+s)\omega - \omega_2 & \text{if } \omega \geq \omega_2. \end{cases}$$

Instrument 2 can be interpreted as a composite formed by combining the underlying stock with a put option at  $\omega_1$  and a shorted call option at  $\omega_2$  (with  $\omega_1$  and  $\omega_2$  being at levels where the cost of the shorted call exactly covers the cost of the put). However, to avoid the complication of introducing the options as separate instruments, which would increase the number of portfolio variables and mess up the two-dimensional geometry, we prefer to think of Instrument 2 as a single entity.

Let  $F_{[s]}$  denote the cumulative distribution function for the random variable  $r_{[s]}$ . Observe that for  $s = 0$  we have  $F_{[0]}(\gamma) = F(\omega_1 + \gamma)$  for  $\gamma < 0$  but  $F_{[0]}(\gamma) = F(\omega_2 + \gamma)$  for  $\gamma \geq 0$ . The jump of  $F_{[0]}$  at  $\gamma = 0$  from  $F(\omega_1)$  to  $F(\omega_2)$  is the feature we will exploit.

To do this without unnecessary trouble in the calculations, we choose  $\alpha = F(0)$ , so that  $\alpha$  is the probability of the interval  $[-1, 0]$  in  $\Omega$  and is intermediate to the jump of  $F_{[0]}$  at  $\gamma = 0$ , inasmuch as  $F(\omega_1) < F(0) < F(\omega_2)$ . In particular, then,  $\text{VaR}_{\alpha}(r_{[0]}) = 0$ . The  $\alpha$ -tail distribution for  $r_{[0]}$ , which by definition corresponds to the truncated distribution function with values  $\alpha^{-1}F_{[0]}(\gamma)$  when  $F_{[0]}(\gamma) < \alpha$  but 1 when  $F_{[0]}(\gamma) \geq \alpha$ , has density  $\alpha^{-1}F'(\omega_1 + \gamma) > 0$  for  $\gamma$  in  $[-1 - \omega_1, 0]$  and an atom of size  $1 - \alpha^{-1}F(\omega_1)$  at  $\gamma = 0$ , but vanishes elsewhere. Therefore

$$-\text{CVaR}_{\alpha}(r_{[0]}) = \int_{-1-\omega_1}^0 \gamma \alpha^{-1} F'(\omega_1 + \gamma) d\gamma = \frac{1}{\alpha} \int_{-1}^{\omega_1} (\omega - \omega_1) F'(\omega) d\omega = \frac{1}{\alpha} \int_{-1}^0 r_{[0]}(\omega) F'(\omega) d\omega. \quad (32)$$

Since actually  $r_{[s]} = r_{[0]} + sr_1$ , we have more generally even for  $s = \varepsilon > 0$  that  $r_{[\varepsilon]}(\omega) \leq 0$  exactly on the interval  $[-1, 0]$  having probability  $\alpha$ , so that

$$\text{VaR}_{\alpha}(r_{[\varepsilon]}) = 0, \quad \text{CVaR}_{\alpha}(r_{[\varepsilon]}) = -\frac{1}{\alpha} \int_{-1}^0 [r_{[0]}(\omega) + \varepsilon r_1(\omega)] dF_1(\omega) = \text{CVaR}_{\alpha}(r_{[0]}) + \varepsilon \text{CVaR}_{\alpha}(r_1).$$

It follows then in (30) and (31) that  $f'_+(0) = \text{CVaR}_{\alpha}(r_1)$ .

Dealing with  $f'_-(0)$  is harder and requires an estimate. When  $s = -\varepsilon < 0$ , we can only be sure that  $\text{VaR}_\alpha(r_{[-\varepsilon]}) = -\gamma^*$  for some  $\gamma^*$  in the interval  $(-\varepsilon\omega_2, -\varepsilon\omega_1)$  (which contains 0). There are no probability atoms in the distribution of  $r_{[-\varepsilon]}$ , so we have

$$-\text{CVaR}_\alpha(r_{[-\varepsilon]}) = \text{average of } r_{[-\varepsilon]} \text{ on } [-1, \omega'_1] \cup [\omega'_2, \omega''_2], \quad (33)$$

where the two intervals comprise the set of  $\omega$  where  $r_{[-\varepsilon]}(\omega) \leq \gamma^*$ . Here  $\omega'_1$  marks the point where the initial segment in the graph of  $r_{[-\varepsilon]}$  rises past  $\gamma^*$ , while  $\omega'_2$  marks where the middle segment sinks past  $\gamma^*$ , and  $\omega''_2$  marks where the final segment once more rises past  $\gamma^*$ ; cf. Figure 8. The total probability in these intervals is of course  $\alpha$ , and we have  $-1 < \omega'_1 < \omega_1 < \omega'_2 < \omega_2 < \omega''_2$ .

Let  $\alpha_1 = \text{prob}[-1, \omega_1]$  and  $\alpha_2 = \text{prob}[\omega_1, 0]$ ; then  $\alpha_1 + \alpha_2 = \alpha$  with  $\alpha_1 > \text{prob}[-1, \omega'_1]$  and  $\alpha_2 < \text{prob}[\omega'_2, \omega''_2]$ . Let  $\omega_0$  mark the unique point between  $\omega_1$  and  $\omega_2$  for which  $\text{prob}[\omega_0, \omega_2] = \alpha_2$ . Our claim, on the basis of (33), is that

$$-\text{CVaR}_\alpha(r_{[-\varepsilon]}) < \text{average of } r_{[-\varepsilon]} \text{ on } [-1, \omega_1] \cup [\omega_0, \omega_2]. \quad (34)$$

This holds because, in passing from  $[-1, \omega'_1] \cup [\omega'_2, \omega''_2]$  to  $[-1, \omega_1] \cup [\omega_0, \omega_2]$ , we preserve the total  $\omega$ -probability of the union, namely  $\alpha$ , while replacing segments where  $r_{[-\varepsilon]}(\omega) < \gamma^*$  by segments where  $r_{[-\varepsilon]}(\omega) > \gamma^*$ . We obtain from (34) that

$$-\text{CVaR}_\alpha(r_{[-\varepsilon]}) < \frac{1}{\alpha} \left[ \int_{-1}^{\omega_1} [r_{[0]}(\omega) - \varepsilon r_1(\omega)] F'(\omega) d\omega + \int_{\omega_0}^{\omega_2} [r_{[0]}(\omega) - \varepsilon r_1(\omega)] F'(\omega) d\omega \right], \quad (35)$$

where the first integral equals  $\int_{-1}^0 r_{[0]}(\omega) F'(\omega) d\omega - \varepsilon \int_{-1}^0 r_1(\omega) F'(\omega) d\omega + \varepsilon \int_{\omega_1}^0 r_1(\omega) F'(\omega) d\omega$  and the second is just  $-\varepsilon \int_{\omega_0}^{\omega_2} r_1(\omega) F'(\omega) d\omega$ . These observations, applied in (35), give us

$$-\text{CVaR}_\alpha(r_{[-\varepsilon]}) < -\text{CVaR}_\alpha(r_{[0]}) + \varepsilon \text{CVaR}_\alpha(r_1) - \varepsilon \delta \quad (36)$$

for a certain  $\delta$ , which in terms  $\alpha_2$ , the common probability of the two intervals, can be expressed by

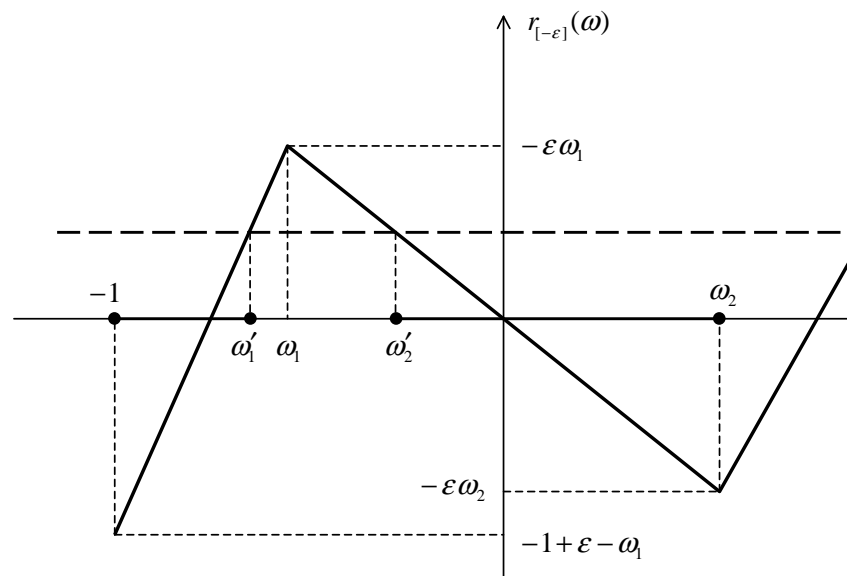
$$\delta = \frac{\alpha_2}{\alpha} \left[ \frac{1}{\alpha_2} \int_{\omega_0}^{\omega_2} r_1(\omega) F'(\omega) d\omega - \frac{1}{\alpha_2} \int_{\omega_1}^0 r_1(\omega) F'(\omega) d\omega \right].$$

Here  $\delta > 0$ , because the average of  $r_1(\omega) = \omega$  over  $[\omega_1, 0]$  is less than its average over  $[\omega_0, \omega_2]$ , an interval further to the right.

The estimate in (36) can be applied in (31) through the fact that  $-\text{CVaR}_\alpha(r_{[-\varepsilon]}) = -f(-\varepsilon)$  and  $\text{CVaR}_\alpha(r_{[0]}) = f(0)$ . This tells us that  $f'_-(0) \leq \text{CVaR}_\alpha(r_1) - \delta$ . Since we already know that  $f'_+(0) = \text{CVaR}_\alpha(r_1)$ , we deduce that  $f'_-(0) < f'_+(0)$ , which is what was needed.

## 8 Conclusions

The replacement of standard deviation by other deviations, such as arise from conditional value-at-risk and other risk notions, in accordance with current trends, by no means causes the classical approach to optimization out-dated. Instead, it enriches that approach by making a degree of customization available. One-fund theorems still reign as a way of simplification, even though the designated funds, in their dependence on the deviation measure, can be different for different classes of investors. Nevertheless, mathematical complications created by instruments involving options, or models set up with only finitely many scenarios, require techniques beyond ordinary differential calculus to formulate and obtain results rigorously.





## References

- [1] C. ACERBI, Spectral Measures of Risk: a Coherent Representation of Subjective Risk Aversion, *Journal of Banking and Finance* **26** (2002), 1505–1518.
- [2] C. ACERBI, P. SIMONETTI, Portfolio Optimization with Spectral Measures of Risk, preprint 2002.
- [3] C. ACERBI, D. TASCHE, Expected Shortfall: A Natural Coherent Alternative to Value at Risk, *Economic Notes* **31** (2002), 379–388.
- [4] G. ALEXANDER, A. BAPTISTA, Economic Implications of Using Mean-VaR Model for Portfolio Selection: A Comparison with Mean-Variance Analysis, *Economic Dynamics and Control* **7** (2002).
- [5] P. ARTZNER, F. DELBAEN, J.-M. EBER, D. HEATH, Coherent Measures of Risk, *Mathematical Finance* **9** (1999), 203–227.
- [6] V.S. BAWA, E.B. LINDENBERG, Capital Market Equilibrium in a Mean-Lower Partial Moment Framework, *Journal of Financial Economics* **5** (1977), 189–200.
- [7] F. DELBAEN, Coherent Risk Measures on General Probability Spaces, preprint 2000.
- [8] H. FÖLLMER, A. SCHIED, *Stochastic Finance*, deGruyter, 2002.
- [9] C.F. HUANG, R.H. LITZENBERGER, *Foundations for Financial Economics*, North Holland, 1988.
- [10] Y. KROLL, H. LEVY, H.M. MARKOWITZ, Mean-Variance versus Direct Utility Maximization, *The Journal of Finance* **3** (1984), 47–61.
- [11] H. KONNO, H. SHIRAKAWA, Equilibrium Relations in the Mean-Absolute Deviation Capital Market, *Asia-Pacific Financial Markets* **1** (1994), 21–35.
- [12] D.G. LUENBERGER, *Investment Science*, Oxford University Press, Oxford, New York, 1998.
- [13] Y. MALEVERGNE, D. SORNETTE, Multi-Moments Method for Portfolio Management: Generalized Capital Asset Pricing Model in Homogeneous and Heterogeneous markets, preprint 2002.
- [14] R. MANSINI, W. OGRYCZAK, M. GRAZIA SPERANZA, Conditional Value at Risk and Related Linear Programming Models for Portfolio Optimization, Warsaw University of Technology, Institute of Control and Computation Engineering, Report 03-02 (2003).
- [15] H.M. MARKOWITZ, Foundations of Portfolio Theory, *Journal of Finance* **46** (1991), 469–477.
- [16] H.M. MARKOWITZ, Portfolio Selection, *The Journal of Finance* **7** (1952), 77–91.
- [17] W. OGRYCZAK, A. RUSZCZYNSKI, From Stochastic Dominance to Mean-Risk Models: Semideviations as Risk Measures, *European Journal of Operational Research* **116** (1999), 33–50.
- [18] W. OGRYCZAK, A. RUSZCZYNSKI, Dual Stochastic Dominance and Related Mean-Risk Models, *SIAM Journal on Optimization* **13** (2002), 60–78.

- [19] R.T. ROCKAFELLAR, *Convex Analysis*, Princeton University Press, Princeton, New Jersey, 1970.
- [20] R.T. ROCKAFELLAR, *Conjugate Duality and Optimization*, No. 16 in Conference Board of Math. Sciences Series, SIAM Publications, 1974.
- [21] R.T. ROCKAFELLAR, S.P. URYASEV, Optimization of Conditional Value-at-Risk, *Journal of Risk* **2** (2000), 21–42.
- [22] R.T. ROCKAFELLAR, S.P. URYASEV, Conditional Value-at-Risk for General Loss Distributions, *Journal of Banking and Finance* **26** (2002), 1443–1471.
- [23] R.T. ROCKAFELLAR, S. URYASEV, M. ZABARANKIN, Deviation Measures in Risk Analysis and Optimization, *Research Report 2002-7*, Department of Industrial and Systems Engineering, University of Florida, 2002.
- [24] R.T. ROCKAFELLAR, S. URYASEV, M. ZABARANKIN, Deviation Measures in Generalized Linear Regression, *Research Report 2002-9*, Department of Industrial and Systems Engineering, University of Florida, 2002.
- [25] R.T. ROCKAFELLAR, S. URYASEV, M. ZABARANKIN, CAPM-like Relations in Portfolio Analysis with General Deviation Measures, forthcoming.
- [26] W.F. SHARPE, Capital Asset Prices With and Without Negative Holdings, *The Journal of Finance* **46** (1991), 489–509.
- [27] W.F. SHARPE, Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk, *The Journal of Finance* **19** (1964), 425–442.
- [28] D. TASCHE, Expected Shortfall and Beyond, *Journal of Banking and Finance* **26** (2002), 1519–1533.
- [29] J. TOBIN, Liquidity Preference as Behavior Towards Risk, *The Review of Economic Studies* **25** (1958), 65–86.