

# Markowitz Revisited: Mean-Variance Models in Financial Portfolio Analysis\*

---

Marc C. Steinbach<sup>†</sup>

**Abstract.** Mean-variance portfolio analysis provided the first quantitative treatment of the tradeoff between profit and risk. We describe in detail the interplay between objective and constraints in a number of single-period variants, including semivariance models. Particular emphasis is laid on avoiding the penalization of overperformance. The results are then used as building blocks in the development and theoretical analysis of multiperiod models based on scenario trees. A key property is the possibility of removing surplus money in future decisions, yielding approximate downside risk minimization.

**Key words.** mean-variance analysis, downside risk, multiperiod model, stochastic optimization

**AMS subject classifications.** 90A09, 90C15, 90C20

**PII.** S0036144500376650

---

**0. Introduction.** The classical mean-variance approach for which Harry Markowitz received the 1990 Nobel Prize in Economics offered the first systematic treatment of a dilemma that each investor faces: the conflicting objectives of *high profit* versus *low risk*. In dealing with this fundamental issue Markowitz came up with a parametric optimization model that was both sufficiently general for a significant range of practical situations and simple enough for theoretical analysis and numerical solution. As the Swedish Academy of Sciences put it [154], “his primary contribution consisted of developing a rigorously formulated, operational theory for portfolio selection under uncertainty.”

Indeed, the subject is so complex that Markowitz’s seminal work of the 1950s [134, 135, 137] probably raised more questions than it answered, thus initiating a tremendous amount of related research. Before placing the present paper into perspective, the following paragraphs give a coarse overview of these issues. A substantial number of references are included, but we have not attempted to compile a complete list. (The 1982 research bibliography [10] contains 400 references on just one of the topics.) However, we have tried to cite (mostly in chronological order) at least several major papers on each subject to provide some starting points for the interested reader.

An important aspect of pareto-optimal (efficient) portfolios is that each determines a von Neumann–Morgenstern utility function [202] for which it maximizes the *expected utility* of the return on investment. This allowed Markowitz to interpret his

---

\*Received by the editors August 30, 1999; accepted for publication (in revised form) August 4, 2000; published electronically February 2, 2001.

<http://www.siam.org/journals/sirev/43-1/37665.html>

<sup>†</sup>Konrad-Zuse-Zentrum für Informationstechnik Berlin (ZIB), Takustr. 7, 14195 Berlin, Germany (steinbach@zib.de, <http://www.zib.de/steinbach>).

approach by the theory of rational behavior under uncertainty [135], [137, Part IV]. Further, certain measures of *risk averseness* evolved as a basic concept in economic theory. These are derived from utility functions and justified by their relationship to the corresponding risk premiums. Work in this area includes Tobin [198], Pratt [160], Lintner [129], Arrow [2], Rubinstein [170], Kihlstrom and Mirman [100, 101], Fishburn and Porter [59], Duncan [46], Kira and Ziemba [107], Ross [169], Chamberlain [28], Hubermann and Ross [87], Epstein [55], Pratt and Zeckhauser [161], and Li and Ziemba [126, 127]. Applications of utility theory and risk averseness measures to portfolio selection were reported, e.g., by Tobin [199], Mossin [150], Hanoch and Levy [75], Levy and Markowitz [123], Kallberg and Ziemba [96], Kroll, Levy, and Markowitz [119], Jewitt [91, 92], King and Jensen [106], Kijima and Ohnishi [103], and Kroll et al. [118].

A fundamental (and still debated) question is how risk should be measured. Markowitz discussed the pros and cons of replacing the variance by alternative risk measures in a more general mean-risk approach [137, Chap. XIII]. These considerations and the theory of stochastic dominance (see Bawa [8, 9, 10], Fishburn [58], Levy [121], Kijima and Ohnishi [104], and Levy and Wiener [125]) stimulated the research in asymmetric risk measures like expectation of loss and semivariance; cf. Bawa and Lindenberg [11], Harlow and Rao [76], Konno [108], Konno and Yamazaki [115], King [105], Markowitz et al. [139], Zenios and Kang [206], Embrechts, Klüppelberg, and Mikosh [51], Gaese [66], Ogryczak and Ruszczyński [155], Rockafellar and Uryasev [166], and Uryasev [200]. The properties of real return distributions also led to risk models involving higher moments; see Ziemba [207], Kraus and Litzenberger [117], Konno, Shirakawa, and Yamazaki [112], and Konno and Suzuki [113]. More recently the theoretical concept of *coherent* risk measures was introduced and further developed by Artzner et al. [3, 4] and Embrechts, Resnick, and Samorodnitsky [52], while portfolio tracking (or replication) approaches became popular in practice; see Dembo [39], Guerard, Takano, and Yamane [69], King [105], Konno and Watanabe [114], Buckley and Korn [26], and Dembo and Rosen [40].

It is quite interesting that the mean-variance approach has received comparatively little attention in the context of long-term investment planning. Although Markowitz did consider true multiperiod models (where the portfolio may be readjusted several times during the planning horizon) [137, Chap. XIII], these considerations used a utility function based on the consumption of wealth over time rather than mean and variance of the final wealth. Other long-term and simplified multiperiod approaches were discussed, e.g., by Phelps [158], Tobin [199], Mossin [150], Samuelson [172], Fama [57], Hakansson [70, 71, 72, 73, 74], Stevens [197], Roll [168], Merton and Samuelson [145], Machina [132], Konno, Pliska, and Suzuki [110], Luenberger [131], and Elliott and van der Hoek [49]. Combined consumption-portfolio strategies have been investigated in many of the works just cited (using discounted utility of consumption and/or utility of final wealth as optimization criteria), and the relation to myopic (short-term) mean-risk efficiency with long-term optimality and long-term capital growth has been studied. Research in the closely related field of continuous-time portfolio management models has often been based on similar approaches; see Merton [141, 142], Sengupta [183], Heath et al. [83], Karatzas, Lehoczky, and Shreve [98], Cox and Huang [32], Karatzas [97], Richardson [163], Dohi and Osaki [43], and Bajeux-Besnainou and Portait [6].

Over roughly the past decade, large-scale real-life models, in particular detailed multiperiod models, have become tractable due to progress in computing technology

(both algorithms and hardware); see, e.g., Perold [156], Mulvey [151], Dempster [42], Glover and Jones [67], Mulvey and Vladimirov [152], Dantzig and Infanger [37], Cariño et al. [27], Consigli and Dempster [31], Beltratti, Consiglio, and Zenios [13], and Gondzio and Kouwenberg [68]. With the exception of the references just given, much of the work cited above neglected details like asset liquidity or transaction costs. At least the second idealization (no transaction costs) causes serious errors when many transactions are performed, as in continuous-time models. Imperfect markets were briefly discussed by Markowitz [137, p. 297ff.] and later (in both discrete and continuous time) by Pogue [159], Chen, Jen, and Zions [29], Perold [156], He and Pearson [81, 82], Karatzas et al. [99], Cvitanic and Karatzas [35, 36], Jacka [90], Shirakawa and Kassai [189], Shirakawa [188], Morton and Pliska [148], Atkinson, Pliska, and Wilmott [5], and Buckley and Korn [26].

A final issue in the context of portfolio selection concerns the assumptions of the investor about the future, which are represented by probability distributions of the asset returns. Being based on assessments of financial analysts or estimated from historical data (or both), these distributions are never exact. (Markowitz called them *probability beliefs*.) The question of the sensitivity of optimization results with respect to errors in the distribution was discussed, e.g., by Best and Grauer [16], Jobson [93], Broadie [25], Chopra and Ziemba [30], Best and Ding [15], and MacLean and Weldon [133].

Another central topic in modern finance is the theory (and prediction) of the behavior of asset prices in capital markets. The original *capital asset pricing model* (CAPM) is based directly on Markowitz's static mean-variance analysis and on the assumption of market equilibrium; cf. Sharpe [184] (who shared the 1990 Nobel Prize jointly with Markowitz and Miller), Lintner [128], and Mossin [149]. The model was later extended to a dynamic setting by Merton [143]; further work on the behavior of asset prices and interest rates includes Vasicek [201], Cox, Ingersoll, and Ross [33, 34], Ho and Lee [85], Bollerslev, Engle, and Wooldridge [23], Hull and White [88], Levy and Samuelson [124], Konno and Shirakawa [111], Konno [109], and Levy [122]. The observation that volatilities of asset returns and other factors change over time has led to the development of the *generalized autoregressive conditional heteroskedasticity* (GARCH) models; see Engle [53], Bollerslev [20], Bollerslev, Chou, and Kroner [21], Bollerslev, Engle, and Nelson [22], and Engle and Kroner [54]. An alternative approach for the multivariate case was given by Harvey, Ruiz, and Shephard [80].

A final major field concerns the hedging of options or, more generally, contingent claims. The typical objective in hedging an option is to eliminate (or reduce) the risk of a future commitment to some asset. This involves an optimal dynamic trading strategy that also determines the fair price of the option. For their pioneering work in that area, Black and Scholes [19] and Merton [144] received the Nobel Prize in 1997. Thereafter, Harrison and Kreps [77] and Harrison and Pliska [78, 79] introduced martingales and semimartingales in the theoretical treatment; these concepts also replaced the earlier stochastic dynamic programming perspective in continuous-time consumption-investment models. A quadratic risk measure for hedging strategies was proposed by Föllmer and Sondermann [62] for the case in which the asset price process is a martingale; Schweizer [175] extended this to the semimartingale case. Further research in the general area of hedging is largely concerned with the investigation and extension of similar concepts, particularly in incomplete markets; cf. Bouleau and Lambertson [24], Föllmer and Schweizer [61], Rabinovich [162], Long [130], Duffie and Richardson [45], Duffie [44], Hofmann, Platen, and Schweizer [86], Schweizer [176,

177, 178, 179, 180, 181, 182], Edirisinghe, Naik, and Uppal [47], Schäl [174], Delbaen and Schachermayer [38], El Karoui and Quenez [48], Monat and Stricker [147], Soner, Shreve, and Cvitanic [190], Kramkov [116], Lamberton, Pham, and Schweizer [120], Pham, Rheinländer, and Schweizer [157], Föllmer and Leukert [60], and Heath and Schweizer [84].

We conclude the general discussion by pointing out that dynamic programming (Bellman [12]) and its extension to stochastic differential equations play a central role in much of the early work involving multiperiod and continuous-time models; this pertains to theoretical considerations as well as actual solution procedures. Prominent examples, to name just a few, include Markowitz [137, Chap. XIII], Mossin [150], Samuelson [172], Fama [57] in discrete time, and Merton [142] in continuous time. Since straightforward dynamic programming becomes computationally expensive in complex problems (particularly in the presence of inequality constraints), its practical applicability is basically limited to structurally simple models. In such idealized cases, however, closed-form expressions are often obtained for the optimal strategies. Closed-form solutions in continuous time, with asset prices being described by Brownian motions, were also given by Richardson [163] (for mean-variance optimization of a portfolio consisting of a riskless bond and a single stock) and by Duffie and Richardson [45] (for futures hedging policies under mean-variance and quadratic objectives).

Additional material and references can be found in a more recent book by Markowitz [138] or in any standard text on mathematical finance, like Sharpe [185, 186], Elton and Gruber [50], Ingersoll [89], Alexander and Sharpe [1], Merton and Samuelson [146], Zenios [205], and Ziemba and Mulvey [208].

The present paper develops a fairly complete theoretical understanding of the multiperiod mean-variance approach based on scenario trees. This is achieved by analyzing various portfolio optimization problems with gradually increasing complexity. Primal and dual solutions of these problems are derived, and dual variables are given an interpretation where possible. The most important aspect in our discussion is the precise interaction of objective (or risk measure) and constraints (or set of feasible wealth distributions), a subject that has not much been studied in the previous literature. It should be obvious that arguing the properties of risk measures may be meaningless in an optimization context unless it is clear which distributions are possible. A specific goal in our analysis is to avoid penalization due to overperformance. In this context we discuss the role of cash and, in some detail, variance versus semivariance. A key ingredient of our most complex multiperiod model is an artificial arbitrage-like mechanism involving riskless though inefficient portfolios and representing a choice between immediate consumption and future profit.

Each of the problems considered tries to isolate a certain aspect, usually under the most general conditions even if practical situations typically exhibit more specific characteristics. However, we give higher priority to a clear presentation, and inessential generality will sometimes be sacrificed for technical simplicity. In particular, no inequality constraints are included except where necessary. (A separate section is devoted to the influence of such restrictions.) Neither do we attempt to model liquidity constraints or short-selling correctly, nor to include transaction costs; we consider only idealized situations without further justification. The present work grew out of a close cooperation with the Institute of Operations Research at the University of St. Gallen. It is based on a multiperiod mean-variance model that was first proposed by Frauendorfer [63], then refined by Frauendorfer and Siede [65], and later extended to a complete application model including transaction costs and market restrictions.

That model raised some of the theoretical questions treated here; it will be presented later in a joint paper.

Due to future uncertainty the portfolio optimization problems in this paper are all stochastic. More precisely, they are *deterministic equivalents* of convex stochastic programs; cf. Wets [203]. Except for the semivariance problems, they are also quadratic programs involving a second-order approximation of the return distribution in some sense; cf. Samuelson [173]. Based on earlier work in nonlinear optimal control [191, 192, 196], we previously developed structure-exploiting numerical algorithms for multistage convex stochastic programs like the ones discussed here [193, 194, 195]. Closely related but more general problem classes and duality are studied by Rockafellar and Wets [167] and Rockafellar [165]. For background material on stochastic programming we refer the reader to Kall [94], Dempster [41], Ermoliev and Wets [56], Kall and Wallace [95], Birge [17], Birge and Louveaux [18], and Ruszczyński [171]; for discrete-time stochastic control, see Bertsekas and Shreve [14]. A comprehensive treatment of probability theory was given by Bauer [7], and advanced convex analysis was treated by Rockafellar [164].

The paper is organized as follows. Our analysis begins with single-period models in section 1. Although many of the results are already known, the systematic discussion of subtle details adds insight that is essential in the multiperiod case. To some extent this section has a tutorial character; the problems may serve as examples in an introductory course on optimization. Next, multiperiod mean-variance models are analyzed in section 2, where the final goal consists of constructing an approximate downside risk minimization through appropriate constraints. To the best of our knowledge, this material is new; the research was motivated by practical experience with the application model mentioned above. Some concluding remarks are given in section 3.

**I. Single-Period Mean-Variance Analysis.** Consider an investment in  $n$  assets over a certain period of time. Denote by  $x^\nu$  the capital invested in asset  $\nu$ , by  $x \in \mathbb{R}^n$  the *portfolio* vector, and by  $r \in \mathbb{R}^n$  the random vector of asset returns, yielding asset capitals  $r^\nu x^\nu$  at the end of the investment period. Suppose that  $r$  is given by a joint probability distribution with expectation  $\bar{r} := \mathbf{E}(r)$  and covariance matrix

$$\Sigma := \mathbf{E}[(r - \bar{r})(r - \bar{r})^*] = \mathbf{E}[rr^*] - \bar{r}\bar{r}^*.$$

(The existence of these two moments is assumed throughout the paper.) The choice of a specific portfolio determines a certain distribution of the associated total return (or final wealth)  $w \equiv r^*x$ . Mean-variance analysis aims at forming the most desirable return distribution through a suitable portfolio, where the investor's idea of desirability depends solely on the first two moments.

DEFINITION 1.1 (reward). *The reward of a portfolio is the mean of its return,*

$$\rho(x) := \mathbf{E}(r^*x) = \bar{r}^*x.$$

DEFINITION 1.2 (risk). *The risk of a portfolio is the variance of the return,*

$$R(x) := \sigma^2(r^*x) = \mathbf{E}[(r^*x - \mathbf{E}(r^*x))^2] = \mathbf{E}[x^*(r - \bar{r})(r - \bar{r})^*x] = x^*\Sigma x.$$

Various formulations of the mean-variance problem exist. Although Markowitz was well aware that “the Rational Man, like the unicorn, does not exist” [137, p. 206], he related his approach to the utility theory of von Neumann and Morgenstern [202] from the very beginning. This provides an important theoretical justification on the

grounds that “the ‘fun of the game’ can be ignored in deciding on a rationale for the selection of a portfolio, especially when this involves the allocation of large amounts of other people’s money” [137, p. 226]. As we shall see later, maximizing the expectation of a concave quadratic utility function leads to a formulation like

$$(1.1) \quad \begin{aligned} & \max_x \quad \mu\rho(x) - \frac{1}{2}R(x) \\ & \text{subject to (s.t.)} \quad e^*x = 1, \end{aligned}$$

where  $e \in \mathbb{R}^n$  denotes the vector of all 1s. The objective models the actual goal of the investor, a tradeoff between risk and reward,<sup>1</sup> while the *budget equation*  $e^*x = 1$  simply specifies the initial wealth  $w_0$  (normalized without loss of generality to  $w_0 = 1$ ). Our preferred formulation comes closer to the original one; it minimizes risk subject to the budget equation and subject to the condition that a certain target reward  $\rho$  be obtained,

$$(1.2) \quad \begin{aligned} & \min_x \quad \frac{1}{2}R(x) \\ & \text{s.t.} \quad e^*x = 1, \\ & \quad \quad \rho(x) = \rho. \end{aligned}$$

Here the investor’s goal is *split* between objective and reward condition.

In this section we study the precise relation of problems (1.1) and (1.2) and a number of increasingly general single-period variants. We will include a cash account, then consider certain inequality constraints, utility functions, and finally downside risk. Many of the results are already known, but usually in a different form. Here we choose a presentation that facilitates the study of nuances in the optimization problems and that integrates seamlessly with the more general case of multiperiod problems in section 2.

**1.1. Risky Assets Only.** The simplest situation is given by portfolios consisting exclusively of risky assets. In this case we impose two conditions on the return distribution.

BASIC ASSUMPTIONS.

(A1) The covariance matrix is positive definite,  $\Sigma > 0$ .

(A2) The expectation  $\bar{r}$  is not a multiple of  $e$ .

*Remarks.* The first assumption means that all  $n$  assets (and any convex combination) are indeed risky; riskless assets like cash will be treated separately if present. The second assumption implies  $n \geq 2$  and guarantees a nondegenerate situation, otherwise problem (1.1) would always have the same optimal portfolio  $x = \Sigma^{-1}e/(e^*\Sigma^{-1}e)$  regardless of the tradeoff parameter, and problem (1.2) would have inconsistent constraints except for one specific value of the target reward:  $\rho = \bar{r}^*e/n$ . Notice that no formal restrictions are imposed on the value of  $\bar{r}$ , although  $\bar{r} > 0$  (and even  $\bar{r} > e$ ) will usually hold in practice.

Due to assumption (A1) we can define the following constants that will be used throughout this section:

$$\alpha := e^*\Sigma^{-1}e, \quad \beta := e^*\Sigma^{-1}\bar{r}, \quad \gamma := \bar{r}^*\Sigma^{-1}\bar{r}, \quad \delta := \alpha\gamma - \beta^2.$$

<sup>1</sup>Many authors attach a tradeoff parameter  $\theta$  to the *risk* term and maximize  $\rho(x) - \theta R(x)/2$ , which is equivalent to (1.1) if  $\mu \equiv \theta^{-1} > 0$ . However, this problem becomes unbounded for  $\theta \leq 0$ , whereas (1.1) remains solvable for  $\mu \leq 0$ . This is important in our analysis.

LEMMA 1.3. *The constants  $\alpha$ ,  $\gamma$ , and  $\delta$  are positive. More precisely,*

$$\alpha \in \left[ \frac{n}{\lambda_{\max}(\Sigma)}, \frac{n}{\lambda_{\min}(\Sigma)} \right], \quad \gamma \in \left[ \frac{\|\bar{r}\|_2^2}{\lambda_{\max}(\Sigma)}, \frac{\|\bar{r}\|_2^2}{\lambda_{\min}(\Sigma)} \right], \quad |\beta| < \frac{\sqrt{n} \|\bar{r}\|_2}{\lambda_{\min}(\Sigma)},$$

where  $\lambda_{\min}, \lambda_{\max}$  denote the minimal and maximal eigenvalue of  $\Sigma$ , respectively.

*Proof.* Since  $\Sigma > 0$  (by (A1)), we have

$$\alpha = e^* \Sigma^{-1} e \in [\|e\|_2^2 \lambda_{\min}(\Sigma^{-1}), \|e\|_2^2 \lambda_{\max}(\Sigma^{-1})] = \left[ \frac{n}{\lambda_{\max}(\Sigma)}, \frac{n}{\lambda_{\min}(\Sigma)} \right].$$

The inclusion for  $\gamma$  is analogous. Since  $\bar{r}$  and  $e$  are linearly independent and  $\Sigma > 0$  (by (A2) and (A1)), the  $2 \times 2$  matrix

$$\begin{pmatrix} e^* \\ \bar{r}^* \end{pmatrix} \Sigma^{-1} \begin{pmatrix} e \\ \bar{r} \end{pmatrix} = \begin{pmatrix} \alpha & \beta \\ \beta & \gamma \end{pmatrix} > 0$$

has positive determinant  $\delta$ . Thus  $|\beta| < \sqrt{\alpha\gamma} \leq \sqrt{n} \|\bar{r}\|_2 / \lambda_{\min}(\Sigma)$ .  $\square$

*Remark.* The inclusions for  $\alpha$  and  $\gamma$  are sharp but not the bound on  $|\beta|$ , and neither  $\alpha < \gamma$  nor  $\beta > 0$  hold in general. In any case, we need only  $\alpha, \gamma, \delta > 0$ .

PROBLEM 1. Let us first consider the standard tradeoff formulation. To simplify the comparison with our preferred formulation, we minimize negative utility

$$\begin{aligned} \min_x \quad & \frac{1}{2} x^* \Sigma x - \mu \bar{r}^* x \\ \text{s.t.} \quad & e^* x = 1. \end{aligned}$$

The Lagrangian is

$$L(x, \lambda; \mu) = \frac{1}{2} x^* \Sigma x - \mu \bar{r}^* x - \lambda (e^* x - 1).$$

THEOREM 1.4. *Problem 1 has the unique primal-dual solution*

$$x = \Sigma^{-1} (\lambda e + \mu \bar{r}), \quad \lambda = (1 - \mu \beta) / \alpha$$

and associated reward

$$\rho = \lambda \beta + \mu \gamma = (\beta + \mu \delta) / \alpha.$$

*Proof.* From the Lagrangian one obtains the system of first-order necessary conditions

$$\begin{pmatrix} \Sigma & e \\ e^* & \end{pmatrix} \begin{pmatrix} x \\ -\lambda \end{pmatrix} = \begin{pmatrix} \mu \bar{r} \\ 1 \end{pmatrix}.$$

Its first row (dual feasibility) yields the optimal portfolio  $x$ . The optimal multiplier  $\lambda$  and reward  $\rho$  are obtained by substituting  $x$  into the second row (primal feasibility) and the definition of  $\rho$ , respectively. Uniqueness of the solution follows from strong convexity of the objective and full rank of the constraint.  $\square$

*Remark.* Although the qualitative interpretation of the tradeoff function is clear, the precise value of the tradeoff parameter  $\mu$  should also have an interpretation. In particular, the resulting reward is of interest. This is one reason why we prefer a different formulation of the mean-variance problem. (Other important reasons are greater modeling flexibility and sparsity in the multiperiod formulation; see section 2.)

PROBLEM 2. The mean-variance problem with prescribed reward reads

$$\begin{aligned} \min_x \quad & \frac{1}{2}x^*\Sigma x \\ \text{s.t.} \quad & e^*x = 1, \\ & \bar{r}^*x = \rho. \end{aligned}$$

Its Lagrangian is

$$L(x, \lambda, \mu; \rho) = \frac{1}{2}x^*\Sigma x - \lambda(e^*x - 1) - \mu(\bar{r}^*x - \rho).$$

We refer to the dual variables  $\lambda, \mu$  as the *budget multiplier* and the *reward multiplier*, respectively. It will soon be shown that the optimal reward multiplier  $\mu$  is precisely the tradeoff parameter of Problem 1.

THEOREM 1.5. *Problem 2 has the unique primal-dual solution*

$$x = \Sigma^{-1}(\lambda e + \mu \bar{r}), \quad \lambda = (\gamma - \beta\rho)/\delta, \quad \mu = (\alpha\rho - \beta)/\delta.$$

*Proof.* The system of first-order optimality conditions reads

$$\begin{pmatrix} \Sigma & e & \bar{r} \\ e^* & & \\ \bar{r}^* & & \end{pmatrix} \begin{pmatrix} x \\ -\lambda \\ -\mu \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ \rho \end{pmatrix}.$$

As in Theorem 1.4, the optimal portfolio  $x$  is obtained from the first row. Substitution of  $x$  into rows two and three yields the optimal multipliers

$$\begin{pmatrix} \lambda \\ \mu \end{pmatrix} = \left[ \begin{pmatrix} e^* \\ \bar{r}^* \end{pmatrix} \Sigma^{-1} (e \ \bar{r}) \right]^{-1} \begin{pmatrix} 1 \\ \rho \end{pmatrix} = \begin{pmatrix} \alpha & \beta \\ \beta & \gamma \end{pmatrix}^{-1} \begin{pmatrix} 1 \\ \rho \end{pmatrix} = \frac{1}{\delta} \begin{pmatrix} \gamma - \beta\rho \\ \alpha\rho - \beta \end{pmatrix}.$$

Uniqueness of the solution follows as in Theorem 1.4.  $\square$

THEOREM 1.6. *Problem 1 with parameter  $\mu$  and Problem 2 with parameter  $\rho$  are equivalent if and only if  $\mu$  equals the optimal reward multiplier of Problem 2 or, equivalently,  $\rho$  equals the optimal reward of Problem 1.*

*Proof.* The required conditions,  $\mu = (\alpha\rho - \beta)/\delta$  and  $\rho = (\beta + \mu\delta)/\alpha$ , are clearly equivalent. It follows that the optimal budget multipliers of both problems are identical,

$$\frac{1 - \mu\beta}{\alpha} = \frac{\delta - \alpha\beta\rho + \beta^2}{\alpha\delta} = \frac{\alpha\gamma - \alpha\beta\rho}{\alpha\delta} = \frac{\gamma - \beta\rho}{\delta}.$$

Hence optimal portfolios also agree. The “only if” direction is trivial.  $\square$

*Remarks.* Apparently, the optimality conditions of Problem 2 include the optimality conditions of Problem 1, and additionally the reward condition. These  $n + 2$  equations define a one-dimensional affine subspace for the  $n + 3$  variables  $x, \lambda, \mu, \rho$ , which is parameterized by  $\mu$  in Problem 1 and by  $\rho$  in Problem 2. As an immediate consequence, the optimal risk is a quadratic function of  $\rho$ , denoted by  $\sigma^2(\rho)$ . Its graph is called the *efficient frontier*.<sup>2</sup>

<sup>2</sup>More generally, the efficient frontier refers to the set of all pareto-optimal solutions in any multi-objective optimization problem. The solutions (portfolios) are also called efficient. Strictly speaking, this applies only to the *upper branch* here, that is,  $\rho \geq \hat{\rho}$ , or, equivalently,  $\mu \geq 0$  (see the following discussion).



THEOREM 1.7. *In Problems 1 and 2, the optimal risk is*

$$\sigma^2(\rho) = (\alpha\rho^2 - 2\beta\rho + \gamma)/\delta = (\mu^2\delta + 1)/\alpha.$$

*Its global minimum over all rewards is attained at  $\hat{\rho} = \beta/\alpha$  and has the positive value  $\sigma^2(\hat{\rho}) = 1/\alpha$ . The associated solution is  $\hat{x} = \Sigma^{-1}e/\alpha$ ,  $\hat{\lambda} = 1/\alpha$ ,  $\hat{\mu} = 0$ .*

*Proof.* By Definition 1.2 and Theorem 1.5,

$$\begin{aligned}\sigma^2(\rho) &= x^*\Sigma x = (\lambda e + \mu\bar{r})^*\Sigma^{-1}(\lambda e + \mu\bar{r}) \\ &= \lambda^2\alpha + 2\lambda\mu\beta + \mu^2\gamma = \lambda(\lambda\alpha + \mu\beta) + \mu(\lambda\beta + \mu\gamma).\end{aligned}$$

Using  $\lambda, \rho$  from Theorem 1.4 and  $\lambda, \mu$  from Theorem 1.5 gives

$$\sigma^2(\rho) = \lambda + \mu\rho = (\alpha\rho^2 - 2\beta\rho + \gamma)/\delta = (\mu^2\delta + 1)/\alpha.$$

The remaining statements follow trivially.  $\square$

*Discussion.* The optimal portfolio is clearly a reward-dependent linear combination of the reward-independent portfolios  $\Sigma^{-1}e$  and  $\Sigma^{-1}\bar{r}$ . Moreover, it is an affine function of  $\rho$ . The efficient frontier and optimal investments into two risky assets are depicted in Figure 1.1. Here, since  $n = 2$ , the optimal portfolio is completely determined by the budget condition and the reward condition; it does not depend on  $\Sigma$  and is thus correlation-independent. Not so the risk: for negatively correlated assets, it has a pronounced minimum at a fairly large reward  $\hat{\rho}$ . As the correlation increases, the lowest possible risk is attained at a smaller reward and has a larger value. (These statements do not generalize simply to the case  $n > 2$ .)

A serious drawback of the model (in this form) is the fact that positive deviations from the prescribed reward are penalized, and hence the “risk” increases when  $\rho$  is reduced below  $\hat{\rho}$ . Indeed, the penalization cannot be avoided, indicating that the model is somehow incomplete. We will see, however, that unnecessary positive deviations from  $\rho$  do not occur if the model is extended appropriately. For the moment let us accept that only the upper branch is relevant in practice.

**1.2. Risky Assets and Riskless Cash.** Now consider  $n$  risky assets and an additional cash account  $x^c$  with deterministic return  $r^c \equiv \bar{r}^c$ . The portfolio is  $(x, x^c)$ , and  $x, r, \bar{r}, \Sigma$  refer only to its risky part.

BASIC ASSUMPTIONS. Assumption (A2) is replaced by a similar condition on the extended portfolio, which may now consist of just one risky asset and cash.

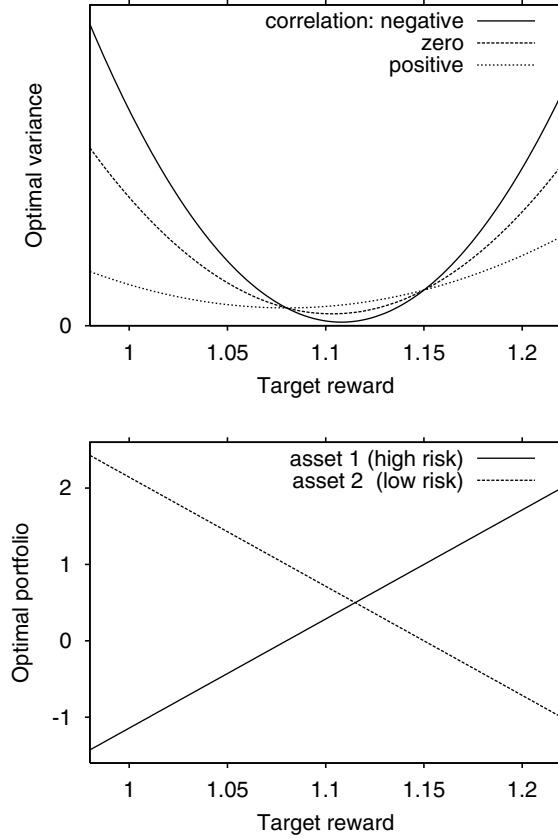
$$(A1) \quad \Sigma > 0.$$

$$(A3) \quad \bar{r} \neq r^c e.$$

*Remarks.* Again, the second assumption excludes degenerate situations, and no restrictions are imposed to ensure realistic returns. In practice one can typically assume  $\bar{r} > r^c e > 0$  (or even  $r^c e \geq e$ ), which satisfies (A3). The constants  $\alpha, \beta, \gamma$  are defined as before; they are related to the risky part of the portfolio only. (Condition (A3) makes  $\delta = 0$  possible, but  $\delta$  plays no role here.)

PROBLEM 3. Any covariance associated with cash vanishes, so that the risk and reward are  $R(x, x^c) = x^*\Sigma x$  and  $\rho(x, x^c) = \bar{r}^*x + r^c x^c$ , respectively, and the optimization problem reads

$$\begin{aligned}\min_{x, x^c} \quad & \frac{1}{2} \begin{pmatrix} x \\ x^c \end{pmatrix}^* \begin{pmatrix} \Sigma & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ x^c \end{pmatrix} = \frac{1}{2} x^* \Sigma x \\ \text{s.t.} \quad & e^* x + x^c = 1, \\ & \bar{r}^* x + r^c x^c = \rho.\end{aligned}$$



**Fig. 1.1** Portfolio with two risky assets having expected returns  $\bar{r}^1 = 1.15$  and  $\bar{r}^2 = 1.08$ . Top: Efficient frontier for negatively correlated, uncorrelated, and positively correlated assets. Bottom: Optimal portfolio versus reward.

THEOREM 1.8. Problem 3 has the unique primal-dual solution

$$\begin{aligned} x &= \Sigma^{-1}(\lambda e + \mu \bar{r}) = \mu \Sigma^{-1}(\bar{r} - r^c e), & \lambda &= -r^c \mu, \\ x^c &= 1 - \mu(\beta - r^c \alpha), & \mu &= (\rho - r^c)/\delta^c, \end{aligned}$$

where  $\delta^c := (r^c)^2 \alpha - 2r^c \beta + \gamma > 0$ . The resulting optimal risk is

$$\sigma^2(\rho) = (\rho - r^c)^2 / \delta^c.$$

Its global minimum over all rewards is attained at  $\hat{\rho} = r^c$  and has value zero. The associated solution has 100% cash:  $(\hat{x}, \hat{x}^c) = (0, 1)$ ,  $\hat{\lambda} = \hat{\mu} = 0$ .

*Proof.* The system of optimality conditions is

$$\begin{pmatrix} \Sigma & 0 & e & \bar{r} \\ 0 & 0 & 1 & r^c \\ e^* & 1 & & \\ \bar{r}^* & r^c & & \end{pmatrix} \begin{pmatrix} x \\ x^c \\ -\lambda \\ -\mu \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \\ \rho \end{pmatrix}.$$

The optimal budget multiplier  $\lambda$  is obtained from row 2. Substitution into row 1 yields the expression for  $x$ , and substituting  $x$  into row 3 yields  $x^c$ . Substitution of  $x$

and  $x^c$  into row 4 gives

$$\rho = \bar{r}^*x + r^cx^c = \mu(\gamma - r^c\beta) + r^c - \mu(r^c\beta - (r^c)^2\alpha) = r^c + \mu\delta^c,$$

yielding  $\mu$ . The positivity of  $\delta^c$  follows (with (A3)) from

$$\delta^c = (\bar{r} - r^ce)^*\Sigma^{-1}(\bar{r} - r^ce).$$

Finally, the second formula for  $x$  yields

$$\sigma^2(\rho) = \mu^2(\bar{r} - r^ce)^*\Sigma^{-1}(\bar{r} - r^ce) = \mu^2\delta^c = (\rho - r^c)^2/\delta^c.$$

The remaining statements ( $\hat{\rho} = r^c$ , etc.) follow trivially.  $\square$

PROBLEM 4. Problem 3 also has a tradeoff version:

$$\begin{aligned} \min_{x, x^c} \quad & \frac{1}{2}x^*\Sigma x - \mu(\bar{r}^*x + r^cx^c) \\ \text{s.t.} \quad & e^*x + x^c = 1. \end{aligned}$$

THEOREM 1.9. *Problem 3 with parameter  $\rho$  and Problem 4 with parameter  $\mu$  are equivalent if and only if  $\rho = r^c + \mu\delta^c$ .*

*Proof.* The proof is analogous to the proof of Theorem 1.6 and is therefore omitted.  $\square$

*Discussion.* Basically the situation is quite similar to Problem 2, the only qualitative difference being the existence of one zero-risk portfolio: for  $\rho = r^c$ , the capital is completely invested in cash and the risk vanishes. Otherwise a fraction of  $e^*x = \mu(\beta - r^c\alpha)$  is invested in risky assets and the risk is positive; see Figure 1.2. The optimal portfolio is now a mix of the (reward-independent) risky portfolio  $(\Sigma^{-1}(\bar{r} - r^ce), 0)$  and cash  $(0, 1)$ . The following comparison shows precisely how the cash account reduces risk when added to a set of (two or more) risky assets.

THEOREM 1.10. *The risk in Problem 3 is almost always lower than it is in Problem 2: If  $\beta \neq r^c\alpha$ , then the efficient frontiers touch in the single point*

$$\rho = r^c + \frac{\delta^c}{\beta - r^c\alpha} = \frac{\gamma - r^c\beta}{\beta - r^c\alpha}, \quad \sigma^2(\rho) = \frac{\delta^c}{(\beta - r^c\alpha)^2}$$

(see Figure 1.2), where the solutions of both problems are “identical”:  $x$  or  $(x, 0)$ . If  $\beta = r^c\alpha$ , then  $x^c \equiv 1$  and  $e^*x \equiv 0$ , and the risks differ by the constant  $1/\alpha$ :

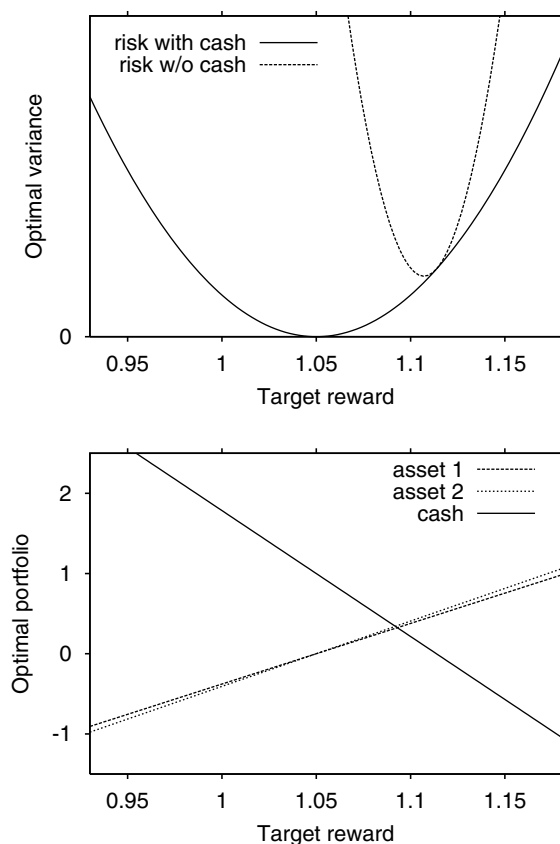
$$\frac{(\rho - r^c)^2}{\delta^c} + \frac{1}{\alpha} = \frac{\alpha\rho^2 - 2\beta\rho + \gamma}{\delta}.$$

*Proof.* If  $\beta \neq r^c\alpha$ , then Problem 3 has a unique zero-cash solution,  $x^c = 0$ , with

$$\mu = \frac{1}{\beta - r^c\alpha}, \quad \lambda = -\frac{r^c}{\beta - r^c\alpha}.$$

This gives the stated values of  $\rho$  and  $\sigma^2(\rho)$  by Theorem 1.8. Substituting  $\rho$  into the formulae for  $\lambda, \mu$  in Theorem 1.5 yields identical values in both problems. Hence the portfolios agree, too. The curvatures of the efficient frontiers,  $d^2\sigma^2(\rho)/d\rho^2$ , are  $2\alpha/\delta$  and  $2/\delta^c$ , respectively. Now,

$$\alpha\delta^c - \delta = (r^c)^2\alpha^2 - 2r^c\alpha\beta + \alpha\gamma - \alpha\gamma + \beta^2 = (r^c\alpha - \beta)^2 > 0.$$



**Fig. 1.2** Portfolio with two risky assets and cash, having expected returns  $\bar{r}^1 = 1.15$  and  $\bar{r}^2 = 1.08$  (as before), and  $r^c = 1.05$ . Top: Efficient frontiers with and without cash. Bottom: Optimal portfolio versus reward.

Thus  $2\alpha/\delta > 2/\delta^c > 0$ , implying that Problem 3 has lower risk if  $x^c \neq 0$ . The case  $\beta = r^c\alpha$  is trivial: both efficient frontiers have  $\hat{\rho} = r^c$  and identical curvatures.  $\square$

To conclude this section, we show that it does not make sense to consider portfolios with more than one riskless asset (and no further restrictions).

**LEMMA 1.11 (arbitrage).** *Any portfolio having at least two riskless assets  $x^c, x^d$  with different returns  $r^c, r^d$  can realize any desired reward at zero risk.*<sup>3</sup>

*Proof.* Choose  $x^c = (\rho - r^d)/(r^c - r^d)$ ,  $x^d = 1 - x^c$ , and invest nothing in other assets.  $\square$

**1.3. Risky Assets, Cash, and Guaranteed Total Loss.** Let us now consider a portfolio with  $n \geq 1$  risky assets, a riskless cash account as in Problem 3, and in addition an “asset”  $x^l$  with guaranteed total loss, i.e.,  $r^l \equiv \bar{r}^l = 0$ . (Notice that  $x^l$  is not “risky” in the sense of an uncertain future.) At first glance this situation seems strange, but it will turn out to be useful.<sup>4</sup>

<sup>3</sup>Here and in what follows, we use an abstract notion of arbitrage, meaning any opportunity to generate riskless profit. This differs from more specific standard definitions in finance.

<sup>4</sup>The suggestive notion of an “asset with guaranteed total loss” is perhaps the most obvious but least reasonable interpretation of  $x^l$ ; this is precisely what we wish to stress by using it.

BASIC ASSUMPTIONS. In addition to the conditions of the previous section we now require positive cash return ( $r^c \leq r^l$  does not make sense).

(A1)  $\Sigma > 0$ .

(A3)  $\bar{r} \neq r^c e$ .

(A4)  $r^c > 0$ .

PROBLEM 5. All covariances associated with  $x^c$  or  $x^l$  vanish, so that the risk and reward are  $R(x, x^c, x^l) = x^* \Sigma x$  and  $\rho(x, x^c, x^l) = \bar{r}^* x + r^c x^c$ , respectively, and the optimization problem reads

$$\begin{aligned} \min_{x, x^c, x^l} \quad & \frac{1}{2} \begin{pmatrix} x \\ x^c \\ x^l \end{pmatrix}^* \begin{pmatrix} \Sigma & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ x^c \\ x^l \end{pmatrix} = \frac{1}{2} x^* \Sigma x \\ \text{s.t.} \quad & e^* x + x^c + x^l = 1, \quad x^l \geq 0, \\ & \bar{r}^* x + r^c x^c = \rho. \end{aligned}$$

Note that the no-arbitrage condition  $x^l \geq 0$  must be imposed; otherwise one could borrow arbitrary amounts of money without having to repay. However, Lemma 1.11 still works for *sufficiently small*  $\rho$ . This is precisely our intention.

THEOREM 1.12. *Problem 5 has unique primal and dual solutions  $x, x^c, x^l, \lambda, \mu, \eta$ , where  $\eta$  is the multiplier of the nonnegativity constraint  $x^l \geq 0$ . For  $\rho > r^c$ , the optimal solution has  $x^l = 0$  and  $\eta = -\lambda > 0$  and is otherwise identical to the solution of Problem 3. Any reward  $\rho \leq r^c$  is obtained at zero risk by investing in a linear combination of the two riskless assets, with primal-dual solution*

$$x = 0, \quad x^c = \frac{\rho}{r^c}, \quad x^l = 1 - \frac{\rho}{r^c}, \quad \lambda = \mu = \eta = 0.$$

*Proof.* The system of necessary conditions can be written

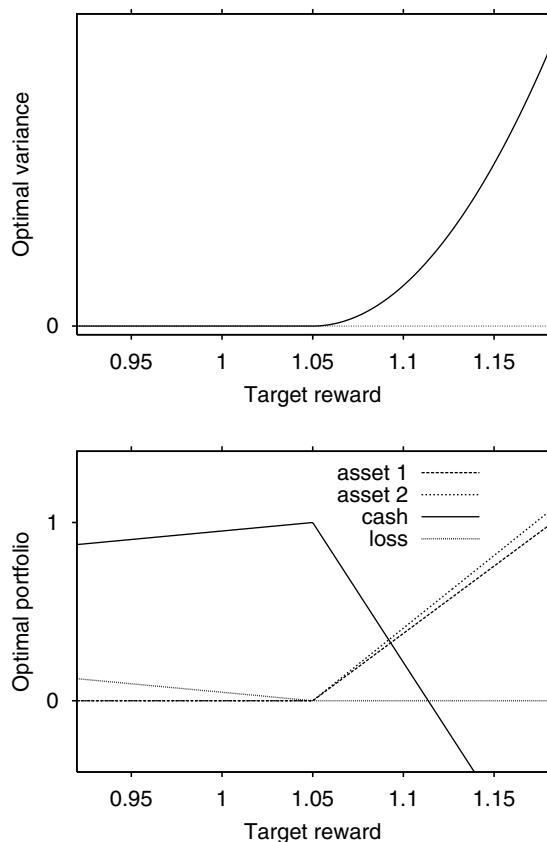
$$\begin{pmatrix} \Sigma & 0 & 0 & e & \bar{r} \\ 0 & 0 & 0 & 1 & r^c \\ 0 & 0 & 0 & 1 & 0 \\ e^* & 1 & 1 & & \\ \bar{r}^* & r^c & 0 & & \end{pmatrix} \begin{pmatrix} x \\ x^c \\ x^l \\ -\lambda \\ -\mu \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \eta \\ 1 \\ \rho \end{pmatrix}, \quad x^l \geq 0, \quad \eta \geq 0, \quad x^l \eta = 0.$$

As in Theorem 1.8, the first two rows yield  $\lambda = -r^c \mu$  and  $x = \mu \Sigma^{-1}(\bar{r} - r^c e)$ . The third row yields  $\eta = -\lambda = r^c \mu$ . Hence, by complementarity of  $x^l$  and  $\eta$ , rows 4 and 5 yield either  $x^c$  and  $\mu$  as in Problem 3 (if  $x^l = 0$  and  $\eta \geq 0$ ; case 1), or  $x^c + x^l = 1$  and  $r^c x^c = \rho$  (if  $x^l \geq 0$  and  $\eta = 0$ ; case 2). Due to the nonnegativity of  $x^l$  and  $\eta$ , case 1 can hold only for  $\rho \geq r^c$ , and case 2 only for  $\rho \leq r^c$ . (Indeed, for  $\rho = r^c$  both cases coincide so that all variables are continuous with respect to the parameter  $\rho$ .)  $\square$

PROBLEM 6. The tradeoff version of Problem 5 reads

$$\begin{aligned} \min_{x, x^c, x^l} \quad & \frac{1}{2} x^* \Sigma x - \mu(\bar{r}^* x + r^c x^c) \\ \text{s.t.} \quad & e^* x + x^c + x^l = 1, \quad x^l \geq 0. \end{aligned}$$

THEOREM 1.13. *Problem 6 with  $\mu > 0$  is equivalent to Problem 5 (with  $\rho > r^c$ ) iff  $\rho = r^c + \mu \delta^c$ . Every solution of Problem 5 with  $\rho \leq r^c$  is optimal for Problem 6 with  $\mu = 0$ . Problem 6 is unbounded for  $\mu < 0$ : no solution exists.*



**Fig. 1.3** Portfolio with two risky assets, cash, and an asset with guaranteed loss. Expected returns are  $\bar{r}^1 = 1.15$ ,  $\bar{r}^2 = 1.08$ , and  $r^c = 1.05$  (as before). Top: Efficient frontier. Bottom: Optimal portfolio versus reward.

*Proof.* The necessary conditions for both problems are identical, except that in the tradeoff problem  $\mu$  is given and the reward condition is missing. The condition  $\eta = r^c \mu$  together with nonnegativity and complementarity of  $x^l, \eta$  leads immediately to the three given cases.  $\square$

*Discussion.* Apparently, at the price of slightly increased complexity, Problem 5 correctly captures the case of an overly pessimistic investor. It minimizes something that qualitatively resembles a quadratic *downside risk* (or *shortfall risk*): the risk of obtaining less than the desired amount; see Figure 1.3. In that sense the model is now more realistic. (In contrast, its tradeoff version becomes degenerate for  $\mu = 0$  and does not extend to  $\mu < 0$ .) But what does it mean to “invest” knowingly in an asset with guaranteed total loss? Does it not imply that one might as well burn the money?

Let us first give the provocative answer, “Yes, why not?” From the point of view of the model, the investor’s goal is minimizing the “risk” of earning less *or* more than the specified reward. Therefore, it makes sense to get rid of money whenever this reduces the variance, which it indeed does for  $\rho < \hat{\rho}$ . The model cannot know and consequently does not care how the investor will interpret that, and it will use *any* possible means to take out capital if appropriate.

Of course, we can also offer a better interpretation. The fraction invested in  $x^l$  is simply surplus capital: the target reward  $\rho$  is achieved at zero risk without that amount, so it need not be invested in the first place—at least not into the portfolio under consideration. The investor may enjoy a free lunch instead or support her favorite artist, if she prefers that to burning the money. Or she may reconsider and decide to pursue a more ambitious goal; the model does not suggest *how* to spend the surplus money. This interpretation of the new riskless (but inefficient) solutions becomes obvious after the following observation.

LEMMA 1.14. *Problem 5 is equivalent to the modification of Problem 3, where the budget equation  $e^*x + x^c = 1$  is replaced by the inequality  $e^*x + x^c \leq 1$ , i.e., less than 100% investment is allowed.*

*Proof.* With a slack variable  $s \geq 0$ , the modified condition is clearly equivalent to  $e^*x + x^c + s = 1$ , and the modified Problem 3 becomes *identical* to Problem 5: the ominous loss asset is simply a slack variable,  $x^l \equiv s$ .  $\square$

**1.4. Utility Functions.** Let us start a brief excursion into utility-based portfolio optimization by considering Problem 1, the tradeoff formulation of the mean-variance model for  $n$  risky assets. In utility theory, the portfolio is chosen so that some function  $U(w)$ , the investor's (subjective) *utility* of final wealth  $w = r^*x$ , has maximal expectation for the given return distribution. The connection is apparent: minimizing the tradeoff function with parameter  $\mu_\rho$  is equivalent to maximizing the expectation  $\mathbf{E}[U_\rho(r^*x)]$  if we define the family of concave quadratic utility functions

$$U_\rho(w) := \mu_\rho w - \frac{1}{2}(w - \rho)^2, \quad \mu_\rho \equiv \frac{\alpha\rho - \beta}{\delta}.$$

( $\mu_\rho$  is the optimal budget multiplier of the target reward  $\rho$ .) If  $\rho + \mu_\rho > 0$ , then this equivalence remains valid for the normalized utility functions

$$\bar{U}_\rho(w) := \frac{1}{(\rho + \mu_\rho)^2} \left[ U_\rho(w) + \frac{1}{2}\rho^2 \right] = \frac{w}{\rho + \mu_\rho} - \frac{w^2}{2(\rho + \mu_\rho)^2},$$

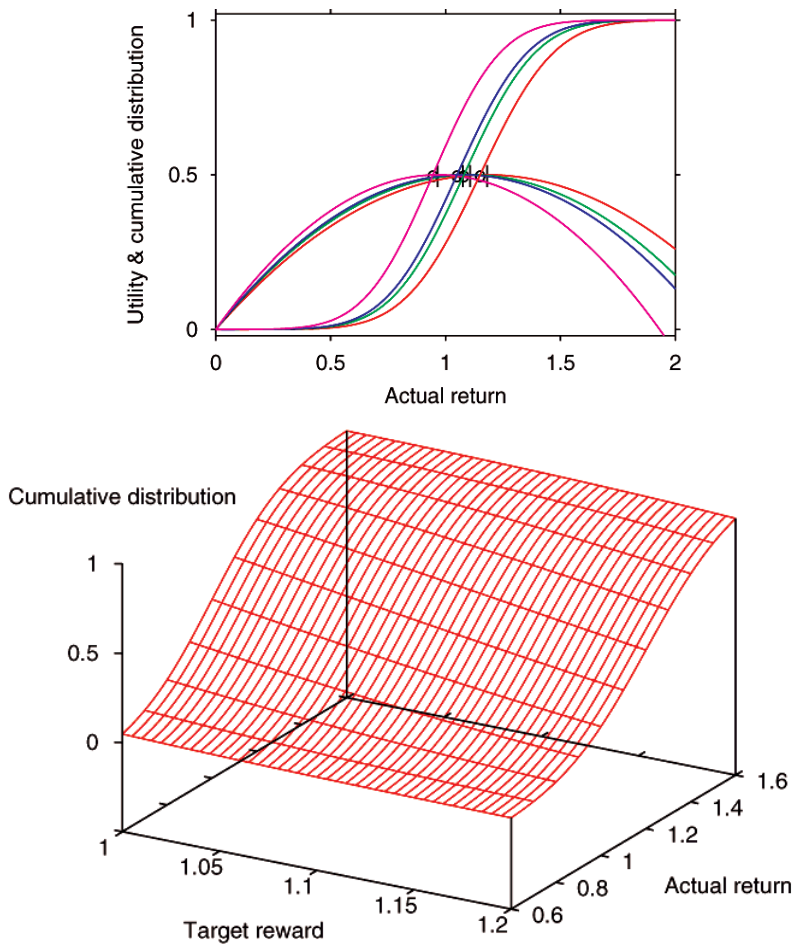
satisfying  $\bar{U}_\rho(0) = 0$  and  $\max_{w \in \mathbb{R}} \bar{U}_\rho(w) = \bar{U}_\rho(\rho + \mu_\rho) = \frac{1}{2}$ . For a portfolio with two positively correlated risky assets, Figure 1.4 shows the normalized utility functions associated with several target rewards, and the resulting optimal wealth distribution functions given normally distributed returns,  $r \sim \mathcal{N}(\bar{r}, \Sigma)$ . The optimal cumulative wealth distributions have the explicit form

$$\Phi_\rho(w) := \frac{1}{\sqrt{2\pi}\sigma_\rho} \int_{-\infty}^w \exp\left(-\frac{(t - \rho)^2}{2\sigma_\rho^2}\right) dt = \Phi\left(\frac{w - \rho}{\sqrt{2}\sigma_\rho}\right),$$

where  $\sigma_\rho^2 := \sigma^2(\rho) = (\mu_\rho^2\delta + 1)/\alpha$ , and  $\Phi$  is the standard error integral,

$$\Phi(w) := \frac{1}{\sqrt{\pi}} \int_{-\infty}^w \exp(-t^2) dt.$$

When cash is included in the portfolio (Problem 4), then utility functions  $U_\rho, \bar{U}_\rho$  and distribution functions  $\Phi_\rho$  have precisely the same form, except that the final wealth becomes  $w = r^*x + r^c x^c$  and that  $\mu_\rho := (\rho - r^c)/\delta^c$  and  $\sigma_\rho^2 := \mu_\rho^2\delta^c$ . The properties of  $\Phi_\rho$  give another indication of the risk reduction mechanism described in Theorem 1.10: Figure 1.5 shows that the slope of  $\Phi_\rho$  is rather steep and becomes a jump in the zero-risk case,  $\rho = r^c$ .



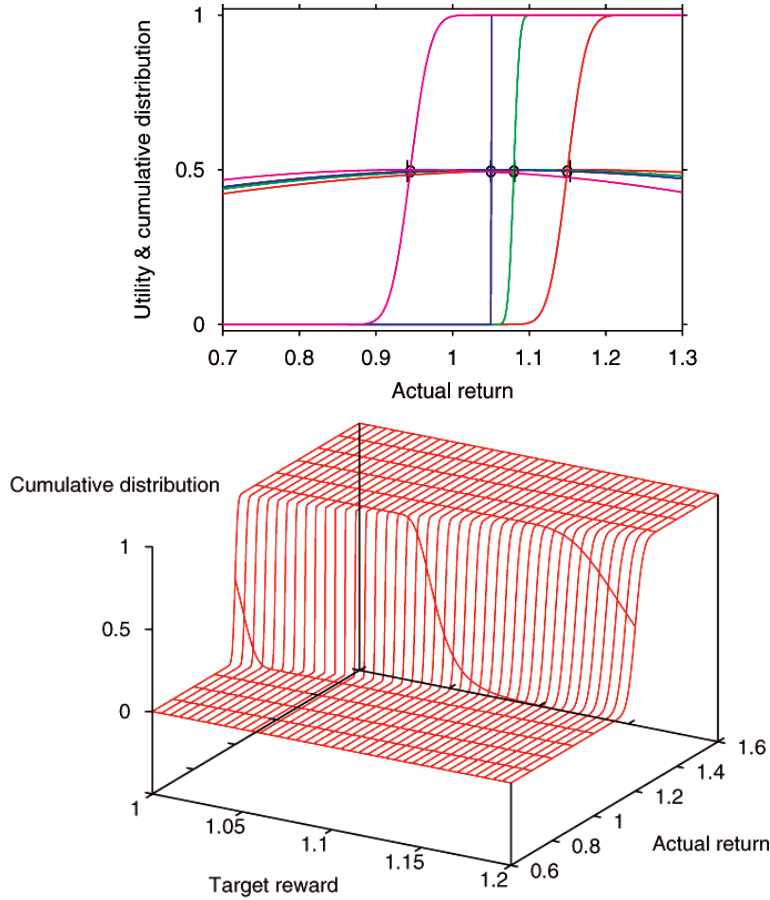
**Fig. 1.4** *Utility-based portfolio optimization for two positively correlated assets. Top: Normalized utility functions  $\bar{U}_\rho$  and optimal cumulative wealth distribution functions  $\Phi_\rho$  for the values  $\rho \in \{0.9r^c, r^c, \bar{r}^1, \bar{r}^2\}$ . Center of symmetry of  $\Phi_\rho$  curves at  $(\rho, \frac{1}{2})$  marked by 'o'; maximum of  $\bar{U}_\rho$  parabolas at  $(\rho + \mu_\rho, \frac{1}{2})$  marked by 'l'. Bottom: Family of optimal wealth distribution functions  $\Phi_\rho$  over the range of target rewards  $\rho \in [1.0, 1.2]$ .*

When a loss asset is also added to the portfolio in Problem 6, the utility functions  $U_\rho, \bar{U}_\rho$  are exactly identical to the previous case, and even their optimal wealth distributions for  $\rho \geq r^c$  coincide; see Figure 1.6. For  $\rho \leq r^c$ , however, the wealth distributions  $\Phi_\rho := \chi_{[\rho, \infty)}$  become indicator functions rather than normal distributions: they all have  $\mu_\rho = \sigma_\rho = 0$  and a jump discontinuity at  $\rho$ , producing zero risk. (In Problem 4 this happens only for  $\rho = r^c$ .)

In this paper we do not wish to pursue the subject further. The interested reader should refer to the original considerations of Markowitz [137, Part IV], the literature cited in the introduction (especially [92, 103, 105, 106, 119, 123, 150]), and the references therein.

**1.5. Influence of Inequalities.** Except for the no-arbitrage condition  $x^l \geq 0$ , all the problems considered so far have been purely equality constrained. Now we study





**Fig. 1.5** The same situation as in Figure 1.4 with cash included in the portfolio.

problems with inequalities. Let us first view  $\rho$  as a lower bound (not an exact value) for the desired reward.

**THEOREM 1.15.** Consider the following modifications of Problems 2, 3, and 5: the reward equation is replaced by the inequality  $\bar{r}^*x \geq \rho$  in Problem 2 and similarly by  $\bar{r}^*x + r^c x^c \geq \rho$  in Problems 3 and 5. Then the following hold:

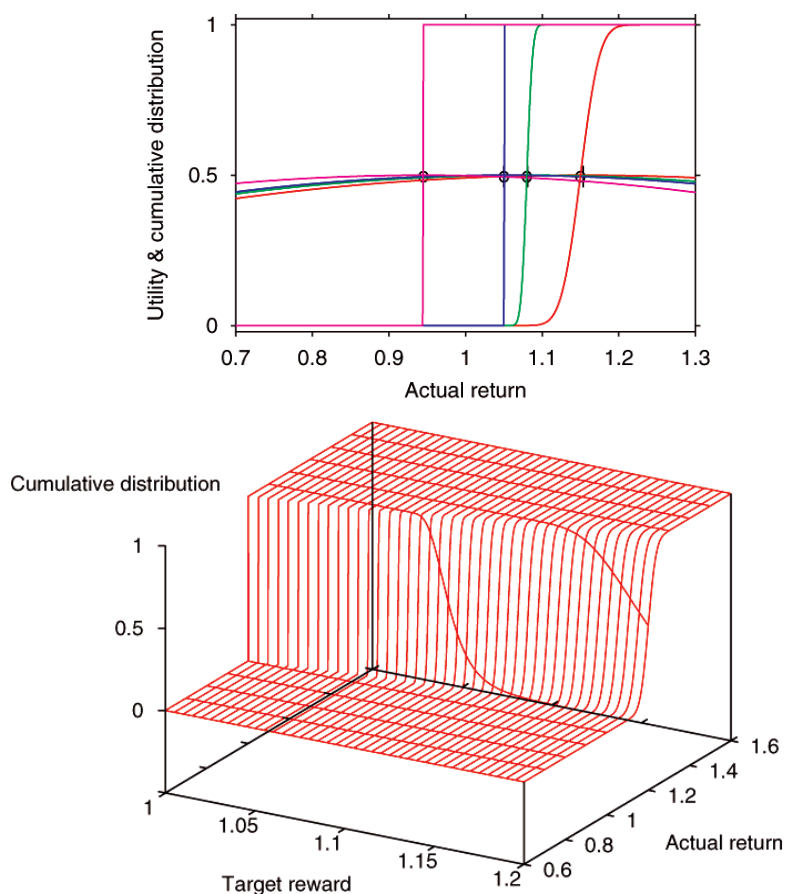
(1) The solution of each original problem for  $\rho \geq \hat{\rho}$  is also the unique solution of the corresponding modified problem (upper branch).

(2) The solution of Problem 2 or 3 with reward  $\hat{\rho}$  is also the unique solution of the corresponding modified problem for any  $\rho \leq \hat{\rho}$  (lower branch).

(3) Any solution of Problem 5 with  $\rho(x, x^c, x^l) \in [\rho, r^c]$  is a riskless solution of the modified problem with  $\rho < r^c$ . That is, any portfolio  $(0, x^c, 1 - x^c)$  with  $x^c \in [\rho/r^c, 1]$  is optimal.

*Proof.* The proof is obvious.  $\square$

*Discussion.* Specifying the desired reward as a lower bound rather than an exact value leads to reasonable behavior on the lower branch of the efficient frontier. The minimal-risk solution is simply extended to all sufficiently small rewards, yielding again a quadratic downside-like risk in each case. In Problem 3 this provides



**Fig. 1.6** The same situation as in Figure 1.4 with cash and loss included in the portfolio.

an alternative to introducing a loss asset. (By statement (3) of Theorem 1.15, the combination yields nonunique solutions but no further advantages.) The lower bound formulation, as originally introduced by Markowitz, might appear more natural than the loss asset, but mathematically both are equivalent: optimal investments in risky assets and the resulting risk are identical, but in the latter case the surplus money is put in  $x^l$  (removed immediately) and in the former case it is invested in cash  $x^c$ —to be removed afterwards. This difference could be interpreted as reflecting certain attitudes of the investor toward surplus money:  $e^*x + x^c \leq 1$  and  $\rho(x) \geq \rho$  would model respective preferences for immediate consumption or future profit, whereas the combination would express indecision. Instead, we simply interpret the loss model as giving the investor a *choice* between consumption and profit in some situations. Leaving the choice open seems preferable in view of the multiperiod case.

In practice, nonnegativity constraints  $x \geq 0$ ,  $x^c \geq 0$  will usually be included to prohibit short-selling assets or borrowing cash.<sup>5</sup> The budget equation then implies that only *convex combinations* of the assets are permitted. (That is, strictly speaking,

<sup>5</sup>Even if borrowing is allowed, it should also be modeled as a (separate) nonnegative asset in practice since the interest rate differs from the one for investing.

convex combinations of single-asset portfolios.) Of course, the unconstrained solution remains valid if and only if it is nonnegative anyway. Otherwise some constraints become tight, excluding the corresponding assets from the portfolio and increasing risk. More precisely, the following simple facts hold.

**THEOREM 1.16.** *Include nonnegativity constraints  $x \geq 0$ ,  $x^c \geq 0$  in Problems 2, 3, and 5 and assume for simplicity that  $\bar{r} > r^c e > 0$ . Denote by  $\bar{r}^{\min}, \bar{r}^{\max}$  the minimal and maximal expected return in the portfolio, and choose corresponding assets  $x^{\min}, x^{\max}$ . Then the following hold:*

- (1) *Problems 3 and 5 have  $x^{\min} = x^c$  and  $x^{\min} = x^l$ , respectively.*
- (2) *In each problem, an optimal solution exists iff  $\rho \in [\bar{r}^{\min}, \bar{r}^{\max}]$ .*
- (3) *The efficient frontier is convex and piecewise quadratic (or linear).*

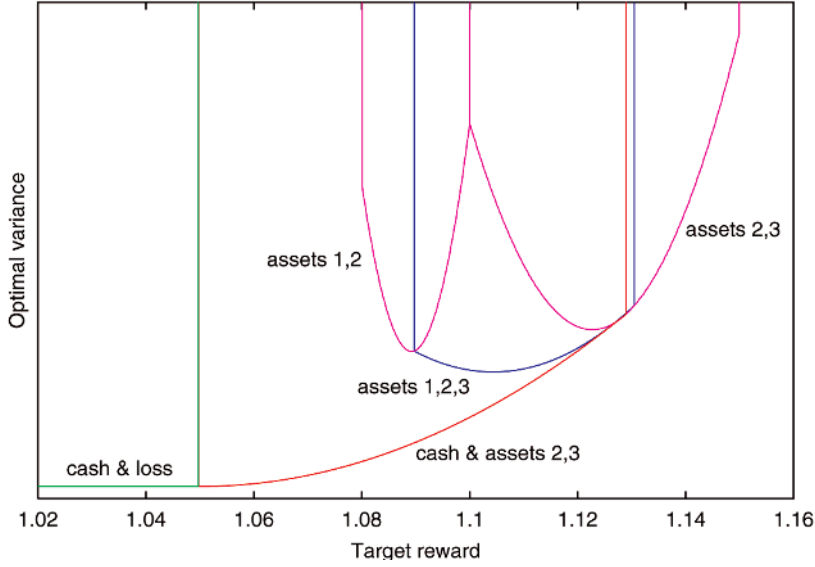
*Proof.* Statement (1) is trivial. Since each problem is convex, an optimal solution exists if and only if the feasible set is nonempty. For  $\rho \in [\bar{r}^{\min}, \bar{r}^{\max}]$ , the feasible set clearly contains a (unique) convex combination of  $x^{\min}$  and  $x^{\max}$ . Conversely, every convex combination of assets yields as reward the same convex combination of individual expected returns, which lies in the range  $[\bar{r}^{\min}, \bar{r}^{\max}]$ . This proves statement (2). To prove statement (3) consider Problem 2 first. Let  $\rho_0, \rho_1 \in [\bar{r}^{\min}, \bar{r}^{\max}]$  with respective solutions  $x_0, x_1$ . Then  $x_t := (1-t)x_0 + tx_1$  is feasible for  $\rho_t := (1-t)\rho_0 + t\rho_1$ ,  $t \in [0, 1]$ , and convexity of the efficient frontier follows from convexity of  $R$ ,

$$\sigma^2(\rho_t) \leq R(x_t) \leq (1-t)R(x_0) + tR(x_1) = (1-t)\sigma^2(\rho_0) + t\sigma^2(\rho_1).$$

At  $\bar{r}^{\min}$  and  $\bar{r}^{\max}$  all the money is invested in one single asset:  $x^{\min}$  or  $x^{\max}$ . Each  $\rho \in (\bar{r}^{\min}, \bar{r}^{\max})$  determines a subset of two or more nonnegative assets whose efficient frontier gives the optimal risk in that point. Since strict positivity is a generic property, each of these subportfolios is optimal either in a single point or on an entire nondegenerate interval. Thus, the efficient frontier (in Problem 2) is composed of finitely many quadratic pieces. Precisely the same arguments hold for Problem 3 since  $R(x, x^c) \equiv R(x)$ . In Problem 5, the efficient frontier consists of the segment  $\sigma^2(\rho) \equiv 0$  on  $[0, r^c]$  and the segments of Problem 3 on  $[r^c, \bar{r}^{\max}]$ .  $\square$

*Discussion.* The theorem gives a simple characterization of the influence of standard nonnegativity constraints. In a portfolio with three risky assets, the respective efficient frontiers of subportfolios that contribute to the optimal solution in Problems 2, 3, and 5 might look as in Figure 1.7. Other inequalities, like upper bounds on the assets or limits on arbitrary asset combinations, will further restrict the range of feasible rewards and increase the risk in a similar manner. This situation was already considered by Markowitz: he handles general linear inequalities by *dummy assets* (slacks) and constraints  $Ax = b$ ,  $x \geq 0$ ; the case  $A = e^*$  (with  $x \geq 0$  and  $\rho(x) \geq \rho$ ) is called the *standard case* [137, p. 171]. Moreover, Markowitz devised an algorithm to trace the *critical lines*, that is, the segments of the efficient frontier [136], [137, Chap. VIII]. Criteria for including or excluding assets in an optimal portfolio were developed by McEntire [140] and Kijima [102], and the number of assets in an optimal portfolio was investigated, e.g., by Nakasato and Furukawa [153].

**1.6. Downside Risk.** In the discussion of section 1.3 we made the remark that Problem 5 resembles a downside risk. This will now be investigated in detail. We have to work with the distribution of returns, but the entire analysis can be given in geometric terms using its *support* and elementary facts of convex analysis. Consider a probability space  $(\mathbb{R}^n, \mathcal{B}, P)$  and let  $\Xi \in \mathcal{B}$  denote the support of  $P$ , i.e., the smallest closed Borel set with measure 1,



**Fig. 1.7** Efficient frontiers for Problems 2, 3, and 5 with nonnegativity constraints for all assets. Efficient frontier for Problem 2 covers range  $[1.08, 1.15]$  with three quadratic segments (purple, blue, purple). Efficient frontier for Problem 3 covers range  $[1.05, 1.15]$  with three quadratic segments (red, blue, purple). Efficient frontier for Problem 5 covers range  $[0.00, 1.15]$  with one linear segment and three quadratic segments (green, red, blue, purple).

$$\Xi = \text{supp}(P) := \bigcap_{S \in \mathcal{B}: P(S)=1} \bar{S}.$$

(Of course, if  $P$  has a density  $\phi$ , then  $\Xi = \text{supp}(\phi) = \overline{\{x \in \mathbb{R}^n : \phi(x) > 0\}}$ .) In the following we will actually use the convex hull of the support most of the time, denoted by  $C := \text{conv}(\Xi)$ .

**DEFINITION 1.17** (downside risk). *For a function  $w$  of the random vector  $r$  with distribution  $P$ , the downside risk of order  $q > 0$  with target  $\tau \in \mathbb{R}$  is*

$$R_\tau^q(w) := \mathbf{E}[|\min(w(r) - \tau, 0)|^q] = \int_{\mathbb{R}^n} |\min(w(r) - \tau, 0)|^q dP.$$

*Remarks.* Without the risk context such expectations are neutrally called *lower partial moments*, with *downside expected value* or *semideviation* (order 1) and *downside variance* or *semivariance* (order 2) as special cases. In [137, Chap. XIII], Markowitz gives a qualitative discussion of the linear case (*expected value of loss*,  $q = 1$ ), the quadratic case (*semivariance*,  $q = 2$ ), and some other measures of risk, by examining the associated utility functions. Expectation of loss has recently gained interest as a *coherent* replacement for the popular *Value-at-Risk (VaR)*, often under alternative names like *mean shortfall*, *tail VaR*, or *conditional VaR*; cf. [3, 51, 66, 166, 200].

In the following we are only interested in quadratic downside risk of portfolio returns like  $w_{x,x^c}(r, r^c) = r^*x + r^c x^c$ . Moreover, we always use the target reward as a natural choice for the shortfall target,  $\tau = \rho$ , and write simply  $R_\rho(x, x^c)$  instead of  $R_\rho^2(w_{x,x^c})$ . The problems considered in this section are downside risk versions of Problems 3 and 5 and of the modification of Problem 3 with  $\rho(x, x^c) \geq \rho$ . In each

case only the objective is changed: standard risk  $R$  is replaced by downside risk  $R_\rho$ . Before considering these problems we need some technical preparations.

For  $x \neq 0$  and  $c \in \mathbb{R}$  let us introduce open and closed half-spaces

$$H(x, c) := \{r \in \mathbb{R}^n : r^*x < c\}, \quad \bar{H}(x, c) := \{r \in \mathbb{R}^n : r^*x \leq c\}$$

and portfolio-dependent semivariance matrices

$$\Sigma(x) := \int_{\bar{r}+H(x,0)} (r - \bar{r})(r - \bar{r})^* dP, \quad x \neq 0.$$

For  $x = 0$ , let  $\Sigma(0) := \frac{1}{2}\Sigma$ , where  $\Sigma$  is the usual covariance matrix,

$$\Sigma := \int_{\mathbb{R}^n} (r - \bar{r})(r - \bar{r})^* dP.$$

LEMMA 1.18. Denote by  $\uplus$  a disjoint union. Then, for  $x \neq 0$  and  $a > 0$ ,

- (1)  $H(ax, c) = H(x, a^{-1}c)$ ,  $H(x, ac) = H(a^{-1}x, c)$ ,  $H(ax, ac) = H(x, c)$ ;
- (2)  $H(-x, -c) = \mathbb{R}^n \setminus \bar{H}(x, c)$ ;
- (3)  $\bar{H}(x, c) = H(x, c) \uplus \partial H(x, c)$ ,  $\bar{H}(x, 0) = H(x, 0) \uplus \{x\}^\perp$ .

Statements (1) and (2) remain valid when  $H$  and  $\bar{H}$  are exchanged everywhere.

*Proof.* The proof is immediate from the definitions.  $\square$

LEMMA 1.19. For  $x \in \mathbb{R}^n$  and  $a > 0$ ,

- (1)  $\Sigma(ax) = \Sigma(x)$ ;
- (2)  $0 \leq \Sigma(x) \leq \Sigma$  (in particular, each  $\Sigma(x)$  is positive semidefinite);
- (3)  $x^*\Sigma(x)x = \mathbf{E}[\min((r - \bar{r})^*x, 0)^2]$ ;
- (4)  $x^*[\Sigma(x) + \Sigma(-x)]x = x^*\Sigma x$ .

*Proof.* Statements (1), (2), (3) are obvious from the definitions and the first identity in Lemma 1.18. The expressions in statement (4) are identical for  $x = 0$ ; otherwise they differ by the integral of  $((r - \bar{r})^*x)^2$  over  $\bar{r} + \{x\}^\perp$ , which is clearly zero.  $\square$

LEMMA 1.20. For any random vector  $r$  the following holds.

- (1) The expectation lies in the convex hull of the support:  $\bar{r} \in C$ .
- (2) The covariance matrix and all semivariance matrices are positive definite iff  $\Xi$  has full dimension in the sense that its convex hull has nonempty interior:

$$\text{int}(C) \neq \emptyset \iff \Sigma > 0 \iff \Sigma(x) > 0 \forall x \in \mathbb{R}^n.$$

- (3) If  $r$  is discrete with  $\Sigma > 0$ , then it has at least  $n + 1$  realizations.

*Proof.* Assume  $\bar{r} \notin C$ . Then  $\bar{r}$  has positive distance to  $C$ , and a vector  $x \neq 0$  exists so that  $(r - \bar{r})^*x > 0 \forall r \in C$ . Since expectation is the integral over  $\Xi \subseteq C$ , this yields the contradiction  $0 < \mathbf{E}[(r - \bar{r})^*x] = 0$ , proving statement (1). Now assume  $\text{int}(C) = \emptyset$ . Then  $C$  is contained in some hyperplane  $\bar{r} + \{x\}^\perp$  with  $x \neq 0$ , implying

$$x^*\Sigma x = \mathbf{E}[(r - \bar{r})^*x]^2 = 0.$$

Hence  $\Sigma$  is only positive semidefinite. Conversely, assume  $\text{int}(C) \neq \emptyset$  and  $x \neq 0$ . Then  $(r - \bar{r})^*x < 0 \forall r \in \bar{r} + H(x, 0)$ . By Lemma 1.21 below,  $\bar{r} + H(x, 0)$  has positive measure. Therefore

$$x^*\Sigma(x)x = \int_{\bar{r}+H(x,0)} ((r - \bar{r})^*x)^2 dP > 0,$$

showing that  $\Sigma(x) > 0$ . The proof of statement (2) is complete since  $\Sigma \geq \Sigma(x) \forall x$ . Now statement (3) is an immediate consequence.  $\square$

LEMMA 1.21. *Let  $\text{int}(C) \neq \emptyset$  and  $x \neq 0$ . Then  $\bar{r} + H(x, 0)$  has positive measure.*

*Proof.* The inner product  $s(x) := (r - \bar{r})^*x$  is negative, zero, and positive on the respective sets  $\bar{r} + H(x, 0)$ ,  $\bar{r} + \{x\}^\perp$ , and  $\bar{r} + H(-x, 0)$ . Furthermore,

$$\int_{\bar{r} + H(x, 0)} s(x) dP + \int_{\bar{r} + \{x\}^\perp} s(x) dP + \int_{\bar{r} + H(-x, 0)} s(x) dP = \mathbf{E}[s(x)] = 0.$$

Therefore,  $\bar{r} + H(x, 0)$  and  $\bar{r} + H(-x, 0)$  have either both positive measure or both measure zero. The second case implies  $\Xi \subseteq \bar{r} + \{x\}^\perp$ , which leads to the contradiction  $\text{int}(C) = \emptyset$ .  $\square$

Let us now study the downside risk versions of Problems 3 and 5 under the same assumptions as before ((A1) and (A3), respectively, (A1), (A3), and (A4)). It will be seen that in these cases the qualitative behavior does not change significantly. This is mainly because the constraints are linear and  $\Sigma(x)$  depends only on the direction and not on the magnitude of  $x$  (cf. Lemma 1.19).

PROBLEM 7. We minimize downside risk  $R_\rho(x, x^c)$  for risky assets and cash, with fixed target reward  $\rho(x, x^c) = \rho$ ,

$$\begin{aligned} \min_{x, x^c} \quad & \frac{1}{2} \int_{\mathbb{R}^n} \min(r^*x + r^c x^c - \rho, 0)^2 dP \\ \text{s.t.} \quad & e^*x + x^c = 1, \\ & \bar{r}^*x + r^c x^c = \rho. \end{aligned}$$

PROBLEM 8. Now minimize downside risk  $R_\rho(x, x^c, x^l)$  for risky assets, cash, and loss, with fixed target reward  $\rho(x, x^c, x^l) = \rho$ ,

$$\begin{aligned} \min_{x, x^c, x^l} \quad & \frac{1}{2} \int_{\mathbb{R}^n} \min(r^*x + r^c x^c - \rho, 0)^2 dP \\ \text{s.t.} \quad & e^*x + x^c + x^l = 1, \quad x^l \geq 0, \\ & \bar{r}^*x + r^c x^c = \rho. \end{aligned}$$

LEMMA 1.22. *With  $x^c \equiv 1 - e^*x - x^l$  and  $\theta \equiv r^c x^l$ , Problem 8 is equivalent to*

$$\begin{aligned} \min_{x, \theta} \quad & \frac{1}{2} x^* \Sigma(x) x \\ \text{s.t.} \quad & (\bar{r} - r^c e)^* x = \rho + \theta - r^c, \quad \theta \geq 0. \end{aligned}$$

*When fixing  $\theta = 0$ , the resulting problem is equivalent to Problem 7.*

*Proof.* The modified reward condition is immediately obtained by the identity  $x^c = 1 - e^*x - x^l$ . Using  $r^*x + r^c x^c - \rho = (r - \bar{r})^*x$  gives the downside risk

$$\int_{\mathbb{R}^n} \min((r - \bar{r})^*x, 0)^2 dP = \int_{\bar{r} + H(x, 0)} ((r - \bar{r})^*x)^2 dP = x^* \Sigma(x) x.$$

(The special case  $x = 0$  is easily verified.) Clearly,  $\theta = 0$  means  $x^l = 0$ , yielding Problem 7.  $\square$

We are now ready to analyze Problems 7 and 8. In general, closed-form solutions cannot be found due to the nonlinearity of downside risk with respect to the risky assets. However, we can derive some important properties of the solutions and give a qualitative comparison to Problems 3 and 5.

LEMMA 1.23. *Optimal solutions always exist in Problems 7 and 8. The resulting downside risk is nonnegative and not greater than the optimal risk in Problem 3 or 5, respectively. Moreover, the riskless solutions of Problems 7 and 3 (8 and 5) are identical. (In general the solutions are not unique.)*

*Proof.* Convexity of  $\min(w, 0)^2$  implies convexity of downside risk  $x^*\Sigma(x)x$  and thus of Problems 7 and 8. The existence of solutions and the stated inclusion follow since  $0 \leq \Sigma(x) \leq \Sigma$  by Lemma 1.18. By assumption (A1) and Lemma 1.20, zero risk requires  $x = 0$ , which holds under the same conditions as in the standard risk case:  $\rho = r^c$  in Problem 7 and  $\rho \leq r^c$  in Problem 8.  $\square$

THEOREM 1.24. *In Problem 7, choose respective optimal portfolios  $(x_{\pm}, x_{\pm}^c)$  for  $\rho_{\pm} := r^c \pm 1$ . Then, for  $a \geq 0$ ,  $(ax_{\pm}, ax_{\pm}^c - a + 1)$  is optimal for  $\rho = r^c \pm a$ . Moreover,  $x_{\pm} \neq 0$  and  $x_+ \neq x_-$ .*

*Proof.* If  $ax_+$  is optimal, the transformation of  $x_+^c$  follows from  $e^*x + x^c = 1$ . Suppose that  $ax_+$  is not optimal for  $\rho = r^c + a > r^c$ . Then, by Lemma 1.22,  $x \neq ax_+$  exists so that  $(\bar{r} - r^c e)^*x = \rho - r^c$  and

$$x^*\Sigma(x)x < ax_+^*\Sigma(ax_+)ax_+ = a^2x_+^*\Sigma(x_+)x_+,$$

where the last equality holds by Lemma 1.19. Hence, letting  $y = a^{-1}x$ ,

$$(\bar{r} - r^c e)^*y = a^{-1}(\rho - r^c) = 1 = \rho_+ - r^c$$

and

$$y^*\Sigma(y)y = a^{-2}x^*\Sigma(x)x < x_+^*\Sigma(x_+)x_+.$$

Thus  $x_+$  cannot be optimal for  $\rho_+$ , which is a contradiction. The case  $\rho < r^c$  is analogous, and  $\rho = r^c$  is trivial. Finally,  $(\bar{r} - r^c e)^*x_- < 0 < (\bar{r} - r^c e)^*x_+$  implies  $x_{\pm} \neq 0$  and  $x_+ \neq x_-$ .  $\square$

THEOREM 1.25. *Constants  $c_{\pm} \in (0, 1)$  exist so that the optimal risk in Problem 7 is  $c_+$  ( $c_-$ ) times the optimal risk of Problem 3 on the upper (lower) branch.*

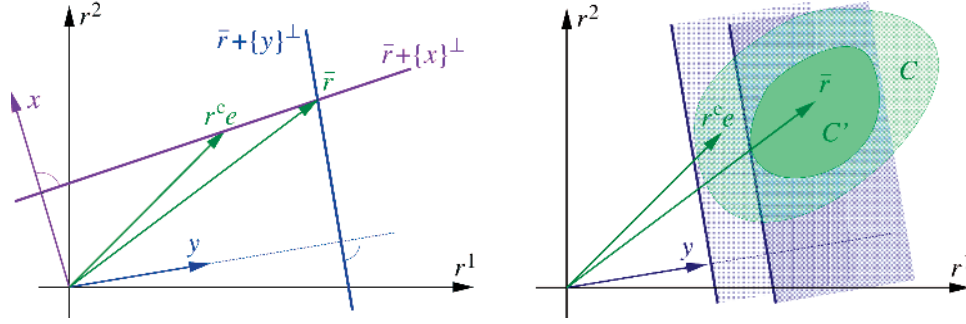
*Proof.* The existence of  $c_{\pm} \in (0, 1]$  with the stated properties follows from Lemma 1.23 and Theorem 1.24. Statement (4) of Lemma 1.19 implies  $c_{\pm} < 1$ .  $\square$

THEOREM 1.26. *The same statements as in Theorems 1.24 and 1.25 hold on the upper branch in Problem 8. On the lower branch one has the (unique) riskless solution  $(x, x^c, x^l) = (0, \rho/r^c, 1 - \rho/r^c)$ .*

*Proof.* This is a simple case distinction.  $\square$

*Remarks.* When assumption (A1) ( $\Sigma > 0$ ) is dropped, the following can be shown using Lemmas 1.18–1.22. A feasible portfolio with  $x \neq 0$  has zero risk in Problem 7 or 8 iff  $\Xi$  is contained in the hyperplane  $\bar{r} + \{x\}^{\perp}$ . In Problem 7 such a portfolio exists for  $\rho = r^c$  iff  $\Xi$  lies in a hyperplane containing both  $\bar{r}$  and  $r^c e$  (see  $x$  in Figure 1.8), and for  $\rho \neq r^c$  iff  $\Xi$  lies in a hyperplane containing  $\bar{r}$  but not  $r^c e$  (see  $y$  in Figure 1.8). Likewise, in Problem 8 such a portfolio exists for  $\rho \leq r^c$  iff  $\Xi$  lies in any hyperplane containing  $\bar{r}$  (see  $x, y$  in Figure 1.8), and for  $\rho > r^c$  iff  $\Xi$  lies in a hyperplane containing  $\bar{r}$  but not  $r^c e$  (see  $y$  in Figure 1.8). In both cases, an arbitrage is thus possible if  $\Xi$  lies in a hyperplane containing  $\bar{r}$  but not  $r^c e$ .

*Discussion.* Up to now, downside risk has behaved qualitatively similarly to standard risk: the efficient frontier is still piecewise quadratic, and optimal portfolios are always combinations of reward-independent portfolios. Only uniqueness is not guaranteed any more, and the curvatures of the upper and lower branches of the efficient frontier may differ. (Optimal portfolios will usually also differ from their standard risk



**Fig. 1.8** *Left: Zero-risk hyperplanes for singular  $\Sigma$  in Problems 7 and 8 ( $n = 2$ ). Downside risk vanishes if, depending on  $\rho$ , the blue or purple line contains the convex hull  $C$ . Right: Zero-risk half-spaces for Problem 9. Downside risk vanishes for  $\rho < r^c$  with convex hull  $C \supseteq \Xi$  and for any  $\rho \in \mathbb{R}$  with convex hull  $C' \supseteq \Xi$  not containing  $r^c e$ .*

counterparts, of course.) The similarity is caused by fixing the reward: this places the mean  $\bar{r}$  on the boundary of semivariance half-spaces so that the properties of  $\Sigma(x)$  come into play. The two risk measures become identical if the return distribution is symmetric with respect to rotations about  $\bar{r}$ . In that case  $\Sigma(x) = \frac{1}{2}\Sigma \forall x$ , and  $c_+ = c_- = \frac{1}{2}$ .

The last problem considered in this section is the downside risk version of the modification of Problem 3. Again the riskless solutions are of interest.

**PROBLEM 9.** We minimize downside risk  $R_\rho(x, x^c)$  for risky assets and cash, with desired minimal reward  $\rho(x, x^c) \geq \rho$ ,

$$\begin{aligned} \min_{x, x^c, \theta} \quad & \frac{1}{2} \int_{\mathbb{R}^n} \min(r^*x + r^c x^c - \rho, 0)^2 dP \\ \text{s.t.} \quad & e^*x + x^c = 1, \\ & \bar{r}^*x + r^c x^c = \rho + \theta, \quad \theta \geq 0. \end{aligned}$$

*Remark.* Note that downside risk is still calculated with respect to the *desired* reward  $\rho$ , whereas the *actual* reward is now  $\rho + \theta$ . Otherwise the problem would be equivalent to Problem 8.

**LEMMA 1.27.** *The pure cash portfolio  $(x, x^c) = (0, 1)$  is feasible for Problem 9 iff  $\rho \leq r^c$ . Otherwise, with  $x^c \equiv 1 - e^*x$ , Problem 9 is equivalent to*

$$\begin{aligned} \min_{x, \theta} \quad & \frac{1}{2} \int_{\bar{r} + H(x, -\theta)} (\theta + (r - \bar{r})^*x)^2 dP \\ \text{s.t.} \quad & (\bar{r} - r^c e)^*x = \rho + \theta - r^c, \quad \theta \geq 0. \end{aligned}$$

*Proof.* The first part is trivial. The second part is proved similar to Lemma 1.22, the sole difference being that the reward condition now yields  $r^*x + r^c x^c - \rho = (r - \bar{r})^*x + \theta$  instead of  $(r - \bar{r})^*x$ .  $\square$

**THEOREM 1.28.** *If  $r^c e \in \text{int}(C)$ , then the following holds in Problem 9.*

- (1) *For every  $\rho \leq r^c$ ,  $(x, x^c) = (0, 1)$  is a riskless solution.*
- (2) *A portfolio with  $x \neq 0$  has zero risk iff  $\Xi \subseteq r^c e + \bar{H}(-x, r^c - \rho)$ .*
- (3) *For  $\rho < r^c$  such a portfolio exists iff  $\Xi$  is contained in any closed half-space.*
- (4) *For  $\rho \geq r^c$  such a portfolio does not exist.*



*Proof.* Statement (1) is trivial. If  $x \neq 0$ , then downside risk clearly vanishes iff  $\Xi$  does not intersect  $\bar{r} + H(x, -\theta)$ , that is, iff  $(r - \bar{r})^*x + \theta \geq 0$  for  $r \in \Xi$ . Substituting  $\theta = (\bar{r} - r^c e)^*x - \rho + r^c$  from the reward equation yields the equivalent condition  $(r - r^c e)^*x \geq \rho - r^c$  for  $r \in \Xi$ , which proves statement (2). (This condition implies feasibility of  $x$  since  $\bar{r} \in C$ .) Now, if  $\Xi$  is contained in some closed half-space, then  $y \neq 0$  exists so that  $(r - r^c e)^*y \geq -1$  for  $r \in \Xi$ ; see Figure 1.8. For  $\rho < r^c$  let  $x := (r^c - \rho)y$  to satisfy the zero-risk condition. The “only if” direction of statement (3) is trivial. Finally observe that  $C$  contains an open ball centered at  $r^c e \in \text{int}(C)$ . On such a ball the inner product  $(r - r^c e)^*x$  takes positive and negative values for any  $x \neq 0$ , showing that for  $\rho \geq r^c$  the zero-risk condition  $(r - r^c e)^*x \geq \rho - r^c \geq 0$  cannot be satisfied.  $\square$

*Remarks.* Similar arguments show that for  $\rho > r^c$  ( $\rho = r^c$ ) a zero-risk portfolio with  $x \neq 0$  exists iff  $\Xi$  lies in a closed half-space not containing  $r^c e$  (not containing  $r^c e$  in its interior); see  $C'$  in Figure 1.8. This is why we need an additional no-arbitrage condition. Although  $r^c e \in C$  would suffice, we choose the stronger condition  $r^c e \in \text{int}(C)$  to ensure a unique riskless solution (100% cash) for  $\rho = r^c$ . Thus  $x \neq 0$  will produce positive risk for any  $\rho \geq r^c$ .

*Discussion.* Risk vanishes on the lower branch in Problem 9, but for the upper branch we know only that it is convex; even the optimal portfolios for different  $\rho > r^c$  may be unrelated. This is because the actual reward may exceed the shortfall target, resulting in semivariance half-spaces far from  $\bar{r}$  and producing asymmetric (or decentral) risk integrals instead of  $\Sigma(x)$ . One might expect such truly nonlinear behavior from any downside risk measure, but it occurs only if the actual reward may differ from the shortfall target. One might also associate zero risk on the lower branch with downside risk, but this property occurs for standard risk as well and has nothing to do with the objective; it is caused by either of the inequalities  $\rho(x) \geq \rho$  or  $e^*x + x^c \leq 1$ .

**1.7. Summary.** We have discussed various formulations of the classical mean-variance approach to obtain single-period models that give a qualitatively correct description of risk, particularly for unreasonably small target rewards. Positivity constraints and other inequalities have been studied, and downside risk models have been analyzed in detail. Thus we have clarified the effects and interaction of all components in the portfolio optimization problems. In what follows we use the results of this section in developing multiperiod models. The goal is to achieve an approximate minimization of downside risk, which turns out to be essential in the generalized situation: if the investor’s wealth happens to increase rapidly, surplus money in the sense of section 1.3 will appear in *intermediate* periods. This will increase the variance even if the entire capital is invested in cash from then on.

As in the previous section, elementary convex analysis plays a central role in the investigation. Tradeoff formulations or utility functions will not be considered any more since extra constraints provide higher modeling flexibility and facilitate a better understanding of subtle details.

**2. Multiperiod Mean-Variance Analysis.** For multiperiod mean-variance models we consider a planning horizon of  $T + 1$  periods (not necessarily of equal length) in discrete time  $t = 0, \dots, T + 1$ . The portfolio is allocated at  $t = 0$  and thereafter restructured at  $t = 1, \dots, T$ , before the investor obtains the reward after the final period, at time  $T + 1$ . The portfolios and return vectors are  $x_t, r_{t+1} \in \mathbb{R}^n$ ,  $t = 0, \dots, T$ , yielding asset capitals  $r_t^v x_{t-1}^v$  just before the decision at time  $t$ . Cash, its return, and loss assets (if present) are denoted by  $x_t^c, r_t^c$ , and  $x_t^l$ ; the wealth is

$w_t = r_t x_{t-1}$  (or  $r_t x_{t-1} + r_t^c x_{t-1}^c$ ). Cash returns  $r_t^c$  are assumed to be known a priori, whereas the evolution of asset returns is of course random. The decision at time  $t$  is made after observing the realizations of  $r_1, \dots, r_t$  but prior to observations of  $r_{t+1}, \dots, r_{T+1}$ , leading to a *nonanticipative* policy  $x = (x_0, \dots, x_T)$ .

Suppose that the distribution of returns until  $T$  is given by a *scenario tree*: each  $r_t$  has finitely many realizations  $r_j$  with probabilities  $p_j > 0$ ,  $j \in L_t$ , so that  $L_t$  forms a level set in the tree. The set of all nodes is  $V := \bigcup_{t=0}^T L_t$ , and the set of leaves, each representing a scenario, is  $L := L_T$ . We denote by  $0 \in L_0$  the root, by  $j \in L_t$  the current node (a partial scenario), by  $i \equiv \pi(j) \in L_{t-1}$  its parent node, and by  $S(j) \subseteq L_{t+1}$  the set of child nodes (successors). The return in the final period may be given by continuous distributions in each leaf. Thus  $r_t, x_t$  are random vectors on a discrete-continuous probability space that possesses a filtration generated by the tree. The conditional expectation  $\bar{r}_T := \mathbf{E}(r_{T+1}|L_T)$  and its covariance matrix

$$\Sigma_T := \mathbf{E}[(r_{T+1} - \bar{r}_T)(r_{T+1} - \bar{r}_T)^*|L_T] = \mathbf{E}(r_{T+1}r_{T+1}^*|L_T) - \bar{r}_T\bar{r}_T^*$$

define random variables on the same space, with realizations  $\bar{r}_j, \Sigma_j$  on  $L_T$ .

The discrete decision vector is denoted  $x = (x_j)_{j \in V}$ . As before, reward and risk are defined as mean and variance of the final wealth,  $w_{T+1}$ . In the absence of cash these definitions read

$$\rho(x) = \mathbf{E}(r_{T+1}^*x_T) = \mathbf{E}(\bar{r}_T^*x_T) = \sum_{j \in L} p_j \bar{r}_j^* x_j$$

and

$$R(x) = \sigma^2(r_{T+1}^*x_T) = \mathbf{E}[(r_{T+1}^*x_T - \rho(x))^2].$$

LEMMA 2.1 (Frauendorfer and Siede [65]). *The risk is given by*

$$R(x) = \mathbf{E}[x_T^*(\Sigma_T + \bar{r}_T\bar{r}_T^*)x_T] - \rho(x)^2 = \sum_{j \in L} p_j x_j^*(\Sigma_j + \bar{r}_j\bar{r}_j^*)x_j - \rho(x)^2.$$

*Proof.* By definition,

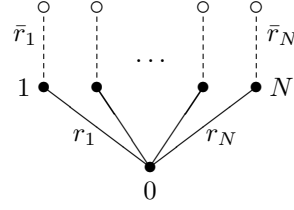
$$\begin{aligned} R(x) &= \mathbf{E}[(r_{T+1}^*x_T - \rho(x))^2] = \mathbf{E}(x_T^*r_{T+1}r_{T+1}^*x_T) - \rho(x)^2 \\ &= \mathbf{E}[\mathbf{E}(x_T^*r_{T+1}r_{T+1}^*x_T|L_T)] - \rho(x)^2 \\ &= \mathbf{E}[x_T^*\mathbf{E}(r_{T+1}r_{T+1}^*|L_T)x_T] - \rho(x)^2 = \mathbf{E}[x_T^*(\Sigma_T + \bar{r}_T\bar{r}_T^*)x_T] - \rho(x)^2. \end{aligned}$$

The discrete representation follows immediately.  $\square$

*Remarks.* Notice that this representation yields a block-diagonal risk matrix because of the separate term  $\rho(x)^2$ . If the Hessian of the latter were included in the risk matrix, it would add a completely dense rank-1 term: the dyadic product containing all the covariances  $-p_j p_k \bar{r}_j \bar{r}_k^*$  between different leaves  $j, k \in L$ . Since  $\rho(x) = \rho$  is fixed in the optimization problems below, we can neglect the term  $\rho^2$  except when considering the reward-dependence of the optimal risk.

COROLLARY 2.2. *Denote by  $\rho_T(x_T) := \bar{r}_T^*x_T$  and  $R_T(x_T) := x_T^*\Sigma_T x_T$  the conditional reward and risk of the final period, respectively, with realizations  $\rho_j(x_j) = \bar{r}_j^*x_j$  and  $R_j(x_j) = x_j^*\Sigma_j x_j$  on  $L_T$ . Then  $R(x) = R_c(x) + R_d(x)$  with*

$$R_c(x) := \mathbf{E}[R_T(x_T)] = \sum_{j \in L} p_j x_j^*\Sigma_j x_j$$



**Fig. 2.1** Scenario tree for the two-period mean-variance model.

and

$$R_d(x) := \mathbf{E}[\rho_T(x_T)^2] - \rho(x)^2 = \sum_{j \in L} p_j \rho_j(x_j)^2 - \rho(x)^2.$$

*Proof.* The proof is obvious from Lemma 2.1.  $\square$

*Remark.* We call  $R_c$  the *continuous* part and  $R_d$  the *discrete* part of the risk. This distinction will be useful in the subsequent analysis:  $R_c$  is the expectation of the conditional variance (of  $w_{T+1}$ ), measuring the average final-period risk, whereas  $R_d$  is the variance of the conditional expectation, measuring how well the individual scenario returns are balanced.

Apparently, period  $T + 1$  with its continuous distribution (of which only the conditional mean and variance enter the problem) corresponds to the single period in the classical model. Indeed, for  $T = 0$  the multiperiod model reduces precisely to the one-period case, where the “scenario tree” consists of the root only, and  $x \equiv x_T$ ,  $r \equiv r_{T+1}$ ,  $\bar{r} \equiv \bar{r}_T$ ,  $\Sigma \equiv \Sigma_T$ ,  $w \equiv w_{T+1}$ . Assuming idealized (frictionless) transactions with no loss of capital in the entire multiperiod situation, the single budget equation  $e^*x = 1$  is supplemented by the set  $\{e^*x_t = r_t^*x_{t-1}\}_{t=1}^T$  with discrete representation  $\{e^*x_j = r_j^*x_{\pi(j)}\}_{j \in V^*}$ , where  $V^* := V \setminus \{0\}$ .

The following analysis requires some lengthy technical proofs; these are moved into the appendix. Moreover, for technical simplicity, increasing amounts of the presentation will be specialized for the two-period problem,  $T = 1$ . In that case it is convenient to number the leaves as  $j = 1, \dots, N$ , so that  $V^* = S(0) = L = \{1, \dots, N\}$ ; see Figure 2.1. We would like to point out, however, that the two-period models exhibit almost all the properties of the general multiperiod case. Comments will be given at the end of each section.

**2.1. Risky Assets Only.** Let us begin the investigation of multiperiod problems with the case of purely risky assets. The model discussed here was originally proposed by Frauendorfer [63] (with slightly different objective) and later refined by Frauendorfer and Siede [65].

In this section we impose regularity conditions similar to (A1) and (A2) in all nodes and in at least one node, respectively. The conditions for  $L_{t-1}$  require certain definitions on  $L_t$  which in turn depend on the conditions for  $L_t$ . To avoid a nested presentation, the conditions will be stated after the definitions.

Let  $\tilde{r}_j := p_j \bar{r}_j$  and  $\tilde{\Sigma}_j := p_j(\Sigma_j + \bar{r}_j \bar{r}_j^*)$  in the leaves  $j \in L$ . By assumption (A5) below,  $\tilde{\Sigma}_j > 0$ ; therefore we can define

$$\tilde{\alpha}_j := e^* \tilde{\Sigma}_j^{-1} e, \quad \tilde{\beta}_j := e^* \tilde{\Sigma}_j^{-1} \tilde{r}_j, \quad \tilde{\gamma}_j := \tilde{r}_j^* \tilde{\Sigma}_j^{-1} \tilde{r}_j, \quad \tilde{\delta}_j := \tilde{\alpha}_j \tilde{\gamma}_j - \tilde{\beta}_j^2.$$

Recursively for  $t = T, \dots, 1$  and  $i \in L_{t-1}$  let

$$\tilde{r}_i := \sum_{j \in S(i)} \frac{\tilde{\beta}_j}{\tilde{\alpha}_j} r_j, \quad \tilde{\Sigma}_i := \sum_{j \in S(i)} \frac{1}{\tilde{\alpha}_j} r_j r_j^*$$

and employ (A5) again to define  $\tilde{\alpha}_i, \dots, \tilde{\delta}_i$  in analogy to  $\tilde{\alpha}_j, \dots, \tilde{\delta}_j$ . In the subsequent analysis (and in the solution algorithm) these quantities will play a similar role to their counterparts in the leaves, but they do not have the same meaning. In particular,  $\tilde{r}_i := \tilde{r}_i/p_i$  and  $\tilde{\Sigma}_i := \tilde{\Sigma}_i/p_i - \tilde{r}_i \tilde{r}_i^*$  are usually not the expectation and covariance matrices of the discrete distribution  $\{r_j\}_{j \in S(i)}$ .

BASIC ASSUMPTIONS.

(A5)  $\forall j \in V: \tilde{\Sigma}_j > 0$ .

(A6)  $\exists j \in V: \tilde{r}_j$  is not a multiple of  $e$ .

*Remarks.* The role of these conditions is analogous to the single-period case: they ensure strict convexity and avoid degenerate constraints. Assumption (A5) also implies  $N \geq n$  as a technical requirement on the return discretization in each period. In practice one will usually have  $N > n$ , otherwise the covariance matrices are only positive semidefinite by Lemma 1.20. Suitable multiperiod discretizations can be generated, e.g., by barycentric approximations [64] or by GARCH models [53, 20].

LEMMA 2.3. *Under assumptions (A5) and (A6), the constants  $\tilde{\alpha}_j, \tilde{\gamma}_j$  are all positive, the  $\tilde{\delta}_j$  are all nonnegative, and at least one  $\tilde{\delta}_j$  is positive.*

*Proof.* Positivity and nonnegativity are proved as in Lemma 1.3, where  $\tilde{\delta}_j = 0$  iff  $\tilde{r}_j, e$  are linearly dependent.  $\square$

PROBLEM 10. The multiperiod mean-variance problem (using  $i \equiv \pi(j)$ ) reads

$$\begin{aligned} \min_x \quad & \sum_{j \in L} \frac{1}{2} x_j^* \tilde{\Sigma}_j x_j - \frac{1}{2} \rho^2 \\ \text{s.t.} \quad & e^* x_0 = 1, \\ & e^* x_j = r_j^* x_i \quad \forall j \in V^*, \\ & \sum_{j \in L} \tilde{r}_j^* x_j = \rho. \end{aligned}$$

Its Lagrangian is

$$\begin{aligned} L(x, \lambda, \mu; \rho) = & \sum_{j \in L} \frac{1}{2} x_j^* \tilde{\Sigma}_j x_j - \frac{1}{2} \rho^2 \\ & - \lambda_0 (e^* x_0 - 1) - \sum_{j \in V^*} \lambda_j (e^* x_j - r_j^* x_i) - \mu \left( \sum_{j \in L} \tilde{r}_j^* x_j - \rho \right). \end{aligned}$$

THEOREM 2.4. *Problem 10 has the unique primal-dual solution*

$$x_j = \tilde{\Sigma}_j^{-1} (\lambda_j e + \mu \tilde{r}_j), \quad \lambda_j = \frac{w_j - \mu \tilde{\beta}_j}{\tilde{\alpha}_j}, \quad \mu = \left( \rho - \frac{\tilde{\beta}_0}{\tilde{\alpha}_0} \right) / \sum_{j \in V} \frac{\tilde{\delta}_j}{\tilde{\alpha}_j},$$

where  $w_0 = 1$  and  $w_j = r_j^* x_i$  for  $j \in V^*$ . The associated optimal risk is

$$R(x) \equiv \sigma^2(\rho) = \frac{1}{\tilde{\alpha}_0} + \left( \rho - \frac{\tilde{\beta}_0}{\tilde{\alpha}_0} \right)^2 / \sum_{j \in V} \frac{\tilde{\delta}_j}{\tilde{\alpha}_j} - \rho^2.$$

Its global minimizer  $\hat{\rho}$  and minimal risk are, respectively,

$$\hat{\rho} = \frac{\tilde{\beta}_0}{\tilde{\alpha}_0} \Big/ \left( 1 - \sum_{j \in V} \frac{\tilde{\delta}_j}{\tilde{\alpha}_j} \right), \quad \sigma^2(\hat{\rho}) = \frac{1}{\tilde{\alpha}_0} - \left( \frac{\tilde{\beta}_0}{\tilde{\alpha}_0} \right)^2 \Big/ \left( 1 - \sum_{j \in V} \frac{\tilde{\delta}_j}{\tilde{\alpha}_j} \right).$$

*Proof.* See the appendix (for  $T = 1$ ).  $\square$

*Discussion.* The proof of Theorem 2.4 is given for the two-period case only, where expressions for the solution variables are derived first in the leaves and then in the root. In the multiperiod case this generalizes readily to a recursive procedure which is actually a highly efficient algorithm for practical computations. We call that recursion the *tree-sparse Schur complement method* [193, 194].

The specialization to a single period gives similar results as in Theorems 1.4, 1.5, and 1.7, but under slightly weaker conditions: only  $\Sigma + \bar{r}\bar{r}^* > 0$  rather than  $\Sigma > 0$  is now required, so that riskless portfolios may exist. (We use the weaker condition since  $\tilde{\Sigma} = \Sigma + \bar{r}\bar{r}^*$  appears naturally in the problem.) The two following lemmas establish the precise relationship.

LEMMA 2.5. *If  $\tilde{\Sigma} > 0$  but not  $\Sigma > 0$ , then the null-space is  $N(\Sigma) = \text{span}(\tilde{\Sigma}^{-1}\bar{r})$ .*

*Proof.* If  $\Sigma x = 0$ , then  $\tilde{\Sigma}x = \bar{r}\bar{r}^*x$  and  $x = (\bar{r}^*x)\tilde{\Sigma}^{-1}\bar{r} \in \text{span}(\tilde{\Sigma}^{-1}\bar{r})$ .  $\square$

LEMMA 2.6. *Consider the single-period case of Problem 10 under assumptions (A1) and (A2), i.e.,  $\Sigma \equiv \Sigma_0 > 0$ . Then*

$$\tilde{\alpha}_0 = \frac{\alpha + \delta}{1 + \gamma}, \quad \tilde{\beta}_0 = \frac{\beta}{1 + \gamma}, \quad \tilde{\gamma}_0 = \frac{\gamma}{1 + \gamma}, \quad \tilde{\delta}_0 = \frac{\delta}{1 + \gamma},$$

and

$$\tilde{x} = x, \quad \tilde{\lambda} = \lambda, \quad \tilde{\mu} = \mu + \rho,$$

where quantities with and without tildes refer to Problems 10 and 2, respectively.

*Proof.* The Sherman–Morrison–Woodbury formula [187, 204] applied to  $\tilde{\Sigma} = \Sigma + \bar{r}\bar{r}^*$  yields

$$\tilde{\Sigma}^{-1} = \Sigma^{-1} - \frac{\Sigma^{-1}\bar{r}\bar{r}^*\Sigma^{-1}}{1 + \bar{r}^*\Sigma^{-1}\bar{r}} = \Sigma^{-1} - \frac{\Sigma^{-1}\bar{r}\bar{r}^*\Sigma^{-1}}{1 + \gamma}.$$

Using this in the definitions of  $\tilde{\alpha}_0, \tilde{\beta}_0, \tilde{\gamma}_0, \tilde{\delta}_0$  gives the first set of identities after a few elementary calculations. The relation between the solutions is similarly obtained by simple but more lengthy calculations (first  $\tilde{\mu}$ , then  $\tilde{\lambda}$ , then  $\tilde{x}$ ).  $\square$

**2.2. Risky Assets and Cash.** Here we define local constants for the two-period case only but state the general optimization problem. Denote by  $r_1^c, r_2^c$  the deterministic (and thus scenario-independent) cash returns in periods 1 and 2, respectively, and by  $r^c := r_2^c r_1^c$  their combined return. As before, let  $\tilde{r}_j = p_j \bar{r}_j$ ,  $\tilde{\Sigma}_j = p_j (\Sigma_j + \bar{r}_j \bar{r}_j^*)$  for  $j \in L$ , and define in addition  $\tilde{r}_j^c := p_j r_2^c$ . With assumption (A7) below let

$$\alpha_j := e^* \Sigma_j^{-1} e, \quad \beta_j := e^* \Sigma_j^{-1} \bar{r}_j, \quad \gamma_j := \bar{r}_j^* \Sigma_j^{-1} \bar{r}_j,$$

and

$$\delta_j^c := (r_2^c)^2 \alpha_j - 2r_2^c \beta_j + \gamma_j = (\bar{r}_j - r_2^c e)^* \Sigma_j^{-1} (\bar{r}_j - r_2^c e).$$

In the root define

$$\tilde{p}_0 := \sum_{j \in S(0)} \frac{p_j}{\delta_j^c + 1}, \quad \tilde{r}_0 := \sum_{j \in S(0)} \frac{p_j}{\delta_j^c + 1} r_j, \quad \tilde{\Sigma}_0 := \sum_{j \in S(0)} \frac{p_j}{\delta_j^c + 1} r_j r_j^*,$$

and furthermore  $\bar{r}_0 = \tilde{r}_0/\tilde{p}_0$ ,  $\Sigma_0 = \tilde{\Sigma}_0/\tilde{p}_0 - \bar{r}_0\bar{r}_0^*$ , and  $\tilde{r}_1^c := \tilde{p}_0 r^c$  (not  $\tilde{p}_0 r_1^c$ ). Using (A7) again,  $\alpha_0, \beta_0, \gamma_0$  are then defined in analogy to  $\alpha_j, \beta_j, \gamma_j$ . Finally, let

$$\delta_0^c := (r_1^c)^2 \alpha_0 - 2r_1^c \beta_0 + \gamma_0$$

and recall that  $t$  is the current time in  $j \in V$ .

BASIC ASSUMPTIONS. For the general multiperiod case we make the following assumptions.

(A7)  $\forall j \in V: \Sigma_j > 0$ .

(A8)  $\exists j \in V: \bar{r}_j \neq r_{t+1}^c e$ .

(A9)  $r^c \neq 0$ .

*Remark.* The conditions here are similar to (A5) and (A6) (or (A1) and (A3)), but in addition we require nonzero cash returns  $r_{t+1}^c$ ,  $t = 0, \dots, T$ . The opposite case would unnecessarily complicate the analysis and is not considered.

LEMMA 2.7. *Under assumptions (A7)–(A9), the constants  $\alpha_j, \gamma_j$  are all positive, the  $\delta_j^c$  are nonnegative, and at least one  $\delta_j^c$  is positive. Moreover,  $\tilde{p}_0 \in (0, 1]$ .*

*Proof.* Positivity and nonnegativity are obvious, where  $\delta_j^c = 0$  iff  $\bar{r}_j = r_{t+1}^c e$ . Now  $\delta_j^c \geq 0$  implies  $p_j/(\delta_j^c + 1) \in (0, p_j]$  and hence  $\tilde{p}_0 \in (0, 1]$ .  $\square$

In the following we use two different formulations for both the reward and the risk, involving again the conditional final-period risk and return. The latter is now  $\rho_T(x_T, x_T^c) := \bar{r}_T^* x_T + r_{T+1}^c x_T^c$ , with realizations  $\rho_j(x_j, x_j^c) := \bar{r}_j^* x_j + r_{T+1}^c x_j^c$ , and the discrete decision vector is  $x = (x_j, x_j^c)_{j \in V}$ .

LEMMA 2.8. *The reward in the presence of cash can be written*

$$\rho(x) = \sum_{j \in L} \bar{r}_j^* x_j + \tilde{r}_j^c x_j^c = \sum_{j \in L} p_j \rho_j(x_j, x_j^c).$$

*The risk has the two representations*

$$\begin{aligned} R(x) &= \sum_{j \in L} \begin{pmatrix} x_j \\ x_j^c \end{pmatrix}^* \begin{pmatrix} \tilde{\Sigma}_j & r_{T+1}^c \tilde{r}_j \\ r_{T+1}^c \tilde{r}_j^* & r_{T+1}^c \tilde{r}_j^c \end{pmatrix} \begin{pmatrix} x_j \\ x_j^c \end{pmatrix} - \rho(x)^2 \\ &= \sum_{j \in L} p_j x_j^* \Sigma_j x_j + \sum_{j \in L} p_j \rho_j(x_j, x_j^c)^2 - \rho(x)^2 =: R_c(x) + R_d(x). \end{aligned}$$

*Proof.* By definition, the continuous representation of the reward reads

$$\rho(x) = \mathbf{E}[r_{T+1}^* x_T + r_{T+1}^c x_T^c] = \mathbf{E}[\bar{r}_T^* x_T + r_{T+1}^c x_T^c] = \mathbf{E}[\rho_T(x_T, x_T^c)].$$

Likewise, using Lemma 2.1,

$$\begin{aligned} R(x) &= \mathbf{E} \left\{ \begin{pmatrix} x_T \\ x_T^c \end{pmatrix}^* \left[ \begin{pmatrix} \Sigma_T & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} r_{T+1} \\ r_{T+1}^c \end{pmatrix} \begin{pmatrix} r_{T+1} \\ r_{T+1}^c \end{pmatrix}^* \right] \begin{pmatrix} x_T \\ x_T^c \end{pmatrix} \right\} - \rho(x)^2 \\ &= \mathbf{E}(x_T^* \Sigma_T x_T) + \mathbf{E}[(\bar{r}_T^* x_T + r_{T+1}^c x_T^c)^2] - \rho(x)^2. \end{aligned}$$

In both cases, the stated discrete formulae are readily obtained.  $\square$

PROBLEM 11. Using the first formulation for reward and risk in Lemma 2.8, the multiperiod mean-variance problem with cash reads

$$\begin{aligned}
 \min_x \quad & \sum_{j \in L} \frac{1}{2} \begin{pmatrix} x_j \\ x_j^c \end{pmatrix}^* \begin{pmatrix} \tilde{\Sigma}_j & r_{T+1}^c \tilde{r}_j \\ r_{T+1}^c \tilde{r}_j^* & r_{T+1}^c \tilde{r}_j^c \end{pmatrix} \begin{pmatrix} x_j \\ x_j^c \end{pmatrix} - \frac{1}{2} \rho^2 \\
 \text{s.t.} \quad & e^* x_0 + x_0^c = 1, \\
 & e^* x_j + x_j^c = r_j^* x_i + r_t^c x_i^c \quad \forall j \in V^*, \\
 & \sum_{j \in L} \tilde{r}_j^* x_j + \tilde{r}_j^c x_j^c = \rho.
 \end{aligned}$$

Its Lagrangian is

$$\begin{aligned}
 L(x, \lambda, \mu; \rho) = & \sum_{j \in L} \frac{1}{2} \begin{pmatrix} x_j \\ x_j^c \end{pmatrix}^* \begin{pmatrix} \tilde{\Sigma}_j & r_{T+1}^c \tilde{r}_j \\ r_{T+1}^c \tilde{r}_j^* & r_{T+1}^c \tilde{r}_j^c \end{pmatrix} \begin{pmatrix} x_j \\ x_j^c \end{pmatrix} - \frac{1}{2} \rho^2 \\
 & - \lambda_0 (e^* x_0 + x_0^c - 1) - \sum_{j \in V^*} \lambda_j (e^* x_j + x_j^c - r_j^* x_i - r_t^c x_i^c) \\
 & - \mu \left( \sum_{j \in L} \tilde{r}_j^* x_j + \tilde{r}_j^c x_j^c - \rho \right).
 \end{aligned}$$

**THEOREM 2.9.** *Problem 11 with  $T = 1$  has the unique primal-dual solution*

$$\begin{aligned}
 x_j &= -\frac{\lambda_j}{\tilde{r}_j^c} \Sigma_j^{-1} (\bar{r}_j - r_2^c e), \quad j \in V^*, & x_0 &= -\frac{\lambda_0}{r_2^c \tilde{r}_1^c} \Sigma_0^{-1} (\bar{r}_0 - r_1^c e), \\
 x_j^c &= \frac{1}{r_2^c} \left[ \frac{\lambda_j}{\tilde{r}_j^c} (\gamma_j - r_2^c \beta_j + 1) + \mu \right], & x_0^c &= \frac{1}{r^c} \left[ \frac{\lambda_0}{\tilde{r}_1^c} (\gamma_0 - r_1^c \beta_0 + 1) + \mu \right], \\
 \lambda_j &= \frac{\tilde{r}_j^c}{\delta_j^c + 1} [r_2^c (r_j^* x_0 + r_1^c x_0^c) - \mu], & \lambda_0 &= \frac{\tilde{r}_1^c}{\delta_0^c + 1} (r^c - \mu) \equiv \tilde{\rho} (r^c - \mu), \\
 \mu &= r^c \frac{\rho - \tilde{\rho}}{r^c - \tilde{\rho}}, & \tilde{\rho} &:= \frac{\tilde{r}_1^c}{\delta_0^c + 1} \in (0, r^c).
 \end{aligned}$$

The associated risk is

$$R(x) \equiv \sigma^2(\rho) = \tilde{\rho} \frac{(r^c - \rho)^2}{r^c - \tilde{\rho}}.$$

Its global minimum is attained at  $\hat{\rho} = r^c$  and has value zero. The associated solution has 100% cash:  $(\hat{x}_0, \hat{x}_0^c) = (0, 1)$ ,  $(\hat{x}_j, \hat{x}_j^c) = (0, r_1^c)$ ,  $\hat{\lambda} = 0$ ,  $\hat{\mu} = r^c$ .

*Proof.* See the appendix.  $\square$

*Remark.* Problem 11 also has a unique solution if  $r_2^c \neq 0$  and  $r_1^c = 0$ , and it has multiple solutions for  $r_2^c = 0$  (regardless of  $r_1^c$ ). These situations are qualitatively different and quite unrealistic, however, and therefore not of interest here.

*Discussion.* Zero risk is now possible (with  $(x_j, x_j^c) = (0, r_1^c)$  for  $j \in L$ ) since balanced scenario returns do not require investments in risky assets. In fact, they require 100% cash in both periods so that the riskless solution is unique.

It can be seen that the whole situation is actually covered by the results of the previous section if in each node one replaces  $x_j$  by  $(x_j, x_j^c)$ , and so on. This way the analysis extends again to the general multiperiod case. The details are obtained precisely as in the previous section when proper replacements are carried out everywhere.

However, the splitting into cash and risky assets makes the definitions of intermediate quantities and the specialization to the single-period case somewhat more involved.

To conclude this section we show that specifying  $\rho$  as a desired *minimal* reward has precisely the same effect as in the single-period case.

PROBLEM 12. We modify Problem 11 by requiring  $\rho$  to be a lower bound on the reward (with associated slack  $\theta \geq 0$  and dual slack  $\eta \geq 0$ ),

$$\begin{aligned} \min_{x, \theta} \quad & \sum_{j \in L} \frac{1}{2} \begin{pmatrix} x_j \\ x_j^c \end{pmatrix}^* \begin{pmatrix} \tilde{\Sigma}_j & r_{T+1}^c \tilde{r}_j \\ r_{T+1}^c \tilde{r}_j^* & r_{T+1}^c \tilde{r}_j^c \end{pmatrix} \begin{pmatrix} x_j \\ x_j^c \end{pmatrix} - \frac{1}{2}(\rho + \theta)^2 \\ \text{s.t.} \quad & e^* x_0 + x_0^c = 1, \\ & e^* x_j + x_j^c = r_j^* x_i + r_i^c x_i^c \quad \forall j \in V^*, \\ & \sum_{j \in L} \tilde{r}_j^* x_j + \tilde{r}_j^c x_j^c = \rho + \theta, \quad \theta \geq 0. \end{aligned}$$

The Lagrangian is now

$$\begin{aligned} L(x, \theta, \lambda, \mu, \eta; \rho) = & \sum_{j \in L} \frac{1}{2} \begin{pmatrix} x_j \\ x_j^c \end{pmatrix}^* \begin{pmatrix} \tilde{\Sigma}_j & r_{T+1}^c \tilde{r}_j \\ r_{T+1}^c \tilde{r}_j^* & r_{T+1}^c \tilde{r}_j^c \end{pmatrix} \begin{pmatrix} x_j \\ x_j^c \end{pmatrix} - \frac{1}{2}(\rho + \theta)^2 \\ & - \lambda_0(e^* x_0 + x_0^c - 1) - \sum_{j \in V^*} \lambda_j(e^* x_j + x_j^c - r_j^* x_i - r_i^c x_i^c) \\ & - \mu \left( \sum_{j \in L} \tilde{r}_j^* x_j + \tilde{r}_j^c x_j^c - \rho - \theta \right) - \eta \theta. \end{aligned}$$

THEOREM 2.10. *Problem 12 has a unique primal-dual solution. For  $\rho \geq r^c$  one obtains  $\theta = 0$ ,  $\eta = \mu - \rho \geq 0$ , and otherwise the same solution as in Problem 11. For  $\rho \leq r^c$  one obtains  $\eta = 0$  and  $\theta = \rho - r^c \geq 0$ , giving reward  $\rho + \theta = r^c$  and the associated riskless solution of Problem 11. (At  $\rho = r^c$  both cases coincide.)*

*Proof.* See the appendix.  $\square$

**2.3. Risky Assets, Cash, and Loss.** We are now entering the main section. Although it will still be a simplification, Problem 13 below covers all the essential aspects of the multiperiod application model mentioned in the introduction. The notation and constants of the previous section remain valid, and the discrete decision vector is  $x = (x_j, x_j^c, x_j^l)_{j \in V}$ . We keep the general multiperiod notation only in the problem statement and Lagrangian; the remaining analysis now concentrates on the two-period case.

BASIC ASSUMPTIONS. In addition to the assumptions of the previous section we require positive cash returns and (as in Theorem 1.28) a no-arbitrage condition on the discrete part of the return distribution, involving  $C_0 := \text{conv}(\{r_j\}_{j \in L})$ .

(A7)  $\forall j \in V: \Sigma_j > 0$ .

(A8)  $\exists j \in V: \tilde{r}_j \neq r_{t+1}^c e$ .

(A10)  $r_1^c > 0, r_2^c > 0$ .

(A11)  $r_1^c e \in \text{int}(C_0)$ .



PROBLEM 13. The multiperiod mean-variance problem with cash and loss reads

$$\begin{aligned} \min_x \quad & \sum_{j \in L} \frac{1}{2} \begin{pmatrix} x_j \\ x_j^c \end{pmatrix}^* \begin{pmatrix} \tilde{\Sigma}_j & r_{T+1}^c \tilde{r}_j \\ r_{T+1}^c \tilde{r}_j^* & r_{T+1}^c \tilde{r}_j^c \end{pmatrix} \begin{pmatrix} x_j \\ x_j^c \end{pmatrix} - \frac{1}{2} \rho^2 \\ \text{s.t.} \quad & e^* x_0 + x_0^c + x_0^l = 1, \quad x_0^l \geq 0, \\ & e^* x_j + x_j^c + x_j^l = r_j^* x_i + r_t^c x_i^c, \quad x_j^l \geq 0 \quad \forall j \in V^*, \\ & \sum_{j \in L} \tilde{r}_j^* x_j + \tilde{r}_j^c x_j^c = \rho. \end{aligned}$$

Its Lagrangian is

$$\begin{aligned} L(x, \lambda, \eta, \mu; \rho) = & \sum_{j \in L} \frac{1}{2} \begin{pmatrix} x_j \\ x_j^c \end{pmatrix}^* \begin{pmatrix} \tilde{\Sigma}_j & r_{T+1}^c \tilde{r}_j \\ r_{T+1}^c \tilde{r}_j^* & r_{T+1}^c \tilde{r}_j^c \end{pmatrix} \begin{pmatrix} x_j \\ x_j^c \end{pmatrix} - \frac{1}{2} \rho^2 \\ & - \lambda_0 (e^* x_0 + x_0^c + x_0^l - 1) - \sum_{j \in V^*} \lambda_j (e^* x_j + x_j^c + x_j^l - r_j^* x_i - r_t^c x_i^c) \\ & - \sum_{j \in V} \eta_j x_j^l - \mu \left( \sum_{j \in L} \tilde{r}_j^* x_j + \tilde{r}_j^c x_j^c - \rho \right). \end{aligned}$$

Theorem A.1 in the appendix characterizes optimal solutions in a way similar to the previous theorems. However, the loss variables  $x_j^l$  still appear in the formulae, and the case distinctions are more involved than in the single-period case (cf. Theorem 1.12) or in Problem 12. Later we will discuss this in part; for the time being, the following results provide more insight.

LEMMA 2.11 (arbitrage in Problem 13). *If assumption (A11) is strictly violated,  $r_1^c e \notin C_0$ , then Problem 13 has a riskless solution for arbitrary  $\rho$ .*

*Proof.* Since  $C_0$  is convex,  $y \in \mathbb{R}^n$  exists so that  $(r - r_1^c e)^* y \geq 1 \forall r \in C_0$ . Given  $\rho \in \mathbb{R}$ , let  $x_0 := (\rho/r_2^c - r_1^c)y$ ,  $x_0^c := 1 - e^* x_0$ , and  $x_0^l := 0$  to obtain

$$w_j = (r_j - r_1^c e)^* x_0 + r_1^c \geq \rho/r_2^c$$

$\forall j \in L$ . Now let  $x_j := 0$ ,  $x_j^c := \rho/r_2^c$ , and  $x_j^l := w_j - \rho/r_2^c \geq 0$ .  $\square$

For the analysis of actual riskless solutions consider in  $\mathbb{R}^n$  the family of closed convex polyhedra

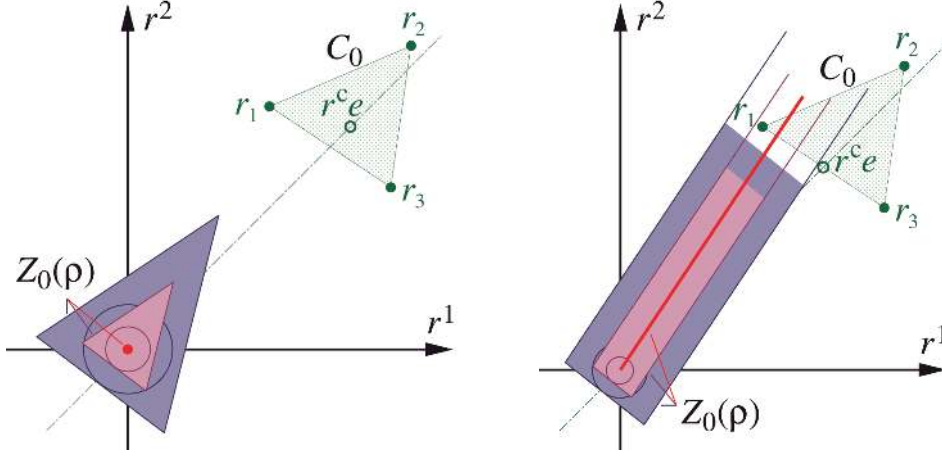
$$Z_0(\rho) := \left\{ x : \left( \frac{r_j}{r_1^c} - e \right)^* x \geq \frac{\rho}{r^c} - 1 \forall j \in L \right\} = \bigcap_{j \in L} \bar{H} \left( e - \frac{r_j}{r_1^c}, 1 - \frac{\rho}{r^c} \right).$$

LEMMA 2.12.  $Z_0(\rho)$  is nonempty (containing the origin) iff  $\rho \leq r^c$ . For  $\rho < r^c$ ,  $0 \in \text{int}(Z_0(\rho))$  and

$$Z_0(\rho) = \left( 1 - \frac{\rho}{r^c} \right) Z_0(0).$$

Each  $Z_0(\rho)$  is bounded and hence compact. In particular,  $Z_0(r^c) = \{0\}$ .

*Remark.* More generally one can show that the following four conditions are equivalent under the weaker no-arbitrage condition  $r_1^c e \in C_0$ .



**Fig. 2.2** First-period returns  $r_j, r_1^c e \in \mathbb{R}^2$  (green) and associated zero-risk polyhedra  $Z_0(\rho)$  with enclosed balls for three decreasing values of  $\rho$  (red =  $r^c$ , purple, blue). Left: Compact case,  $r_1^c e \in \text{int}(C_0)$ . Right: Unbounded case,  $r_1^c e \in \partial C_0$ .

- (1)  $r_1^c e \in \text{int}(C_0)$ .
- (2)  $Z_0(r^c) = \{0\}$ .
- (3)  $\exists \rho \in (-\infty, r^c]: Z_0(\rho)$  is bounded.
- (4)  $\forall \rho \in (-\infty, r^c]: Z_0(\rho)$  is bounded.

As in Theorem 1.28 we choose the stronger condition  $r_1^c e \in \text{int}(C_0)$  to avoid unnecessary technical complications and ambiguities.

*Proof.* By (A11) there exists a convex combination  $\sum_{j \in L} \xi_j (r_j - r_1^c e) = 0$ . Every  $x \in Z_0(\rho)$  then satisfies

$$0 = \sum_{j \in L} \xi_j \left( \frac{r_j}{r_1^c} - e \right)^* x \geq \frac{\rho}{r^c} - 1.$$

Hence  $Z_0(\rho) = \emptyset$  if  $\rho > r^c$  and  $0 \in Z_0(\rho)$  if  $\rho \leq r^c$ . Now let  $\rho < r^c$ . For any  $c > 0$ ,

$$\left( \frac{r_j}{r_1^c} - e \right)^* x \geq -1 \iff \left( \frac{r_j}{r_1^c} - e \right)^* cx \geq -c,$$

showing that  $x \in Z_0(0)$  iff  $(1 - \rho/r^c)x \in Z_0(\rho)$ . To verify  $0 \in \text{int}(Z_0(\rho))$ , observe that  $Z_0(\rho)$  contains the ball around the origin with radius

$$\left(1 - \frac{\rho}{r^c}\right) \left/ \max_{j \in L} \left\| \frac{r_j}{r_1^c} - e \right\|_2 \right. > 0;$$

cf. Figure 2.2. Condition (1) in the remark holds by (A11). We prove (2), (3), and (4) in natural order. Assume first that (2) does not hold and let  $y \in Z_0(r^c) \setminus \{0\}$ . Then  $(r_j/r_1^c - e)^* y \geq 0 \forall j \in L$  and hence  $(r/r_1^c - e)^* y \geq 0$  for  $r \in C_0$ , i.e.,  $C_0 \subseteq r_1^c e + \bar{H}(-y, 0)$ . This yields the contradiction  $r_1^c e \notin \text{int}(C_0)$  and proves (2), which obviously implies (3). Assume now that (4) does not hold, that is, that  $Z_0(\rho)$  is unbounded for some  $\rho \leq r^c$ . Then, since  $Z_0(\rho)$  is convex and  $0 \in Z_0(\rho)$ ,  $y \neq 0$  exists so that  $cy \in Z_0(\rho) \forall c > 0$ ; that is,  $(r_j/r_1^c - e)^* cy \geq \rho/r^c - 1 \forall j \in L$ . This

implies  $(r_j/r_1^c - e)^*cy \geq 0$  and hence  $(r_j/r_1^c - e)^*cy \geq \rho/r^c - 1 \forall c > 0, j \in L$ , and  $\rho \leq r^c$ . Thus (3) implies (4): either none of the polyhedra is bounded or all of them are.  $\square$

**THEOREM 2.13.** *Consider the feasible solutions of Problem 13. Then the following hold:*

- (1) *Risk vanishes iff  $x_j = 0$  and  $x_j^c = \rho/r_2^c$  in all scenarios; it is positive otherwise.*
- (2) *Leaf variables of riskless solutions depend uniquely on the root variables.*
- (3) *For  $\rho < r^c$ , risk vanishes on an  $(n + 1)$ -dimensional cone over  $Z_0(\rho)$ .*
- (4) *For  $\rho = r^c$ , the unique zero-risk portfolio has 100% cash.*
- (5) *For  $\rho > r^c$ , risk is strictly positive.*

*Remarks.* Strictly speaking, the riskless solutions for  $\rho < r^c$  form an  $(n + 1)$ -dimensional cone whose *projection* on the  $(x_0, x_0^l)$ -space has  $Z_0(\rho)$  as base; see Figure 2.3. Likewise, the projection for  $\rho = r^c$  is  $Z_0(r^c) \times \{0\}$ , and in both cases the remaining variables are uniquely determined by feasibility. (If one allows  $r_1^c e \in \partial C_0$ , then the variables  $x_0 \in Z_0(r^c)$ ,  $x_0^c$ , and  $x_j^l, j \in L$ , are not unique in the case  $\rho = r^c$ .)

*Proof.* The condition in (1) is obviously sufficient for zero risk (cf. Lemma 2.11). Let  $p := (p_j)_{j \in L}$  and  $q := (\rho_j(x_j, x_j^c))_{j \in L}$ . Then

$$R_d(x) = \sum_{j \in L} p_j q_j^2 - \left( \sum_{j \in L} p_j q_j \right)^2 = q^* [\text{Diag}(p) - pp^*] q.$$

By Lemma 2.14, this quadratic form vanishes if  $q = \rho e$  and is positive otherwise. Moreover,  $R_c(x) \geq 0 \forall x$ , and  $R_c(x) = 0$  iff  $x_j = 0 \forall j \in L$ . This proves statement (1). The zero-risk condition requires  $w_j \equiv r_j^* x_0 + r_1^c x_0^c \geq \rho/r_2^c$  in all scenarios. If this implied restriction on the root variables holds, then the leaf variables are uniquely determined by statement (1) and  $x_j^l = w_j - x_j^c \geq 0$ , proving statement (2). Since  $x_0^c = 1 - e^* x_0 - x_0^l$ , the inequality  $w_j \geq \rho/r_2^c$  is equivalent to  $x_0 \in Z_0(\rho + r^c x_0^l)$ . By Lemma 2.12 we conclude that  $x$  is optimal for  $\rho < r^c$  iff  $x_0^l \in [0, 1 - \rho/r^c]$  and  $x_0 \in Z_0(\rho + r^c x_0^l)$ ; see Figure 2.3. This proves statement (3). Statements (4) and (5) follow similarly: for  $\rho = r^c$  the zero-risk cone degenerates (only its vertex remains), and for  $\rho > r^c$  it becomes empty.  $\square$

**LEMMA 2.14.** *Let  $p \in \mathbb{R}^N, p > 0, e^* p = 1$ , and  $f(q) := \frac{1}{2} q^* [\text{Diag}(p) - pp^*] q$ . Then*

$$\min_q f(q) \quad \text{s.t.} \quad p^* q = \rho$$

*has the unique minimizer  $q = \rho e$ , with optimal value  $f(\rho e) = 0$ .*

*Proof.* The Lagrangian of the minimization problem is

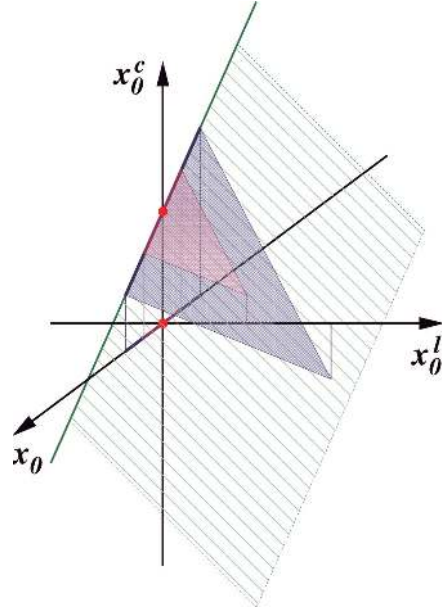
$$L(q, \eta) = \frac{1}{2} q^* [\text{Diag}(p) - pp^*] q - \eta (p^* q - \rho),$$

yielding the optimality condition

$$0 = \frac{\partial L}{\partial q} = \text{Diag}(p)q - pp^* q - \eta p = \text{Diag}(q)p - \rho p - \eta p.$$

Since  $p > 0$ , this gives  $q_j - \rho - \eta = 0$  for  $j = 1, \dots, N$ , and hence  $q = (\rho + \eta)e$ . Now  $p^* q = \rho$  implies  $\eta = 0$ , as required. Clearly,  $f(\rho e) = 0$ .  $\square$

Statement (3) of Theorem 2.13 says that  $x_0^l > 0$  (or  $x_j^l > 0$ ) is possible when the total cash return  $r^c = r_2^c r_1^c w_0$  (or the second-period cash return  $r_2^c w_j$ ) exceeds the



**Fig. 2.3** Zero-risk cones over  $Z_0(\rho)$  for three decreasing  $\rho$  values (red =  $r^c$ , purple, blue); compact case. (Here we have  $n = 1$ , and  $Z_0(\rho)$  is just a point or segment on the  $x_0$  axis.)

desired reward. The following result shows that surplus money is actually *necessary* for positive loss variables: if there is no surplus money, then positive amounts  $x_j^l$  are better invested in cash to reduce risk.

LEMMA 2.15. *Let  $x$  be an optimal solution of Problem 13 for  $\rho > r^c$ .*

- (1) *If  $x_j \neq 0$  in a scenario  $j \in L$ , then  $x_j^l = 0$ .*
- (2)  *$x_0^l = 0$ .*

*Remark.* Statement (1) of the lemma is also obtained from Theorem A.1, as well as the plausible fact that  $x_j^l = 0$  in scenario  $j$  if  $\rho_j < \rho$ . However, the proof of that theorem is not constructive.

*Proof.* Assume  $x_j^l > 0$ . We modify the local variables in scenario  $j$  to construct a better feasible solution. Let  $a := (\bar{r}_j - r_2^c) * x_j$ , define  $\epsilon \in (0, 1]$  as

$$\epsilon := \begin{cases} \min(1, r_2^c x_j^l / a) & \text{if } a > 0, \\ 1 & \text{else,} \end{cases}$$

and replace  $(x_j, x_j^c, x_j^l)$  by

$$(\hat{x}_j, \hat{x}_j^c, \hat{x}_j^l) := (x_j, x_j^c, x_j^l) - \epsilon(x_j, -\bar{r}_j^* x_j / r_2^c, a / r_2^c).$$

Then  $\hat{x}_j^l \geq 0$ ,  $\hat{w}_j = w_j$ ,  $\hat{\rho}_j = \rho_j$ , and the risk is reduced by the positive amount  $[1 - (1 - \epsilon)^2] x_j^* \Sigma_j x_j$ . This proves statement (1). For  $x_0^l > 0$  we modify  $x_0^c$  and all  $x_j^l$ . Let  $(\hat{x}_0, \hat{x}_0^c, \hat{x}_0^l) := (x_0, x_0^c + x_0^l, 0)$  and  $(\hat{x}_j, \hat{x}_j^c, \hat{x}_j^l) := (x_j, x_j^c, x_j^l + r_1^c x_0^l)$  for  $j \in L$ . Clearly,  $\hat{x}$  is feasible,  $R(\hat{x}) = R(x)$ , and  $\hat{x}_j^l \geq r_1^c x_0^l > 0$  in all scenarios. Thus, by statement (1),  $\hat{x}$  and consequently  $x$  are not optimal.  $\square$

*Discussion.* As in the one-period case, the 100% cash solution plays a key role: it has the largest reward among all riskless solutions. But now the solutions become

degenerate for small rewards  $\rho < r^c$ , even if loss is allowed only in the second period, i.e., if  $x_0^l = 0$  is fixed. This does not happen in Problem 12 (the modification of Problem 11 with minimal reward  $\rho(x) \geq \rho$ ), which behaves precisely as the corresponding single-period problem. Of course, the degeneracy occurs only for practically irrelevant rewards, and even then it can easily be avoided. (One may choose the vertex of the zero-risk cone, i.e.,  $x_0^l = 1 - \rho/r^c$ . This removes any surplus money immediately and gives a unique solution.)

The only case of practical interest is  $\rho > r^c$ , when the solution of Problem 11 remains optimal in Problem 12. Why do we prefer the loss formulation, Problem 13? Obviously the risk cannot be higher than in Problem 11 since every optimal solution of the latter remains feasible in the former problem. Actually it turns out that the loss formulation gives *strictly lower* risk in most cases; i.e., it allows better solutions than Problem 12. To develop a geometric understanding for this observation, we compare Problems 11 and 13 in a simplified situation. A reformulation eliminates all the budget equations and most of the portfolio variables in favor of the individual scenario returns. The two risk terms  $R_c, R_d$  are then used to explain in which cases (and how) an optimal solution of Problem 11 can be modified to give a better feasible solution of Problem 13.

PROBLEM 14. Consider as an example a portfolio consisting of just one risky asset and cash, using the second formulation of reward and risk in Lemma 2.8. Include loss assets in the leaves but not in the root, and write the problem with scenario returns  $\rho_j$  as additional variables,

$$\begin{aligned} \min_{x, \{\rho_j\}} \quad & \sum_{j \in L} \frac{1}{2} p_j [\sum_j x_j^2 + \rho_j^2] - \frac{1}{2} \rho^2 \\ \text{s.t.} \quad & \rho_j = \bar{r}_j x_j + r_2^c x_j^c \quad \forall j \in L, \\ & x_0 + x_0^c = 1, \\ & x_j + x_j^c + x_j^l = r_j x_0 + r_1^c x_0^c, \quad x_j^l \geq 0 \quad \forall j \in L, \\ & \sum_{j \in L} p_j \rho_j = \rho. \end{aligned}$$

This specialization of Problem 13 is only considered for  $\rho \geq r^c$ , but for arbitrary  $\rho \in \mathbb{R}$  as a specialization of Problem 11, i.e., when all the loss variables  $x_j^l = 0$  are fixed. (In these cases the solution is unique by Theorems 2.9 and A.1.)

LEMMA 2.16. *For simplicity assume  $\bar{r}_j \neq r_2^c \forall j \in L$ . (Assumption (A8) guarantees this for just one  $j \in L$ .) Let  $\phi_j := \sum_j / (\bar{r}_j - r_2^c)^2 > 0$ ,  $\psi_j := r_2^c(r_j - r_1^c)$ , and  $\theta_j \equiv r_2^c x_j^l$ . Then Problem 14 is equivalent to*

$$\begin{aligned} \min_{x_0, \{\rho_j, \theta_j\}} \quad & \sum_{j \in L} \frac{1}{2} p_j [\phi_j (\rho_j - r^c - \psi_j x_0 + \theta_j)^2 + \rho_j^2] - \frac{1}{2} \rho^2 \\ \text{s.t.} \quad & \sum_{j \in L} p_j \rho_j = \rho, \quad \theta_j \geq 0 \quad \forall j \in L. \end{aligned} \tag{2.1}$$

The optimal scenario returns are

$$\rho_j = \frac{\phi_j (r^c + \psi_j x_0 - \theta_j) + \mu}{\phi_j + 1}, \quad j \in L. \tag{2.2}$$

Moreover, the optimal reward multiplier  $\mu$  has the same value in these equivalent problems.

*Proof.* Eliminate

$$x_0^c = 1 - x_0, \quad x_j^c = (r_j - r_1^c)x_0 + r_1^c - x_j - x_j^l.$$

Then substitute  $x_j^c$  into the  $\rho_j$  equation and use  $\psi_j$  and  $\theta_j$  to obtain

$$\rho_j = (\bar{r}_j - r_2^c)x_j + r^c + \psi_j x_0 - \theta_j.$$

Solving for  $x_j$ , inserting it into the objective, and using  $\phi_j$  yields problem (2.1). Differentiating the Lagrangian with respect to the returns  $\rho_j$  gives optimality conditions

$$\phi_j(\rho_j - r^c - \psi_j x_0 + \theta_j) + \rho_j = \mu,$$

from which one obtains the expression (2.2). This derivation holds for the case without loss, too: one just has to set  $x_j^l = \theta_j = 0$  everywhere. Finally, when the problem transformations above are applied to the full primal-dual system of optimality conditions, it is observed that  $\mu$  has the same value in both problems.  $\square$

*Discussion.* Consider the case  $\theta_j = 0$  first (Problem 11). We have to choose optimal values for  $x_0$  and for the scenario returns  $\rho_j$  so that their mean equals  $\rho$ . Defining  $d_j := r^c + \psi_j x_0$  gives the continuous risk part

$$R_c = \sum_{j \in L} p_j \phi_j (\rho_j - d_j)^2 \geq 0.$$

This is a weighted average of scenario risks  $\phi_j(\rho_j - d_j)^2$ , each of which defines a parabola characterized by its offset  $d_j$  and curvature  $\phi_j$ . Both magnitude and distance of the offsets are influenced by the common “spread factor”  $x_0$ : they all coincide with  $r^c$  if  $x_0 = 0$  (100% cash), whereas  $x_0 = 1$  (no cash) yields the discrete distribution  $d_j = r_2^c r_j$  with mean  $d := r_2^c \mathbf{E}(\{r_j\})$ . Clearly, the continuous risk  $R_c$  is small when all the scenario returns are close to their respective offsets, while the discrete risk  $R_d$  is small when they are close to each other. Thus, loosely speaking,  $x_0$  has the job of balancing the scenarios by adjusting  $d$  (close to  $\rho$ ) without spreading the offsets too much.

Only one detail changes when  $\theta_j > 0$  is allowed (Problem 13): each offset  $d_j$  is replaced by  $d_j - \theta_j$ ; that is, the parabolas may be shifted to the left *separately* in each scenario.

We can now explain the risk reduction mechanism. Consider an optimal solution of Problem 14 *without* loss. Typically there will be “good” scenarios (fortunate for the investor, with large offsets  $d_j > \rho$ ) and “bad” scenarios (unfortunate for the investor, with small offsets  $d_j < \rho$ ). Moreover, one expects that some of the scenario returns will lie on their local upper branch ( $\rho_j > d_j$ ) and some on their local lower branch ( $\rho_j < d_j$ ). If in a good scenario the optimal return lies on the lower branch,  $\rho < \rho_j < d_j$ , then its contribution to the risk is canceled by shifting the parabola to the left,  $\theta_j := d_j - \rho_j > 0$ . Clearly, nothing else changes, so that this gives a suboptimal feasible solution of Problem 13 which is better than the optimal solution of Problem 11. The mechanism here is precisely the same as in a single period (Problem 5) except that it now occurs *locally* in individual scenarios.

Of course, the *optimal* solution of Problem 13 will readjust all the variables *globally*. Now  $x_0$  still has the job of balancing scenarios by adjusting  $d$  close to  $\rho$ , but only

without spreading the *small* offsets  $d_j < \rho$  too much. The large offsets  $d_j > \rho$  produce surplus money in *all* sufficiently good scenarios. These scenarios do not contribute to  $R_c$ , and all their returns are equal (and slightly larger than  $\rho$  to balance the bad ones). This is proved in Theorem A.1: surplus money  $x_j^l > 0$  implies  $r_2^c w_j > \mu$  and  $\rho_j = \mu \geq \rho$ . It means that a jump discontinuity is produced in the distribution of final wealth, to which each of the riskless scenarios contributes a fraction. It also means that the bad scenarios dominate the resulting risk; again we have an approximate minimization of downside risk.

When looking for an instance of Problem 11 with  $\rho < \rho_j < d_j$  one might try Problem 2.1 with only two scenarios. However, the following results show that the effect cannot occur if  $n = 1$  and  $N = 2$ : in that case (with  $\rho > r^c$ ) the single degree of freedom in  $x_0$  is sufficient to balance the scenarios well. After proving this we give an example of the risk reduction with  $n = 1$  and  $N = 3$ . A slight modification finally shows that risk reduction can occur even in bad scenarios; that is, optimal solutions of Problem 11 may have  $\rho_j < d_j < \rho$ .

For the comparison of signs we define the equivalence relation

$$a \sim b \quad :\iff \quad \text{sign}(a) = \text{sign}(b) \in \{-1, 0, 1\}$$

and use the fact that  $\{-1, 0, 1\}$  is a multiplicative subgroup in  $\mathbb{R}$ .

LEMMA 2.17. *In Problem 14 without loss assets,  $\rho_j - d_j \sim s_j(\rho - r^c)$ , where*

$$s_j := \sum_{k \in L \setminus \{j\}} \frac{p_k}{\delta_k^c + 1} (r_k - r_j)(r_k - r_1^c).$$

*Proof.* See the appendix.  $\square$

*Remark.* Notice that  $s_j$  depends only on the stochastic data (the return distribution) and in particular that it does *not* depend on  $\rho$ .

COROLLARY 2.18. *In Problem 14 with two scenarios and without loss assets, both  $\rho_1$  and  $\rho_2$  lie on their respective upper branches if  $\rho > r^c$ .*

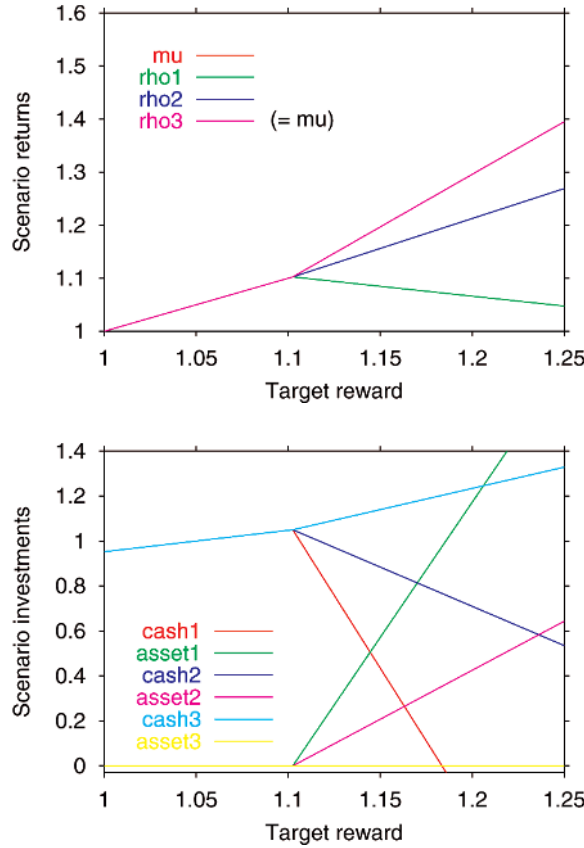
*Proof.* Assumption (A11) yields  $r_1 < r_1^c < r_2$ , and the sums  $s_j$  consist of one single term each since  $N = 2$ . This implies that  $s_1, s_2$  are both positive.  $\square$

*Remark.* A weak version of this result holds for  $r_1^c \in C_0$ . Then  $r_1 \leq r_1^c \leq r_2$ , and  $s_1, s_2$  are both nonnegative. Even that weaker version would prevent any beneficial shift  $\theta_j > 0$ .

The example with  $N = 3$  is simple: let  $(p_1, p_2, p_3) = (0.2, 0.6, 0.2)$ ,  $(r_1, r_2, r_3) = (1.0, 1.1, 1.2)$ ,  $r_1^c = 1.05$ ,  $\Sigma_1 = \Sigma_2 > 0$ , and  $j = 3$ . Then  $\delta_1^c = 0.05^2/\Sigma_1 = \delta_2^c$ , and hence

$$s_3 = \sum_{k=1}^2 \frac{p_k}{\delta_k^c + 1} (r_k - r_3)(r_k - r_1^c) = \frac{0.2(-0.2)(-0.05) + 0.6(-0.1)(0.05)}{\delta_1^c + 1} < 0.$$

The corollary proves that  $\rho_3$  lies on the lower branch even though  $r_1^c \in \text{int}(C_0)$ ; thus introducing the loss asset  $x_3^l$  reduces the optimal risk. (The values of  $\Sigma_3$ ,  $r_2^c$ , and  $\bar{r}_1, \bar{r}_2, \bar{r}_3$  do not matter here as long as (A7), (A8), (A10), and (A11) are satisfied. The reader may check that this holds for  $\Sigma_j = 1$ ,  $r_2^c = 1.05$ , and  $\bar{r}_j = 1.1$ ,  $j \in \{1, 2, 3\}$ .) Setting instead  $r_1^c := 1.15$  and  $\Sigma_2 = \Sigma_3 > 0$  yields  $s_1 < 0$ . This shows that risk reduction can also occur in bad scenarios—at least for unreasonably large  $r_1^c$ . Results for this small example problem (with loss assets included) are displayed in Figures 2.4 and 2.5.



**Fig. 2.4** Example problem with three scenarios; unconstrained asset values. Top: Optimal scenario returns  $\rho_i$  and reward multiplier  $\mu$ . Bottom: Optimal scenario investments in period 2. For  $\rho \leq r^c = 1.1025$ , the optimal portfolios consist of 100% cash, yielding zero risk.

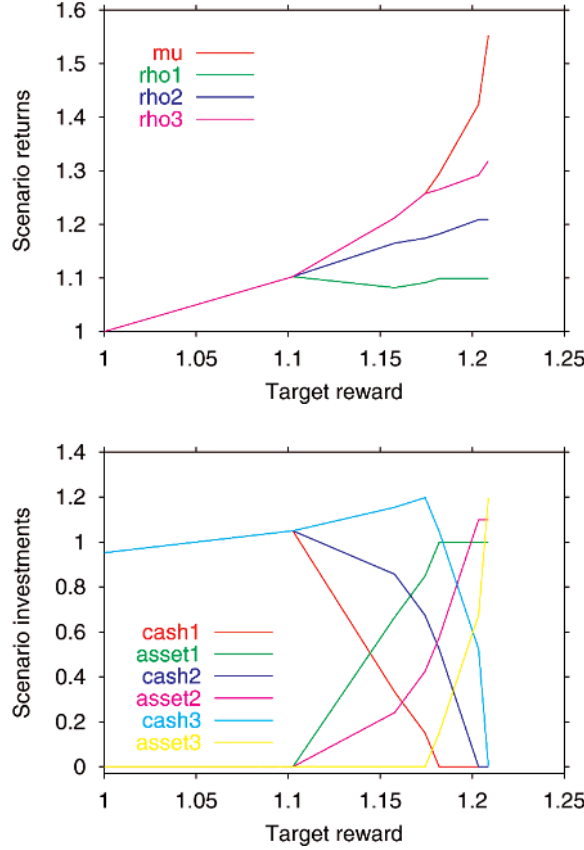
**2.4. The General Multiperiod Case.** The general situation with loss assets is a straightforward extension of the two-period case. The following results can be proved by suitable modifications in the first part of section 2.3 (up to Theorem 2.13).

Surplus money can now appear in any node  $j \in L_t$ ,  $t \in \{0, \dots, T\}$ , if the (partial) scenario is sufficiently good up to that point, that is, if

$$w_j > w_t^{\min} := \frac{\mu}{r_{T+1}^c \cdots r_{t+1}^c}.$$

(As before, the root can be excluded by considering only “large” target rewards  $\rho > r^c$ , with  $r^c := r_{T+1}^c \cdots r_1^c$ .) When there is surplus money, the whole scenario subtree rooted in  $j$  does not contribute to the continuous risk provided that sufficient node capitals  $w_k$  are maintained. Implicitly, this condition defines a zero-risk polyhedron whose geometry is determined by the subtree’s discrete return distribution. Of course, if  $j \in L_{T-1}$ , then one gets a cone similar to the one considered before, but depending on  $w_j$ . These observations imply that *the generic optimal solution is highly degenerate*: any reasonable return discretization will include good and bad partial scenarios on each level, and surplus money will almost always appear somewhere in the tree.





**Fig. 2.5** Example problem with three scenarios; nonnegative asset values ( $\bar{r}^{\max} = 1.21$ ). Top: Optimal scenario returns  $\rho_i$  and reward multiplier  $\mu$ . Bottom: Optimal scenario investments. For  $\rho \leq r^c = 1.1025$ , the optimal portfolios consist of 100% cash, yielding zero risk.

However, even in the zero-risk subtrees, all leaf variables are again uniquely determined by the lower level variables.

Obviously  $w_{t+s}^{\min}$  is sufficient capital for a node  $k \in L_{t+s}$  in the zero-risk subtree, and the easiest way to maintain that amount is to invest precisely  $w_t^{\min}$  in cash and remove the rest (“invest” in  $x_j^l$ ). Thus all surplus money is taken out immediately in the root of the subtree, and each remaining node has 100% cash.

**3. Conclusions.** We have seen that multiperiod mean-variance problems behave much like their single-period counterparts in many respects. Specifically, it is possible to avoid overperformance by allowing the removal of capital. Small target rewards  $\rho \leq r^c$  are met exactly at zero risk. In that case all the capital is either invested in cash or removed; thus, minimizing the variance is trivially equivalent to minimizing the semivariance (or any other downside risk measure) *without* removing capital, but allowing it to exceed the target reward. That is, with  $x = (x_j, x_j^c)_{j \in V}$  and in abbreviated notation, the problem

$$\min_x R(x) \quad \text{s.t.} \quad \rho(x) = \rho, \quad e^* x_j + x_j^c \leq w_j \quad \forall j \in V$$

is equivalent to the downside risk problem

$$\min_x R_\rho^2(x) \quad \text{s.t.} \quad \rho(x) \geq \rho, \quad e^*x_j + x_j^c = w_j \quad \forall j \in V.$$

(Of course, the solutions of the second problem differ insofar as surplus money is invested in cash instead of being removed.) For moderate values  $\rho > r^c$  one cannot avoid overperformance completely, but in effect the first problem still tends to minimize the semivariance. More precisely, the discrete part  $R_d$  approximates its downside version due to the existence of zero-risk subtrees. The quality of that approximation decreases as  $\rho$  increases so that for large values the risk measure becomes a blending of variance and semivariance. Note that there is no such gradual process in the single-period case, but there is a close similarity between single-period downside risk (Theorem 1.28) and multiperiod zero-risk polyhedra (Lemma 2.12). We may conclude that Problem 13 is a reasonable multiperiod model for an investor who wishes to minimize the semivariance rather than the variance of final wealth.

The previous comparison also gives some hints on how an optimal policy obtained from Problem 13 should be interpreted. Again, positive values of  $x_j^l$  do not suggest the burning of that amount. They indicate the presence of surplus money which the investor may spend immediately without risking missing her goal, or which she may invest in cash to obtain a riskless extra profit. Of course, she may also consume part of the surplus and invest the rest. Thus, if the investor implements any optimal policy over the full planning horizon, she will approximately minimize the risk of ending up with less than the desired amount, regardless of her choice. Interestingly, the second alternative (investing) amounts to a *single-period* strategy with *predetermined intermediate decisions*, which may be useful when the investor cannot react to the market until the end of her planning horizon or for some reason does not wish to do so.

However, it should be noted that the problem under consideration is not *time-invariant* in the sense that optimal decisions in later periods still reflect the investor's initial objective when time has passed. This is due to the reward condition, which involves an expectation over all scenarios, that is, over the potential futures at  $t = 0$ . But at  $t = 1$  most of these potential futures become impossible, regardless of which scenario is realized. The terminal condition  $\rho(x) = \rho$  or  $\rho(x) \geq \rho$  usually cannot be satisfied when the restricted expectation over the subtree is taken—unless it happens to be a zero-risk subtree. Therefore only the immediate decision will be of interest for the typical investor. Rather than following the original future policy, she will adjust the reward and solve the problem anew for each decision. Of course, the investor may also build an extended model after each period in pursuing a moving horizon technique.

In any case it seems appropriate to consider all riskless strategies (in addition to the efficient ones) as reasonable choices in multiperiod decision models. This does no harm since it includes all the standard alternatives, but it opens up new possibilities such as the trick described above.

We conclude the discussion by pointing out two issues that might be interesting subjects of future research. First, the model presented here does not include any preferences of consumption, although one can easily specify hard constraints (exact, minimal, or maximal consumption) through a cash flow. However, it is not clear how one should incorporate (soft) preferences and how the result would be related to long-term models based on utility of consumption. Second, the multiperiod setting enables the investor to control higher moments of her distribution of final wealth—at least to some extent. How would risk measures involving skewness, for example, behave in the context of our model?

On the practical side, stochastic programming models are in the process of becoming standard decision support tools in the financial industry. Optimization problems with tens of thousands of scenarios and a million decision variables are nowadays solved within minutes on a standard workstation. Indeed, it is more difficult to generate the stochastic data that are required in setting up the model. Typical applications of stochastic programming include portfolio selection and asset liability management problems. For instance, the real-life version of the model considered here is being employed by a Swiss insurance company in the allocation of pension funds. A second major application field for related technology, such as stochastic differential equations and martingale theory, concerns the valuation and hedging of options and other derivatives and the design of new financial instruments.

**Appendix.** The appendix contains some proofs and a theorem that would have disrupted the line of thought in the main body of the multiperiod section.

*Proof of Theorem 2.4.* The system of optimality conditions (for two periods) can be written

$$\left( \begin{array}{ccc|ccc} 0 & & & e & -r_1 & \dots & -r_N & 0 \\ & \tilde{\Sigma}_1 & & & e & & & \tilde{r}_1 \\ & & \ddots & & & & & \vdots \\ & & & & & & & \tilde{r}_N \\ \hline & & & \tilde{\Sigma}_N & & & e & \\ e^* & & & & & & & \\ -r_1^* & e^* & & & & & & \\ \vdots & & \ddots & & & & & \\ -r_N^* & & & & & & e^* & \\ 0 & \tilde{r}_1^* & \dots & \tilde{r}_N^* & & & & \end{array} \right) \begin{pmatrix} x_0 \\ x_1 \\ \vdots \\ x_N \\ -\lambda_0 \\ -\lambda_1 \\ \vdots \\ -\lambda_N \\ -\mu \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \\ \rho \end{pmatrix}.$$

For  $j \in L$ ,  $x_j = \tilde{\Sigma}_j^{-1}(\lambda_j e + \mu \tilde{r}_j)$  is immediately obtained from the  $j$ th dual feasibility condition. Substitution into the budget equation for  $x_j$  and into dual feasibility condition 0, respectively, yields  $\lambda_j = (r_j^* x_0 - \mu \tilde{\beta}_j) / \tilde{\alpha}_j$  and

$$0 = -\lambda_0 e + \sum_{j \in L} \lambda_j r_j = -\lambda_0 e + \sum_{j \in L} \frac{r_j^* x_0 - \mu \tilde{\beta}_j}{\tilde{\alpha}_j} r_j = -\lambda_0 e + \tilde{\Sigma}_0 x_0 - \mu \tilde{r}_0,$$

giving  $x_0 = \tilde{\Sigma}_0^{-1}(\lambda_0 e + \mu \tilde{r}_0)$ . The budget equation for  $x_0$  now reads

$$1 = e^* x_0 = e^* \tilde{\Sigma}_0^{-1}(\lambda_0 e + \mu \tilde{r}_0) = \lambda_0 \tilde{\alpha}_0 + \mu \tilde{\beta}_0,$$

yielding  $\lambda_0 = (1 - \mu \tilde{\beta}_0) / \tilde{\alpha}_0$ . From the reward equation one finally obtains

$$\begin{aligned} \rho &= \sum_{j \in L} \tilde{r}_j^* x_j = \sum_{j \in L} (\lambda_j \tilde{\beta}_j + \mu \tilde{\gamma}_j) = \sum_{j \in L} \left( \frac{\tilde{\beta}_j}{\tilde{\alpha}_j} r_j^* x_0 - \mu \frac{\tilde{\beta}_j^2}{\tilde{\alpha}_j} + \mu \tilde{\gamma}_j \right) \\ &= \tilde{r}_0^* x_0 + \mu \sum_{j \in L} \frac{\tilde{\delta}_j}{\tilde{\alpha}_j} = \lambda_0 \tilde{\beta}_0 + \mu \tilde{\gamma}_0 + \mu \sum_{j \in L} \frac{\tilde{\delta}_j}{\tilde{\alpha}_j} \\ &= \frac{\tilde{\beta}_0}{\tilde{\alpha}_0} - \mu \frac{\tilde{\beta}_0^2}{\tilde{\alpha}_0} + \mu \tilde{\gamma}_0 + \mu \sum_{j \in L} \frac{\tilde{\delta}_j}{\tilde{\alpha}_j} = \frac{\tilde{\beta}_0}{\tilde{\alpha}_0} + \mu \sum_{j \in V} \frac{\tilde{\delta}_j}{\tilde{\alpha}_j}. \end{aligned}$$

This gives the reward multiplier using Lemma 2.3,

$$\mu = \left( \rho - \frac{\tilde{\beta}_0}{\tilde{\alpha}_0} \right) / \sum_{j \in V} \frac{\tilde{\delta}_j}{\tilde{\alpha}_j}.$$

(Here we need assumption (A6): the denominator vanishes if  $\tilde{\delta}_j = 0 \forall j \in V$ .) As for the global minimum, we have

$$\begin{aligned} x_j^* \tilde{\Sigma}_j x_j &= \lambda_j (\lambda_j \tilde{\alpha}_j + \mu \tilde{\beta}_j) + \mu (\lambda_j \tilde{\beta}_j + \mu \tilde{\gamma}_j) \\ &= \lambda_j r_j^* x_0 + \mu \frac{\tilde{\beta}_j r_j^* x_0 + \mu \tilde{\delta}_j}{\tilde{\alpha}_j} = \frac{(r_j^* x_0)^2}{\tilde{\alpha}_j} + \mu^2 \frac{\tilde{\delta}_j}{\tilde{\alpha}_j} \end{aligned}$$

and similarly

$$x_0^* \tilde{\Sigma}_0 x_0 = \lambda_0 + \mu \frac{\tilde{\beta}_0 + \mu \tilde{\delta}_0}{\tilde{\alpha}_0} = \frac{1}{\tilde{\alpha}_0} + \mu^2 \frac{\tilde{\delta}_0}{\tilde{\alpha}_0}.$$

Therefore,

$$\begin{aligned} \sum_{j \in L} x_j^* \tilde{\Sigma}_j x_j &= \sum_{j \in L} \left( \frac{(r_j^* x_0)^2}{\tilde{\alpha}_j} + \mu^2 \frac{\tilde{\delta}_j}{\tilde{\alpha}_j} \right) = x_0^* \tilde{\Sigma}_0 x_0 + \mu^2 \sum_{j \in L} \frac{\tilde{\delta}_j}{\tilde{\alpha}_j} \\ &= \frac{1}{\tilde{\alpha}_0} + \mu^2 \sum_{j \in V} \frac{\tilde{\delta}_j}{\tilde{\alpha}_j} = \frac{1}{\tilde{\alpha}_0} + \left( \rho - \frac{\tilde{\beta}_0}{\tilde{\alpha}_0} \right)^2 / \sum_{j \in V} \frac{\tilde{\delta}_j}{\tilde{\alpha}_j}. \end{aligned}$$

Subtracting  $\rho^2$  yields a risk expression of the form  $\sigma^2(\rho) = s + (\rho - c)^2 / d - \rho^2$ . Since the optimal portfolio  $x$  is an affine function of  $\rho$ , and  $R$  is convex quadratic, the efficient frontier either is strictly convex (iff  $d < 1$ ) or  $\sigma^2(\rho) \equiv 0$  (iff  $d = 1$  and  $c = s = 0$ ). But  $s = \tilde{\alpha}_0^{-1} > 0$  by Lemma 2.3; therefore  $\sigma^2(\rho)$  has the global minimum  $s + c^2 / (1 - d)$  at  $\hat{\rho} = c / (1 - d)$ , as stated.  $\square$

*Proof of Theorem 2.9.* The system of optimality conditions can be written

$$\begin{aligned} \partial L / \partial x_0 &= 0: & -\lambda_0 e + \sum_{j \in L} \lambda_j r_j &= 0, \\ \partial L / \partial x_0^c &= 0: & -\lambda_0 + \sum_{j \in L} \lambda_j r_1^c &= 0, \\ \partial L / \partial x_j &= 0: & \tilde{\Sigma}_j x_j + r_2^c \tilde{r}_j x_j^c - \lambda_j e - \mu \tilde{r}_j &= 0 \quad \forall j \in L, \\ \partial L / \partial x_j^c &= 0: & r_2^c \tilde{r}_j^* x_j + r_2^c \tilde{r}_j^c x_j^c - \lambda_j - \mu \tilde{r}_j^c &= 0 \quad \forall j \in L, \\ \partial L / \partial \lambda_0 &= 0: & e^* x_0 + x_0^c &= 1, \\ \partial L / \partial \lambda_j &= 0: & e^* x_j + x_j^c - r_j^* x_0 - r_1^c x_0^c &= 0 \quad \forall j \in L, \\ \partial L / \partial \mu &= 0: & \sum_{j \in L} \tilde{r}_j^* x_j + \tilde{r}_j^c x_j^c &= \rho. \end{aligned}$$

The dual feasibility condition  $\partial L / \partial x_j^c = 0$  gives

$$(A.1) \quad x_j^c = \frac{1}{r_2^c} \left[ -\tilde{r}_j^* x_j + \frac{\lambda_j}{\tilde{r}_j^c} + \mu \right].$$

Substitution into  $\partial L/\partial x_j = 0$  yields

$$0 = \tilde{\Sigma}_j x_j + \tilde{r}_j \left[ -\tilde{r}_j^* x_j + \frac{\lambda_j}{\tilde{r}_j^c} + \mu \right] - \lambda_j e - \mu \tilde{r}_j = p_j \Sigma_j x_j + \frac{\lambda_j}{r_2^c} (\bar{r}_j - r_2^c e),$$

which gives  $x_j$  and, upon substitution into (A.1),  $x_j^c$ :

$$x_j = -\frac{\lambda_j}{\tilde{r}_j^c} \Sigma_j^{-1} (\bar{r}_j - r_2^c e), \quad x_j^c = \frac{1}{r_2^c} \left[ \frac{\lambda_j}{\tilde{r}_j^c} (\gamma_j - r_2^c \beta_j + 1) + \mu \right].$$

Therefore the budget equation  $\partial L/\partial \lambda_j = 0$  reads

$$w_j = e^* x_j + x_j^c = -\frac{\lambda_j}{\tilde{r}_j^c} (\beta_j - r_2^c \alpha_j) + \frac{\lambda_j}{r_2^c \tilde{r}_j^c} (\gamma_j - r_2^c \beta_j + 1) + \frac{\mu}{r_2^c} = \frac{\lambda_j}{r_2^c} \frac{\delta_j^c + 1}{\tilde{r}_j^c} + \frac{\mu}{r_2^c},$$

from which one obtains

$$\lambda_j = \frac{\tilde{r}_j^c}{\delta_j^c + 1} [r_2^c (r_j^* x_0 + r_1^c x_0^c) - \mu].$$

Now we can proceed with the root variables. The condition  $\partial L/\partial x_0^c = 0$  reads

$$0 = -\lambda_0 + \sum_{j \in L} r_1^c \frac{\tilde{r}_j^c}{\delta_j^c + 1} [r_2^c (r_j^* x_0 + r_1^c x_0^c) - \mu] = -\lambda_0 + r^c r_2^c \tilde{r}_0^* x_0 + r^c \tilde{r}_1^c x_0^c - \mu \tilde{r}_1^c,$$

giving

$$(A.2) \quad x_0^c = \frac{1}{r^c} \left[ -r_2^c \tilde{r}_0^* x_0 + \frac{\lambda_0}{\tilde{r}_1^c} + \mu \right].$$

Likewise, after inserting  $\lambda_j$  and then  $x_0^c$ , the condition  $\partial L/\partial x_0 = 0$  reads

$$\begin{aligned} 0 &= -\lambda_0 e + \sum_{j \in L} r_j \frac{\tilde{r}_j^c}{\delta_j^c + 1} [r_2^c (r_j^* x_0 + r_1^c x_0^c) - \mu] \\ &= -\lambda_0 e + (r_2^c)^2 \tilde{\Sigma}_0 x_0 + r^c r_2^c \tilde{r}_0^* x_0 - \mu r_2^c \tilde{r}_0 = (r_2^c)^2 \tilde{p}_0 \Sigma_0 x_0 + \frac{\lambda_0}{r_1^c} (\tilde{r}_0 - r_1^c e). \end{aligned}$$

This yields  $x_0$  and, by substitution into (A.2),  $x_0^c$ :

$$x_0 = -\frac{\lambda_0}{r_2^c \tilde{r}_1^c} \Sigma_0^{-1} (\tilde{r}_0 - r_1^c e), \quad x_0^c = \frac{\lambda_0}{r^c \tilde{r}_1^c} (\gamma_0 - r_1^c \beta_0 + 1) + \frac{\mu}{r^c}.$$

Thus the root budget equation is

$$1 = e^* x_0 + x_0^c = -\frac{\lambda_0}{r_2^c \tilde{r}_1^c} (\beta_0 - r_1^c \alpha_0) - \frac{\lambda_0}{r^c \tilde{r}_1^c} (\gamma_0 - r_1^c \beta_0 + 1) + \frac{\mu}{r^c} = \frac{\lambda_0}{r^c} \frac{\delta_0^c + 1}{\tilde{r}_1^c} + \frac{\mu}{r^c},$$

which gives

$$\lambda_0 = \frac{\tilde{r}_1^c}{\delta_0^c + 1} (r^c - \mu) = \tilde{\rho} (r^c - \mu).$$

By Lemma 2.7 we have  $\tilde{p}_0 \in (0, 1]$  and  $\delta_0^c \geq 0$ . Moreover, if  $\tilde{p}_0 = 1$  then  $\delta_j^c = 0$  for  $j \in L$  and thus  $\delta_0^c > 0$ . This proves the inclusion  $\tilde{\rho} \equiv r^c \tilde{p}_0 / (\delta_0^c + 1) \in (0, r^c)$ . Altogether, the previous results give

$$(A.3) \quad \tilde{r}_j^* x_j + \tilde{r}_j^c x_j^c = -\frac{\lambda_j}{r_2^c} (\gamma_j - r_2^c \beta_j) + \frac{\lambda_j}{r_2^c} (\gamma_j - r_2^c \beta_j + 1) + p_j \mu = \frac{\lambda_j}{r_2^c} + p_j \mu$$

and similarly

$$(A.4) \quad r_2^c \tilde{r}_0^* x_0 + \tilde{r}_1^c x_0^c = -\frac{\lambda_0}{r^c} (\gamma_0 - r_1^c \beta_0) + \frac{\lambda_0}{r^c} (\gamma_0 - r_1^c \beta_0 + 1) + \tilde{p}_0 \mu = \frac{\lambda_0}{r^c} + \tilde{p}_0 \mu.$$

Upon insertion of  $\lambda_j$  and then  $\lambda_0$  the reward condition reads

$$\begin{aligned} \rho &= \sum_{j \in L} \tilde{r}_j^* x_j + \tilde{r}_j^c x_j^c = \sum_{j \in L} \frac{p_j}{\delta_j^c + 1} [r_2^c (r_j^* x_0 + r_1^c x_0^c) - \mu] + \mu \\ &= r_2^c \tilde{r}_0^* x_0 + \tilde{r}_1^c x_0^c - \tilde{p}_0 \mu + \mu = \frac{\lambda_0}{r^c} + \mu = \frac{\tilde{\rho}}{r^c} (r^c - \mu) + \mu. \end{aligned}$$

This gives  $\rho = \tilde{\rho} + \mu(r^c - \tilde{\rho})/r^c$  and thus  $\mu = r^c(\rho - \tilde{\rho})/(r^c - \tilde{\rho})$ . To calculate the risk, let  $\bar{R}_j(x_j, x_j^c) := x_j^* \sum_j x_j + \rho_j(x_j, x_j^c)^2$  and use the last expression from equation (A.3) for  $p_j \rho_j(x_j, x_j^c)$ . Then

$$\begin{aligned} \bar{R}_j(x_j, x_j^c) &= \left( \frac{\lambda_j}{\tilde{r}_j^c} \right)^2 \delta_j^c + \left( \frac{\lambda_j}{\tilde{r}_j^c} + \mu \right)^2 = \left( \frac{\lambda_j}{\tilde{r}_j^c} \right)^2 (\delta_j^c + 1) + 2 \frac{\lambda_j}{\tilde{r}_j^c} \mu + \mu^2 \\ &= \frac{[r_2^c (r_j^* x_0 + r_1^c x_0^c) - \mu]^2}{\delta_j^c + 1} + 2 \frac{[r_2^c (r_j^* x_0 + r_1^c x_0^c) - \mu]}{\delta_j^c + 1} \mu + \mu^2 \\ &= \frac{[r_2^c (r_j^* x_0 + r_1^c x_0^c)]^2}{\delta_j^c + 1} + \mu^2 \frac{\delta_j^c}{\delta_j^c + 1}. \end{aligned}$$

Since  $R(x) = \sum_{j \in L} p_j \bar{R}_j(x_j, x_j^c) - \rho^2$ , we get

$$\begin{aligned} R(x) + \rho^2 &= (r_2^c)^2 \sum_{j \in L} \frac{p_j}{\delta_j^c + 1} [(r_j^* x_0)^2 + 2r_1^c x_0^c r_j^* x_0 + (r_1^c x_0^c)^2] + \mu^2 \sum_{j \in L} \frac{p_j \delta_j^c}{\delta_j^c + 1} \\ &= (r_2^c)^2 [x_0^* \tilde{\Sigma}_0 x_0 + 2r_1^c x_0^c \tilde{r}_0^* x_0 + \tilde{p}_0 (r_1^c x_0^c)^2] + \mu^2 (1 - \tilde{p}_0). \end{aligned}$$

Using equation (A.4), the first term equals

$$r_2^c \tilde{r}_j^c [x_0^* \Sigma_0 x_0 + (\tilde{r}_0^* x_0 + r_1^c x_0^c)^2] = r_2^c \tilde{r}_j^c \left[ \left( \frac{\lambda_0}{r_2^c \tilde{r}_1^c} \right)^2 \delta_0^c + \left( \frac{\lambda_0}{r_2^c \tilde{r}_1^c} + \frac{\mu}{r_2^c} \right)^2 \right],$$

which is further simplified to

$$\tilde{p}_0 \left[ \left( \frac{\lambda_0}{\tilde{r}_1^c} \right)^2 (\delta_0^c + 1) + 2 \frac{\lambda_0}{\tilde{r}_1^c} \mu + \mu^2 \right] = \tilde{p}_0 \left[ \frac{(r^c - \mu)^2}{\delta_0^c + 1} + 2 \frac{r^c - \mu}{\delta_0^c + 1} \mu + \mu^2 \right].$$

Therefore we have

$$\begin{aligned} R(x) + \rho^2 &= \tilde{p}_0 \frac{(r^c)^2 + \delta_0^c \mu^2}{\delta_0^c + 1} + \mu^2 (1 - \tilde{p}_0) = r^c \tilde{\rho} + \mu^2 \left( 1 - \frac{\tilde{p}_0}{\delta_0^c + 1} \right) \\ &= r^c \tilde{\rho} + \mu^2 \left( 1 - \frac{\tilde{\rho}}{r^c} \right) = r^c \tilde{\rho} + r^c \frac{(\rho - \tilde{\rho})^2}{r^c - \tilde{\rho}}. \end{aligned}$$

Subtracting  $\rho^2$  gives the stated risk formula whose minimum over all  $\rho$  is easily determined. The remaining statements follow trivially.  $\square$

*Proof of Theorem 2.10.* The optimality conditions include those of Problem 11 except that  $\rho$  must be replaced by  $\rho + \theta$  (in  $\partial L/\partial \mu = 0$ ). Additional conditions are  $\mu - (\rho + \theta) = \eta$  (from  $\partial L/\partial \theta = 0$ ),  $\theta \geq 0$ ,  $\eta \geq 0$ , and complementarity  $\theta\eta = 0$ . The proof now proceeds precisely as in the proof of Theorem 2.9, with  $\rho$  replaced by  $\rho + \theta$  in all instances. Finally, nonnegativity and complementarity of  $\theta$  and  $\eta$  lead to the case distinction: either  $\theta = 0$  and  $\mu \geq \rho$ , or  $\theta > 0$  and  $\mu = \rho + \theta$ . The identity  $\mu = r^c(\rho + \theta - \tilde{\rho})/(r^c - \tilde{\rho})$  gives  $\rho \geq r^c$  in the first case and  $\rho + \theta = r^c$  in the second case.  $\square$

THEOREM A.1. In Problem 13 let  $\theta_j \equiv r_2^c x_j^l$  for  $j \in L$ ,  $\theta_0 \equiv r^c x_0^l$ , and

$$\bar{\theta}_0 := \frac{1}{\tilde{p}_0} \sum_{j \in L} \frac{p_j}{\delta_j^c + 1} \theta_j, \quad s_0 := \frac{1}{\tilde{p}_0} \sum_{j \in L} \frac{p_j}{\delta_j^c + 1} \theta_j r_j.$$

Then every primal-dual solution satisfies

$$\begin{aligned} x_j &= -\frac{\lambda_j}{\tilde{r}_j^c} \Sigma_j^{-1} (\bar{r}_j - r_2^c e), & x_j^c &= \frac{1}{r_2^c} \left[ \frac{\lambda_j}{\tilde{r}_j^c} (\gamma_j - r_2^c \beta_j + 1) + \mu \right], \\ x_0 &= -\frac{\lambda_0}{r_2^c \tilde{r}_1^c} \Sigma_0^{-1} (\bar{r}_0 - r_1^c e) - \frac{1}{r_2^c} \Sigma_0^{-1} (\bar{\theta}_0 \bar{r}_0 - s_0), \\ x_0^c &= \frac{1}{r^c} \left[ \frac{\lambda_0}{\tilde{r}_1^c} (\gamma_0 - r_1^c \beta_0 + 1) + \mu + (\gamma_0 + 1) \bar{\theta}_0 - \bar{r}_0^* \Sigma_0^{-1} s_0 \right], \\ \lambda_j &= \frac{\tilde{r}_j^c}{\delta_j^c + 1} [r_2^c (r_j^* x_0 + r_1^c x_0^c) - \mu - \theta_j] = -\eta_j \leq 0, \\ \lambda_0 &= \frac{\tilde{r}_1^c}{\delta_0^c + 1} [r^c - \mu - \theta_0 - (\gamma_0 - r_1^c \beta_0 + 1) \bar{\theta}_0 + (\bar{r}_0 - r_1^c e)^* \Sigma_0^{-1} s_0] = -\eta_0 \leq 0, \\ \mu &= r^c \frac{\rho - \tilde{\rho}}{r^c - \tilde{\rho}} + \frac{\tilde{\rho}}{r^c - \tilde{\rho}} [(\gamma_0 - r_1^c \beta_0 + 1) \bar{\theta}_0 + \theta_0 - (\bar{r}_0 - r_1^c e)^* \Sigma_0^{-1} s_0]. \end{aligned}$$

In particular, the following case distinction can be made in each scenario. If  $\eta_j > 0$ , then  $x_j^l = 0$ ,  $\rho_j(x_j, x_j^c) < \mu$ , and  $r_2^c w_j < \mu$ . Conversely, if  $\eta_j = 0$ , then  $x_j^l \geq 0$ ,  $\rho_j(x_j, x_j^c) = \mu \geq \rho$ , and  $r_2^c w_j \geq \mu$ . In this case the leaf variables are

$$x_j = 0, \quad x_j^c = \mu/r_2^c, \quad x_j^l = w_j - \mu/r_2^c,$$

giving  $x_j^l = 0$  (100% cash) if  $r_2^c w_j = \mu$  and  $x_j^l > 0$  else.

The following case distinction holds in the root. If  $\eta_0 > 0$ , then  $x_0^l = 0$  and  $\rho < \mu$ . Otherwise, if  $\eta_0 = 0$ , then  $x_0^l \geq 0$  and  $\rho = \mu$ .

*Proof.* The optimality conditions of Theorem 2.9 remain valid except that  $x_j^l$  now appears in all the budget conditions  $\partial L/\partial \lambda_j = 0$ ,  $j \in V$ . Additional optimality conditions in each node are  $\lambda_j = -\eta_j$  (from  $\partial L/\partial x_j^l = 0$ ),  $x_j^l \geq 0$ ,  $\eta_j \geq 0$ , and complementarity  $x_j^l \eta_j = 0$ . The expressions above are obtained precisely as in the proof of Theorem 2.9 when slack variables  $x_j^l$  and derived quantities  $\theta_j, \bar{\theta}_0, s_0$  are included. This derivation also yields intermediate results

$$\rho_j(x_j, x_j^c) = \mu + \frac{\lambda_j}{\tilde{r}_j^c} \leq \mu, \quad j \in L, \quad \rho = \mu + \frac{\lambda_0}{r^c} \leq \mu;$$

cf. equation (A.3) and the reward condition after (A.4). The additional optimality conditions above, together with these identities, lead to the stated case distinctions if it is observed that  $\lambda_j$  has the same sign as  $r_2^c w_j - \mu - \theta_j$  for  $j \in L$ .  $\square$

*Remark.* Note that all the multipliers now have a natural interpretation. The reward multiplier  $\mu$  is the maximal scenario return and a *threshold value for surplus money*: there is surplus money in scenario  $j$  iff  $\rho_j = \mu$  and  $r_2^c w_j > \mu$ . The budget multipliers  $\lambda_j$  (up to a scaling factor) measure the difference between the desired return or the scenario returns and the threshold.

*Proof of Lemma 2.17.* Formula (2.2) for  $\rho_j$  (with  $\theta_j = 0$ ) gives

$$(\phi_j + 1)(\rho_j - d_j) = \phi_j d_j + \mu - (\phi_j + 1)d_j = \mu - d_j$$

and hence  $\rho_j - d_j \sim \mu - d_j \equiv \mu - r^c - \psi_j x_0$ . By Theorem 2.9,

$$x_0 = -\frac{\lambda_0}{r_2^c \tilde{r}_1^c} \frac{\bar{r}_0 - r_1^c}{\Sigma_0} = -\tilde{\rho} \frac{r^c - \mu \bar{r}_0 - r_1^c}{r_2^c \tilde{r}_1^c \Sigma_0}.$$

Therefore we have

$$\begin{aligned} \rho_j - d_j &\sim \mu - r^c + r_2^c (r_j - r_1^c) \tilde{\rho} \frac{r^c - \mu \bar{r}_0 - r_1^c}{r_2^c \tilde{r}_1^c \Sigma_0} \\ &= (\mu - r^c) \left( 1 - \frac{r_j - r_1^c}{\delta_0^c + 1} \frac{\bar{r}_0 - r_1^c}{\Sigma_0} \right). \end{aligned}$$

By Theorem 2.9,  $\mu - r^c = r^c(\rho - r^c)/(r^c - \tilde{\rho}) \sim \rho - r^c$ . The second factor equals

$$\begin{aligned} 1 - \frac{(r_j - r_1^c)(\bar{r}_0 - r_1^c)}{[(\bar{r}_0 - r_1^c)\Sigma_0^{-1}(\bar{r}_0 - r_1^c) + 1]\Sigma_0} &\sim (\bar{r}_0 - r_1^c)^2 + \Sigma_0 - (r_j - r_1^c)(\bar{r}_0 - r_1^c) \\ &= \Sigma_0 + \bar{r}_0^2 - (r_j + r_1^c)\bar{r}_0 + r_j r_1^c \\ &\sim \tilde{\Sigma}_0 - (r_j + r_1^c)\tilde{r}_0 + \tilde{p}_0 r_j r_1^c \\ &= \sum_{k \in L} \frac{p_k}{\delta_k^c + 1} [r_k^2 - (r_j + r_1^c)r_k + r_1^c r_j] \\ &= \sum_{k \neq j} \frac{p_k}{\delta_k^c + 1} (r_k - r_j)(r_k - r_1^c) = s_j. \end{aligned}$$

This completes the proof.  $\square$

**Acknowledgments.** This work results from ongoing cooperation with the Institute of Operations Research at the University of St. Gallen. It would have been impossible without several visits and many intensive discussions in the research group of Karl Frauendorfer. My sincerest thanks especially to him, Heiko Siede, and Detlef Steiner. I am also indebted to Nick Trefethen and to an anonymous referee whose comments and suggestions have improved the quality of the final paper.

#### REFERENCES

- [1] A. J. ALEXANDER AND W. F. SHARPE, *Fundamentals of Investment*, Prentice-Hall, Englewood Cliffs, NJ, 1989.
- [2] K. J. ARROW, *Essays in the Theory of Risk-Bearing*, Markham, Chicago, 1971.



- [3] P. ARTZNER, F. DELBAEN, J.-M. EBER, AND D. HEATH, *Thinking coherently*, RISK Magazine, 10 (1997), pp. 68–71.
- [4] P. ARTZNER, F. DELBAEN, J.-M. EBER, AND D. HEATH, *Coherent measures of risk*, Math. Finance, 9 (1999), pp. 203–228.
- [5] C. ATKINSON, S. R. PLISKA, AND P. WILMOTT, *Portfolio management with transaction costs*, Proc. Roy. Soc. London Ser. A, 453 (1997), pp. 551–562.
- [6] I. BAJEUX-BESNAINOU AND R. PORTAIT, *Dynamic asset allocation in a mean-variance framework*, Management Sci., 44 (1998), pp. S79–S95.
- [7] H. BAUER, *Probability Theory*, de Gruyter Studies in Mathematics 23, de Gruyter, Berlin, 1996. (Translated from the German by Robert B. Burckel.)
- [8] V. S. BAWA, *Optimal rules for ordering uncertain prospects*, J. Finan. Econ., 10 (1975), pp. 849–857.
- [9] V. S. BAWA, *Safety-first, stochastic dominance, and optimal portfolio choice*, J. Finan. Quantit. Anal., 13 (1978), pp. 255–271.
- [10] V. S. BAWA, *Stochastic dominance: A research bibliography*, Management Sci., 28 (1982), pp. 698–712.
- [11] V. S. BAWA AND E. LINDENBERG, *Capital market equilibrium in a mean-lower partial moment framework*, J. Finan. Econ., 3 (1977), pp. 189–200.
- [12] R. E. BELLMAN, *Dynamic Programming*, Princeton University Press, Princeton, NJ, 1957.
- [13] A. BELTRATTI, A. CONSIGLIO, AND S. A. ZENIOS, *Scenario modeling for the management of international bond portfolios*, Ann. Oper. Res., 85 (1999), pp. 227–247.
- [14] D. P. BERTSEKAS AND S. E. SHREVE, *Stochastic Optimal Control: The Discrete-Time Case*, Academic Press, New York, 1978.
- [15] M. J. BEST AND B. DING, *On the continuity of the minimum in parametric quadratic programs*, J. Optim. Theory Appl., 86 (1995), pp. 245–250.
- [16] M. J. BEST AND R. R. GRAUER, *Sensitivity analysis for mean-variance portfolio problems*, Management Sci., 37 (1991), pp. 980–989.
- [17] J. R. BIRGE, *Stochastic programming computations and applications*, INFORMS J. Comput., 9 (1997), pp. 111–133.
- [18] J. R. BIRGE AND F. LOUVEAUX, *Introduction to Stochastic Programming*, Springer-Verlag, New York, 1997.
- [19] F. BLACK AND M. S. SCHOLES, *The pricing of options and corporate liabilities*, J. Polit. Econ., 81 (1973), pp. 637–654.
- [20] T. BOLLERSLEV, *Generalized autoregressive conditional heteroskedasticity*, J. Econometrics, 31 (1986), pp. 307–327.
- [21] T. BOLLERSLEV, R. Y. CHOU, AND K. F. KRONER, *ARCH modeling in finance: A review of the theory and empirical evidence*, J. Econometrics, 52 (1992), pp. 5–59.
- [22] T. BOLLERSLEV, R. F. ENGLE, AND D. B. NELSON, *ARCH models*, in The Handbook of Econometrics 4, R. F. Engle and D. L. McFadden, eds., 1994, Elsevier, Amsterdam, the Netherlands, pp. 2959–3038.
- [23] T. BOLLERSLEV, R. F. ENGLE, AND J. M. WOOLDRIDGE, *A capital asset pricing model with time-varying covariances*, J. Polit. Econ., 96 (1988), pp. 116–131.
- [24] N. BOULEAU AND D. LAMBERTON, *Residual risks and hedging strategies in Markovian markets*, Stochastic Process. Appl., 33 (1989), pp. 131–150.
- [25] M. BROADIE, *Computing efficient frontiers using estimated parameters*, Ann. Oper. Res., 45 (1993), pp. 21–58.
- [26] I. R. C. BUCKLEY AND R. KORN, *Optimal index tracking under transaction costs and impulse control*, Int. J. Theor. Appl. Finance, 1 (1998), pp. 315–330.
- [27] D. R. CARIÑO, T. KENT, D. H. MYERS, C. STACY, M. SYLVANUS, A. L. TURNER, K. WATANABE, AND W. T. ZIEMBA, *The Russell–Yasuda Kasai model: An asset/liability model for a Japanese insurance company using multistage stochastic programming*, Interfaces, 24 (1994), pp. 29–49.
- [28] G. CHAMBERLAIN, *A characterization of the distributions that imply mean-variance utility functions*, J. Econ. Theory, 29 (1983), pp. 185–201.
- [29] A. H. Y. CHEN, F. C. JEN, AND S. ZIOMTS, *The optimal portfolio revision policy*, J. Business, 44 (1972), pp. 51–61.
- [30] V. K. CHOPRA AND W. T. ZIEMBA, *The effects of errors in means, variances, and covariances on optimal portfolio choice*, J. Portfolio Management, 20 (1993), pp. 6–11.
- [31] G. CONSIGLI AND M. A. H. DEMPSTER, *Solving dynamic portfolio problems using stochastic programming*, Z. Angew. Math. Mech., 77 (1997), pp. S535–S536.
- [32] J. COX AND C. F. HUANG, *Optimal consumption and portfolio policies when asset prices follow a diffusion process*, J. Econ. Theory, 49 (1989), pp. 33–83.

- [33] J. COX, J. E. INGERSOLL, JR., AND S. A. ROSS, *An intertemporal general equilibrium model of asset prices*, *Econometrica*, 53 (1985), pp. 363–384.
- [34] J. COX, J. E. INGERSOLL, JR., AND S. A. ROSS, *A theory of the term structure of interest rates*, *Econometrica*, 53 (1985), pp. 385–467.
- [35] J. CVITANIĆ AND I. KARATZAS, *Convex duality in constrained portfolio optimization*, *Ann. Appl. Probab.*, 2 (1992), pp. 767–818.
- [36] J. CVITANIĆ AND I. KARATZAS, *Hedging and portfolio optimization under transaction costs: A martingale approach*, *Math. Finance*, 6 (1996), pp. 133–165.
- [37] G. B. DANTZIG AND G. INFANGER, *Multi-stage stochastic linear programs for portfolio optimization*, *Ann. Oper. Res.*, 45 (1993), pp. 59–76.
- [38] F. DELBAEN AND W. SCHACHERMAYER, *The existence of absolutely continuous local martingale measures*, *Ann. Appl. Probab.*, 5 (1995), pp. 926–945.
- [39] R. S. DEMBO, *Scenario optimization*, *Ann. Oper. Res.*, 30 (1991), pp. 63–80.
- [40] R. S. DEMBO AND D. ROSEN, *The practice of portfolio replication. A practical overview of forward and inverse problems*, *Ann. Oper. Res.*, 85 (1999), pp. 267–284.
- [41] M. A. H. DEMPSTER, ED., *Stochastic Programming*, Academic Press, London, 1980.
- [42] M. A. H. DEMPSTER, ED., *On stochastic programming: II. Dynamic problems under risk*, *Stochastics*, 25 (1988), pp. 15–42.
- [43] T. DOHI AND S. OSAKI, *A note on portfolio optimization with path-dependent utility*, *Ann. Oper. Res.*, 45 (1993), pp. 77–90.
- [44] D. DUFFIE, *Dynamic Asset Pricing Theory*, Princeton University Press, Princeton, NJ, 1992.
- [45] D. DUFFIE AND H. R. RICHARDSON, *Mean-variance hedging in continuous time*, *Ann. Appl. Probab.*, 1 (1991), pp. 1–15.
- [46] G. T. DUNCAN, *A matrix measure of multivariate local risk aversion*, *Econometrica*, 45 (1977), pp. 895–903.
- [47] C. EDIRISINGHE, V. NAIK, AND R. UPPAL, *Optimal replication of options with transaction costs and trading restrictions*, *J. Finan. Quantit. Anal.*, 28 (1993), pp. 117–138.
- [48] N. EL KAROUI AND M.-C. QUENEZ, *Dynamic programming and pricing of contingent claims in an incomplete market*, *SIAM J. Control Optim.*, 33 (1995), pp. 29–66.
- [49] R. J. ELLIOTT AND J. VAN DER HOEK, *An application of hidden Markov models to asset allocation problems*, *Fin. Stoch.*, 1 (1997), pp. 229–238.
- [50] E. J. ELTON AND M. J. GRUBER, *Modern Portfolio Theory and Investment Analysis*, 3rd ed., Wiley, New York, 1987.
- [51] P. EMBRECHTS, C. KLÜPPELBERG, AND T. MIKOSH, *Extremal Events in Finance and Insurance*, Springer-Verlag, New York, 1997.
- [52] P. EMBRECHTS, S. RESNICK, AND G. SAMORODNITSKY, *Extreme value theory as a risk management tool*, *North Amer. Actuarial J.*, 26 (1999), pp. 30–41.
- [53] R. F. ENGLE, *Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation*, *Econometrica*, 50 (1982), pp. 987–1007.
- [54] R. F. ENGLE AND K. F. KRONER, *Multivariate simultaneous generalized ARCH*, *Econometric Theory*, 11 (1995), pp. 122–150.
- [55] L. G. EPSTEIN, *Decreasing risk aversion and mean-variance analysis*, *Econometrica*, 53 (1985), pp. 945–961.
- [56] Y. ERMOLIEV AND R. J.-B. WETS, EDs., *Numerical Techniques for Stochastic Optimization*, Springer-Verlag, New York, 1988.
- [57] E. F. FAMA, *Multiperiod consumption-investment decisions*, *Amer. Econ. Rev.*, 60 (1970), pp. 163–174.
- [58] P. C. FISHBURN, *Mean-risk analysis with risk associated with below-target returns*, *Amer. Econ. Rev.*, 67 (1977), pp. 116–126.
- [59] P. C. FISHBURN AND R. B. PORTER, *Optimal portfolios with one safe and one risky asset: Effects of changes in rate of return and risk*, *Management Sci.*, 22 (1976), pp. 1064–1073.
- [60] H. FÖLLMER AND P. LEUKERT, *Efficient hedges: Cost versus shortfall risk*, *Fin. Stoch.*, 4 (2000), pp. 117–146.
- [61] H. FÖLLMER AND M. SCHWEIZER, *Hedging by sequential regression: An introduction to the mathematics of option trading*, *ASTIN Bull.*, 18 (1989), pp. 147–160.
- [62] H. FÖLLMER AND D. SONDERMANN, *Hedging of non-redundant contingent claims*, in *Contributions to Mathematical Economics*, W. Hildenbrand and A. Mas-Colell, eds., North-Holland, Amsterdam, 1986, pp. 203–223.
- [63] K. FRAUENDORFER, *The stochastic programming extension of the Markowitz approach*, *Int. J. Neural Mass-Parallel Comput. Inform. Syst.*, 5 (1995), pp. 449–460.
- [64] K. FRAUENDORFER, *Barycentric scenario trees in convex multistage stochastic programming*, *Math. Programming*, 75 (1996), pp. 277–293.

- [65] K. FRAUENDORFER AND H. SIEDE, *Portfolio selection using multi-stage stochastic programming*, Central European J. Oper. Res., 7 (2000), pp. 277–290.
- [66] R. GAESE, *Risk Management and Risk Measurement*, Dissertation Nr. 2195, Difo-Druck OHG, Bamberg, 1999.
- [67] F. GLOVER AND C. K. JONES, *A stochastic generalized network model and large-scale mean-variance algorithm for portfolio selection*, J. Inform. Optim. Sci., 9 (1988), pp. 299–316.
- [68] J. GONDZIO AND R. KOUWENBERG, *High Performance Computing for Asset Liability Management*, Preprint MS-99-004, Dept. of Mathematics and Statistics, University of Edinburgh, UK, 1999. Oper. Res., to appear.
- [69] J. B. GUERARD, JR., M. TAKANO, AND Y. YAMANE, *The development of efficient portfolios in Japan with particular emphasis on sales and earnings forecasting*, Ann. Oper. Res., 45 (1993), pp. 91–108.
- [70] N. H. HAKANSSON, *Optimal investment and consumption strategies under risk for a class of utility functions*, Econometrica, 38 (1970), pp. 401–416.
- [71] N. H. HAKANSSON, *Capital growth and the mean-variance approach to portfolio selection*, J. Finan. Quantit. Anal., 6 (1971), pp. 517–557.
- [72] N. H. HAKANSSON, *Multi-period mean-variance analysis: Toward a general theory of portfolio choice*, J. Finance, 26 (1971), pp. 857–884.
- [73] N. H. HAKANSSON, *On optimal myopic portfolio policies, with and without correlation of yields*, J. Business, 44 (1971), pp. 324–334.
- [74] N. H. HAKANSSON, *Convergence to isoelastic utility and policy in multi-period portfolio choice*, J. Finan. Econ., 1 (1974), p. 201.
- [75] G. HANOCH AND H. LEVY, *The efficiency analysis of choices involving risk*, Rev. Econ. Stud., 36 (1969), pp. 335–346.
- [76] W. V. HARLOW AND R. K. S. RAO, *Asset pricing in a generalized mean-lower partial moment framework: Theory and evidence*, J. Finan. Quantit. Anal., 24 (1989), pp. 285–311.
- [77] J. M. HARRISON AND D. M. KREPS, *Martingales and arbitrage in multiperiod securities markets*, J. Econ. Theory, 20 (1979), pp. 381–408.
- [78] J. M. HARRISON AND S. R. PLISKA, *Martingales and stochastic integrals in the theory of continuous trading*, Stochastic Process. Appl., 11 (1981), pp. 215–260.
- [79] J. M. HARRISON AND S. R. PLISKA, *A stochastic calculus model of continuous trading: Complete markets*, Stochastic Process. Appl., 15 (1983), pp. 313–316.
- [80] A. HARVEY, E. RUIZ, AND N. SHEPHARD, *Multivariate stochastic variance models*, Rev. Econ. Stud., 61 (1994), pp. 247–264.
- [81] H. HE AND N. D. PEARSON, *Consumption and portfolio policies with incomplete markets and short-sale constraints: The finite dimensional case*, Math. Finance, 1 (1991), pp. 1–10.
- [82] H. HE AND N. D. PEARSON, *Consumption and portfolio policies with incomplete markets and short-sale constraints: The infinite dimensional case*, J. Econ. Theory, 54 (1991), pp. 259–304.
- [83] D. HEATH, S. OREY, V. PESTIEN, AND W. SUDDERTH, *Minimizing or maximizing the expected time to reach zero*, SIAM J. Control Optim., 25 (1987), pp. 195–205.
- [84] D. HEATH AND M. SCHWEIZER, *Martingales versus PDEs in finance: An equivalence result with examples*, J. Appl. Probab., to appear.
- [85] T. S. Y. HO AND S.-B. LEE, *Term structure movements and pricing of interest claims*, J. Finance, 41 (1986), pp. 1011–1029.
- [86] N. HOFMANN, E. PLATEN, AND M. SCHWEIZER, *Option pricing under incompleteness and stochastic volatility*, Math. Finance, 2 (1992), pp. 153–187.
- [87] G. HUBERMANN AND S. ROSS, *Portfolio turnpike theorems, risk aversion, and regularly varying utility functions*, Econometrica, 51 (1983), pp. 1345–1361.
- [88] J. C. HULL AND A. WHITE, *Pricing interest-rate-derivative securities*, Rev. Fin. Studies, 3 (1990), pp. 573–592.
- [89] J. E. INGERSOLL, JR., *Theory of Financial Decision Making*, Rowman & Littlefield, Lanham, MD, 1987.
- [90] S. D. JACKA, *A martingale representation result and an application to incomplete financial markets*, Math. Finance, 2 (1992), pp. 239–250.
- [91] I. JEWITT, *Risk aversion and the choice between risky prospects: The preservation of comparative statistics results*, Rev. Econ. Stud., 54 (1987), pp. 73–85.
- [92] I. JEWITT, *Choosing between risky prospects: The characterization of comparative statistics results, and location independent risk*, Management Sci., 35 (1989), pp. 60–70.
- [93] J. D. JOBSON, *Confidence regions for the mean-variance efficient set: An alternative approach to estimation risk*, Rev. Quantit. Finan. Account., 1 (1991), pp. 235–257.
- [94] P. KALL, *Stochastic Linear Programming*, Springer-Verlag, New York, 1976.

- [95] P. KALL AND S. W. WALLACE, *Stochastic Programming*, Wiley, New York, 1994.
- [96] K. G. KALLBERG AND W. T. ZIEMBA, *Comparison of alternative utility functions in portfolio selection problems*, *Management Sci.*, 29 (1983), pp. 1257–1276.
- [97] I. KARATZAS, *Optimization problems in the theory of continuous trading*, *SIAM J. Control Optim.*, 27 (1989), pp. 1221–1259.
- [98] I. KARATZAS, J. P. LEHOCZKY, AND S. E. SHREVE, *Optimal portfolio and consumption decisions for a “small investor” on a finite horizon*, *SIAM J. Control Optim.*, 25 (1987), pp. 1557–1586.
- [99] I. KARATZAS, J. P. LEHOCZKY, S. E. SHREVE, AND G.-L. XU, *Martingale and duality methods for utility maximization in an incomplete market*, *SIAM J. Control Optim.*, 29 (1991), pp. 702–730.
- [100] R. E. KIHLLSTROM AND L. J. MIRMAN, *On risk aversion with many commodities*, *J. Econ. Theory*, 8 (1974), pp. 361–388.
- [101] R. E. KIHLLSTROM AND L. J. MIRMAN, *Constant, increasing and decreasing risk aversion with many commodities*, *Rev. Econ. Stud.*, 48 (1981), pp. 271–280.
- [102] M. KIJIMA, *The generalized harmonic mean and a portfolio problem with dependent assets*, *Theory and Decision*, 43 (1997), pp. 71–87.
- [103] M. KIJIMA AND M. OHNISHI, *Mean-risk analysis of risk aversion and wealth effects on optimal portfolios with multiple investment opportunities*, *Ann. Oper. Res.*, 45 (1993), pp. 147–163.
- [104] M. KIJIMA AND M. OHNISHI, *Portfolio selection problems via the bivariate characterization of stochastic dominance relations*, *Math. Finance*, 6 (1996), pp. 237–277.
- [105] A. J. KING, *Asymmetric risk measures and tracking models for portfolio optimization under uncertainty*, *Ann. Oper. Res.*, 45 (1993), pp. 165–177.
- [106] A. J. KING AND D. L. JENSEN, *Linear-quadratic efficient frontiers for portfolio optimization*, *Appl. Stochastic Models Data Anal.*, 8 (1992), pp. 195–207.
- [107] D. KIRA AND W. T. ZIEMBA, *The demand for a risky set*, *Management Sci.*, 26 (1980), pp. 1158–1165.
- [108] H. KONNO, *Piecewise linear risk function and portfolio optimization*, *J. Oper. Res. Soc. Japan*, 33 (1990), pp. 139–156.
- [109] H. KONNO, *The relation between investor’s greediness and the asset price in the mean-variance market*, *J. Oper. Res. Soc. Japan*, 40 (1997), pp. 579–589.
- [110] H. KONNO, S. R. PLISKA, AND K.-I. SUZUKI, *Optimal portfolios with asymptotic criteria*, *Ann. Oper. Res.*, 45 (1993), pp. 187–204.
- [111] H. KONNO AND H. SHIRAKAWA, *Existence of a nonnegative equilibrium price vector in the mean-variance capital market*, *Math. Finance*, 5 (1995), pp. 233–246.
- [112] H. KONNO, H. SHIRAKAWA, AND H. YAMAZAKI, *A mean-absolute deviation-skewness portfolio optimization model*, *Ann. Oper. Res.*, 45 (1993), pp. 205–220.
- [113] H. KONNO AND K.-I. SUZUKI, *A mean-variance-skewness portfolio optimization model*, *J. Oper. Res. Soc. Japan*, 38 (1995), pp. 173–187.
- [114] H. KONNO AND H. WATANABE, *Bond portfolio optimization problems and their applications to index tracking: A partial optimization approach*, *J. Oper. Res. Soc. Japan*, 39 (1996), pp. 295–306.
- [115] H. KONNO AND H. YAMAZAKI, *Mean-absolute deviation portfolio optimization model and its application to Tokyo Stock Exchange*, *Management Sci.*, 37 (1991), pp. 519–531.
- [116] D. O. KRAMKOV, *Optional decomposition of supermartingales and hedging contingent claims in incomplete security markets*, *Probab. Theory Related Fields*, 105 (1996), pp. 459–479.
- [117] A. KRAUS AND R. H. LITZENBERGER, *Skewness preference and the valuation of risky assets*, *J. Finance*, 31 (1976), pp. 1085–1094.
- [118] Y. KROLL, M. LESHNO, H. LEVY, AND Y. SPECTOR, *Increasing risk, decreasing absolute risk aversion and diversification*, *J. Math. Econom.*, 24 (1995), pp. 537–556.
- [119] Y. KROLL, H. LEVY, AND H. M. MARKOWITZ, *Mean-variance versus direct utility maximization*, *J. Finance*, 39 (1984), pp. 47–61.
- [120] D. LAMBERTON, H. PHAM, AND M. SCHWEIZER, *Local risk-minimization under transaction costs*, *Math. Oper. Res.*, 23 (1998), pp. 585–612.
- [121] H. LEVY, *Stochastic dominance and expected utility: Survey and analysis*, *Management Sci.*, 38 (1992), pp. 555–593.
- [122] H. LEVY, *Risk and return: An experimental analysis*, *Int. Econ. Rev.*, 38 (1997), pp. 119–149.
- [123] H. LEVY AND H. M. MARKOWITZ, *Approximating expected utility by a function of mean and variance*, *Amer. Econ. Rev.*, 69 (1979), pp. 308–317.
- [124] H. LEVY AND P. A. SAMUELSON, *The capital asset pricing model with diverse holding periods*, *Management Sci.*, 38 (1992), pp. 1529–1542.

- [125] H. LEVY AND Z. WIENER, *Stochastic dominance and prospect dominance with subjective weighting functions*, J. Risk Uncertain., 16 (1998), pp. 147–163.
- [126] Y. LI AND W. T. ZIEMBA, *Characterization of optimal portfolios by univariate and multivariate measures of risk aversion*, Management Sci., 35 (1989), pp. 259–269.
- [127] Y. LI AND W. T. ZIEMBA, *Univariate and multivariate measures of risk aversion and risk premiums*, Ann. Oper. Res., 45 (1993), pp. 265–296.
- [128] J. LINTNER, *The valuation of risky assets and the selection of risky investments in stock portfolios and capital budgets*, Rev. Econ. Statist., 47 (1965), pp. 13–37.
- [129] J. LINTNER, *The market price of risk, size of market, and investor's risk aversion*, Rev. Econ. Statist., (1970).
- [130] J. B. LONG, JR., *The numeraire portfolio*, J. Finan. Econ., 26 (1990), pp. 29–69.
- [131] D. G. LUENBERGER, *A preference foundation for log mean-variance criteria in portfolio choice problems*, J. Econ. Dyn. Control, 17 (1993), pp. 887–906.
- [132] M. J. MACHINA, *Temporal risk and the nature of induced preferences*, J. Econ. Theory, 33 (1984), pp. 199–231.
- [133] L. C. MACLEAN AND K. WELDON, *Estimating multivariate random effects without replication*, Comm. Statist. Theory Methods, 25 (1996), pp. 1447–1469.
- [134] H. M. MARKOWITZ, *Portfolio selection*, J. Finance, 7 (1952), pp. 77–91.
- [135] H. M. MARKOWITZ, *The utility of wealth*, J. Polit. Econom., 60 (1952), pp. 151–158.
- [136] H. M. MARKOWITZ, *The optimization of a quadratic function subject to linear constraints*, Naval Res. Logist. Quarterly, 3 (1956), pp. 111–133.
- [137] H. M. MARKOWITZ, *Portfolio Selection: Efficient Diversification of Investments*, John Wiley, New York, 1959.
- [138] H. M. MARKOWITZ, *Mean-Variance Analysis in Portfolio Choice and Capital Markets*, Basil Blackwell, Oxford, UK, 1987.
- [139] H. M. MARKOWITZ, P. TODD, G. XU, AND Y. YAMANE, *Computation of mean-semivariance efficient sets by the Critical Line Algorithm*, Ann. Oper. Res., 45 (1993), pp. 307–317.
- [140] P. L. MCENTIRE, *Portfolio theory for independent assets*, Management Sci., 30 (1984), pp. 952–963.
- [141] R. C. MERTON, *Lifetime portfolio selection under uncertainty: The continuous-time case*, Rev. Econ. Statist., 51 (1969), pp. 247–257.
- [142] R. C. MERTON, *Optimal consumption and portfolio rules in a continuous-time model*, J. Econ. Theory, 3 (1971), pp. 373–413.
- [143] R. C. MERTON, *An intertemporal capital asset pricing model*, Econometrica, 41 (1973), pp. 867–887.
- [144] R. C. MERTON, *Theory of rational option pricing*, Bell J. Econ. Manag. Sci., 4 (1973), pp. 141–183.
- [145] R. C. MERTON AND P. A. SAMUELSON, *Fallacy of the log-normal approximation to optimal portfolio decision-making over many periods*, J. Finan. Econ., 1 (1974), pp. 67–94.
- [146] R. C. MERTON AND P. A. SAMUELSON, *Continuous Time Finance*, Basil Blackwell, Oxford, UK, 1990.
- [147] P. MONAT AND C. STRICKER, *Föllmer–Schweizer decomposition and mean-variance hedging of general claims*, Ann. Probab., 23 (1995), pp. 605–628.
- [148] A. J. MORTON AND S. R. PLISKA, *Optimal portfolio management with fixed transaction costs*, Math. Finance, 5 (1995), pp. 337–356.
- [149] J. MOSSIN, *Equilibrium in capital assets markets*, Econometrica, 34 (1966), pp. 261–276.
- [150] J. MOSSIN, *Optimal multiperiod portfolio policies*, J. Business, 41 (1968), pp. 215–229.
- [151] J. M. MULVEY, *Nonlinear network models in finance*, Adv. Math. Progr. Fin. Planning, 1 (1987), p. 253.
- [152] J. M. MULVEY AND H. VLADIMIROU, *Stochastic network optimization models for investment planning*, Ann. Oper. Res., 20 (1989), pp. 187–217.
- [153] M. NAKASATO AND K. FURUKAWA, *On the number of securities which constitute an efficient portfolio*, Ann. Oper. Res., 45 (1993), pp. 333–347.
- [154] *The Sveriges Riksbank (Bank of Sweden) Prize in Economic Sciences in Memory of Alfred Nobel*. Press Release of The Royal Swedish Academy of Sciences, Oct. 16, 1990, <http://www.nobel.se/economics/laureates/1990/press.html>.
- [155] W. OGRYCZAK AND A. RUSZCZYŃSKI, *From stochastic dominance to mean-risk models: Semideviations as risk measures*, European J. Oper. Res., 116 (1999), pp. 33–50.
- [156] A. F. PEROLD, *Large-scale portfolio optimization*, Management Sci., 30 (1984), pp. 1143–1160.
- [157] H. PHAM, T. RHEINLÄNDER, AND M. SCHWEIZER, *Mean-variance hedging for continuous processes: New results and examples*, Fin. Stoch., 2 (1998), pp. 173–198.

- [158] E. PHELPS, *The accumulation of risky capital: A sequential utility analysis*, *Econometrica*, 30 (1962), pp. 729–743.
- [159] G. A. POGUE, *An extension of the Markowitz portfolio selection model to include variable transactions' costs, short sales, leverage policies and taxes*, *J. Finance*, 25 (1970), pp. 1005–1027.
- [160] J. W. PRATT, *Risk aversion in the small and in the large*, *Econometrica*, 32 (1964), pp. 122–136.
- [161] J. W. PRATT AND R. J. ZECKHAUSER, *Proper risk aversion*, *Econometrica*, 55 (1987), pp. 143–154.
- [162] R. RABINOVICH, *Pricing stock and bond options when the default-free rate is stochastic*, *J. Finan. Quantit. Anal.*, 24 (1989), pp. 447–457.
- [163] H. R. RICHARDSON, *A minimum variance result in continuous trading portfolio optimization*, *Management Sci.*, 35 (1989), pp. 1045–1055.
- [164] R. T. ROCKAFELLAR, *Convex Analysis*, Princeton University Press, Princeton, NJ, 1969.
- [165] R. T. ROCKAFELLAR, *Duality and optimality in multistage stochastic programming*, *Ann. Oper. Res.*, 85 (1999), pp. 1–19.
- [166] R. T. ROCKAFELLAR AND S. URYASEV, *Optimization of Conditional Value-at-Risk*, *J. Risk*, 2 (2000), pp. 21–41.
- [167] R. T. ROCKAFELLAR AND R. J.-B. WETS, *Generalized linear-quadratic problems of deterministic and stochastic optimal control in discrete time*, *SIAM J. Control Optim.*, 28 (1990), pp. 810–822.
- [168] R. ROLL, *Evidence on the "growth-optimum" model*, *J. Finance*, 28 (1973), pp. 551–566.
- [169] S. ROSS, *Some stronger measures of risk aversion in the small and in the large*, *Econometrica*, 49 (1981), pp. 621–638.
- [170] M. E. RUBINSTEIN, *A comparative statistics analysis of risk premiums*, *J. Business*, 46 (1973), pp. 605–615.
- [171] A. RUSZCZYŃSKI, *Decomposition methods in stochastic programming*, *Math. Programming*, 79 (1997), pp. 333–353.
- [172] P. A. SAMUELSON, *Lifetime portfolio selection by dynamic stochastic programming*, *Rev. Econ. Stud.*, 51 (1969), pp. 239–246.
- [173] P. A. SAMUELSON, *The fundamental approximation theorem of portfolio analysis in terms of means, variances, and higher moments*, *Rev. Econ. Stud.*, 37 (1970), pp. 537–542.
- [174] M. SCHÄL, *On quadratic cost criteria for option hedging*, *Math. Oper. Res.*, 19 (1994), pp. 121–131.
- [175] M. SCHWEIZER, *Option hedging for semimartingales*, *Stochastic Process. Appl.*, 37 (1991), pp. 339–363.
- [176] M. SCHWEIZER, *Mean-variance hedging for general claims*, *Ann. Appl. Probab.*, 2 (1992), pp. 171–179.
- [177] M. SCHWEIZER, *Approximating random variables by stochastic integrals*, *Ann. Probab.*, 22 (1994), pp. 1536–1575.
- [178] M. SCHWEIZER, *Risk-minimizing hedging strategies under restricted information*, *Math. Finance*, 4 (1994), pp. 327–342.
- [179] M. SCHWEIZER, *On the minimal martingale measure and the Föllmer–Schweizer decomposition*, *Stochastic Anal. Appl.*, 13 (1995), pp. 573–599.
- [180] M. SCHWEIZER, *Variance-optimal hedging in discrete time*, *Math. Oper. Res.*, 20 (1995), pp. 1–32.
- [181] M. SCHWEIZER, *Approximation pricing and the variance-optimal martingale measure*, *Ann. Probab.*, 24 (1996), pp. 206–236.
- [182] M. SCHWEIZER, *A Guided Tour through Quadratic Hedging Approaches*, Preprint, Technical University of Berlin, 1999.
- [183] J. K. SENGUPTA, *Optimal portfolio investment in a dynamic horizon*, *Int. J. Syst. Sci.*, 14 (1983), pp. 789–800.
- [184] W. F. SHARPE, *Capital asset prices: A theory of market equilibrium under conditions of risk*, *J. Finance*, 19 (1964), pp. 425–442.
- [185] W. F. SHARPE, *Portfolio Theory and Capital Market*, McGraw-Hill, New York, 1970.
- [186] W. F. SHARPE, *Investment*, Prentice-Hall, Englewood Cliffs, NJ, 1978.
- [187] J. SHERMAN AND W. J. MORRISON, *Adjustment of an inverse matrix corresponding to a change in one element of a given matrix*, *Ann. Math. Statist.*, 21 (1950), pp. 124–127.
- [188] H. SHIRAKAWA, *Optimal consumption and portfolio selection with incomplete markets and upper and lower bound constraints*, *Math. Finance*, 4 (1994), pp. 1–24.
- [189] H. SHIRAKAWA AND H. KASSAI, *Optimal consumption and arbitrage in incomplete, finite state security markets*, *Ann. Oper. Res.*, 45 (1993), pp. 349–372.

- [190] H. M. SONER, S. E. SHREVE, AND J. CVITANIĆ, *There is no nontrivial hedging portfolio for option pricing with transaction costs*, Ann. Appl. Probab., 5 (1995), pp. 327–355.
- [191] M. C. STEINBACH, *Fast Recursive SQP Methods for Large-Scale Optimal Control Problems*, dissertation, University of Heidelberg, Germany, 1995.
- [192] M. C. STEINBACH, *Structured interior point SQP methods in optimal control*, Z. Angew. Math. Mech., 76 (1996), pp. S59–S62.
- [193] M. C. STEINBACH, *Recursive Direct Optimization and Successive Refinement in Multistage Stochastic Programs*, Preprint SC-98-27, Konrad-Zuse-Zentrum für Informationstechnik Berlin, 1998.
- [194] M. C. STEINBACH, *Recursive direct algorithms for multistage stochastic programs in financial engineering*, in Operations Research Proceedings 1998, P. Kall and H.-J. Lüthi, eds., Springer-Verlag, New York, 1999, pp. 241–250.
- [195] M. C. STEINBACH, *Hierarchical sparsity in multistage convex stochastic programs*, in Stochastic Optimization: Algorithms and Applications, S. Uryasev and P. M. Pardalos, eds., Kluwer Academic Publishers, Dordrecht, the Netherlands, 2000, pp. 363–388.
- [196] M. C. STEINBACH, H. G. BOCK, G. V. KOSTIN, AND R. W. LONGMAN, *Mathematical optimization in robotics: Towards automated high speed motion planning*, Surveys Math. Indust., 7 (1998), pp. 303–340.
- [197] G. V. G. STEVENS, *On Tobin's multiperiod portfolio theorem*, Rev. Econ. Stud., 39 (1972), pp. 461–468.
- [198] J. TOBIN, *Liquidity preference as behavior towards risk*, Rev. Econ. Stud., 25 (1958), pp. 65–86.
- [199] J. TOBIN, *The theory of portfolio selection*, in The Theory of Interest Rates, F. H. Hahn and F. P. R. Brechling, eds., Macmillan, London, 1965, pp. 3–51.
- [200] S. URYASEV, *Conditional Value-at-Risk: Optimization algorithms and applications*, Financial Engineering News, (14) (2000), pp. 1–6.
- [201] O. VASICEK, *An equilibrium characterization of the term structure*, J. Finan. Econ., 5 (1977), pp. 177–188.
- [202] J. VON NEUMANN AND O. MORGENSTERN, *Theory of Games and Economic Behavior*, 3rd ed., Princeton University Press, Princeton, NJ, 1953.
- [203] R. J.-B. WETS, *Programming under uncertainty: The equivalent convex program*, SIAM J. Appl. Math., 14 (1966), pp. 89–105.
- [204] M. A. WOODBURY, *Inverting Modified Matrices*, Memorandum Report 42, Statistical Research Group, Princeton University, Princeton, NJ, 1950.
- [205] S. A. ZENIOS, *Financial Optimization*, Cambridge University Press, Cambridge, UK, 1992.
- [206] S. A. ZENIOS AND P. KANG, *Mean-absolute deviation portfolio optimization for mortgage-backed securities*, Ann. Oper. Res., 45 (1993), pp. 433–450.
- [207] W. T. ZIEMBA, *Choosing investments when the returns have stable distributions*, in Mathematical Programming in Theory and Practice, P. L. Hammer and G. Zoutendijk, eds., North-Holland, Amsterdam, 1974, pp. 443–482.
- [208] W. T. ZIEMBA AND J. M. MULVEY, EDS., *Worldwide Asset and Liability Modeling*, Cambridge University Press, Cambridge, UK, 1998.