Portfolio optimization under a minimax rule revisited

Kaiwen Meng[1] · **Hongyu Yang**[2] · **Xiaoqi Yang**[3] · **Carisa Kwok Wai Yu**[4]

Abstract: In this paper, we revisit the bi-criteria portfolio optimization model where the short selling is permitted, and a trade-off is sought between the expected return rate of a portfolio and the maximum of the uncertainty measured by a general deviation measure for all the investments comprising a portfolio. We solve this bi-criteria model by first converting it into a collection of weighted sum piecewise linear convex programs, and then analyzing their optimality conditions. We not only provide explicit analytical formulas for all the efficient portfolios, but also explore as a whole the set of all the efficient portfolios and its structure such as dimensionality and distribution. We generalize the classical Two-fund Theorem by providing some collections of finitely many efficient portfolios to generate or estimate the set of all the efficient portfolios. We also notice that our efficient portfolios are almost the risk parity ones in the sense that the risks are allocated equally across the investments. Moreover, we illustrate the reliability of our model by carrying out Monte Carlo simulations to test the performance of some efficient portfolios versus inefficient ones.

Keywords: deviation measure · bi-criteria optimization · portfolio selection · risk parity

1 Introduction

Portfolio optimization is of both theoretical and practical interest. In portfolio optimization, the total return rate $\sum_{j=1}^{n} x_j \mathbb{R}_j$ of a portfolio x is normally modeled as a random variable, where x_j is the allocation for investment to the j-th asset whose return rate $\mathbb{R}_j$ is a random variable. The foundation for this line of research was laid by Markowitz with his mean-variance model [6] . This model can be formulated as a bi-criteria optimization problem where a trade-off is sought between expected return and investment risk represented by variance, or equivalently by standard deviation. An analytic derivation of the mean-variance efficient frontier as well as the Two-fund Theorem can be find in Merton [26].

Since then, as summarized in Sawik [19,20], this basic model have been extended or modified from three aspects. The first path is to simplify the type and amount of input data (Bertsimas and Pachamanova [27]). The second direction concentrates on the introduction of an alternative measure of risk (Markowitz [7], Konno [8]). Finally, the third relates to the inclusion of other criteria and/or limitations (Qi et al. [28], Kizys et al. [18]).

Note that variance (or equivalently the standard deviation) is a debatable measure of risk. For

[1]School of Economic Mathematics, Southwestern University of Finance and Economics, Chengdu 611130, China
[2]School of Economic Mathematics, Southwestern University of Finance and Economics, Chengdu 611130, China
[3]Department of Applied Mathematics, The Hong Kong Polytechnic University, Kowloon, Hong Kong
[4]Department of Mathematics, Statistics and Insurance, The Hang Seng University of Hong Kong, Siu Lek Yuen, Shatin, Hong Kong

example, it is a symmetric measure that treats positive and negative deviations from the mean in the same way. But for investors, they are totally different. One brings the excess return of investment, while the other causes the investment risk. Several alternative risk measures have been proposed to better capture investors risk perception, including semivariance (Markowitz [7], Ogryczak and Ruszczynski [21]), the absolute deviation (Konno [8] and Konno and Yamazaki [9]), Conditional Value-at-Risk (Rockafellar and Uryasev [11]). Meanwhile, after capturing some commonalities of the existing risk measures, some axiomatic risk measures are proposed, including coherent risk measures (Artzner et al. [22]), reward measures (De Giorgi [23]), general deviation measures (Rockafellar et al. [12]). It is noted that, most risk measures mentioned above can be included in the framework of general deviation measures.

Konno and Yamazaki [9] noted that the derivation of the covariance matrix in mean-variance model can be cumbersome and there are computational limitations in attempting to solve quadratic model in practice. They suggest employing linear objectives to alleviate these computational limitations, such as the work done by [4], [10] and [24]. In particular, Cai et al. proposed in [4] a different way of dealing with the uncertainty by first measuring the uncertainty in $x_j\mathbb{R}_j$ for all j by the absolute deviation measure, and then taking the maximum of all the individual uncertainty as the uncertainty associated with the portfolio x. Some observations can be obtained from [4]: all the efficient portfolios can be analytically derived, and they are not sensitive with respect to the expectations of the $\mathbb{R}_j$'s.

Despite many benefits from the portfolio optimization model in [4], it should be pointed out that the short selling is not allowed for the model, and that the set of all efficient portfolios along with its structure has not been explored. This motivates us to revisit such a model with some new features. Firstly, the short selling is now permitted (i.e., the allocation x_j can be negative). Secondly, the uncertainty in a portfolio x is defined as the maximum of all the individual measures $\mathcal{D}(x_j\mathbb{R}_j)$, where $\mathcal{D}$ could be any proper deviation measure [12, Definition 1], not necessarily being the absolute deviation adopted in [4]. In particular, deviation measures which are not symmetric with respect to ups and downs can be adopted for our model to deal with the downside/upside uncertainty in a proper way, see Remark 2.1 for more details. Finally, in order to test the performance of the efficient portfolios versus the inefficient ones, we will carry out Monte Carlo simulations based on the assumptions that the return rate $\mathbb{R}_j$ is an essentially bounded random variable, and that its realization value falls into some interval based on the estimated expected return. In this paper, we will not only provide explicit analytical formulas for all the efficient portfolios, but also explore as a whole the set of all the efficient portfolios and its structure such as dimensionality and distribution. We will generalize the classical Two-fund Theorem by providing some collections of finitely many efficient portfolios to generate or estimate the set of all the efficient portfolios.

Some new trends and developments in portfolio optimization outside the classical framework are also worth noting, including robust optimization (Shadabfar et al. [16], Caçador et al. [14]), multi-period portfolio optimization (Guo et al. [15]), research and development project portfolio selection (Mavrotas et al. [17]). In particular, Kolm et al. [25] paid attention to constructing risk parity portfolios. The risk parity approach in portfolio construction aims to build portfolios where the overall portfolio risk is diversified by allocating the risk equally across the different investment strategies and/or securities, which can be compared to equal weighted portfolios where $x_j = 1/n$ for all $j = 1, \dots, n$. In Remark 3.1 below, we will explain why the efficient portfolios generated from our model are almost the risk parity ones.

The rest of this paper is organized as follows. In Section 2, we introduce our bi-criteria piecewise linear convex portfolio optimization model. In Section 3, by analyzing the optimality conditions for a weighted sum parametric optimization problem, we derive some explicit analytical formulas for all the efficient portfolios and discuss the structure of the set of all the efficient portfolios. In Section 4, by an example, we illustrate the set of efficient portfolios and the set of efficient frontier, and test the performance of the efficient portfolios versus the inefficient ones by doing some Monte Carlo simulations. In Section 5, we conclude the paper.

Throughout the paper we use the standard notations of convex analysis; see the seminal book [2] by Rockafellar. The inner product of vectors x and y is denoted by $\langle x, y \rangle$. $e_j \in \mathcal{R}^n$ is a vector whose j-th entry is 1 while all the other entries are zero. Let $A \subset \mathcal{R}^n$ be a nonempty set. We denote by $\operatorname{conv} A$, $\operatorname{aff} A$ and $\operatorname{cone} A$ the convex hull, the affine hull and convex cone generated by A, respectively. Let $\lambda \in \mathcal{R}$ be a scalar and $K \subset \mathcal{R}^n$ be another nonempty set. Then the scalar multiple of A and the sum of A and K are respectively given by

$$\lambda A = \{\lambda x \mid x \in A\},$$

$$A + K = \{x + y \mid x \in A, y \in K\}.$$

Let D be a nonempty convex set in $\mathcal{R}^n$, We shall say that D recedes in the direction of y, where $y \neq 0$, if and only if $x + \lambda y \in D$ for every $\lambda \geq 0$ and $x \in D$. The set of all vectors $y \in \mathcal{R}^n$ satisfying this condition, including $y = 0$, will be called the recession cone of D. The dimension of a convex set means that of its affine hull. A polyhedral convex set in $\mathcal{R}^n$ is by definition a set which can be expressed as the intersection of some finite collection of closed half-spaces. Let $P \subset \mathcal{R}^n$ be a convex polyhedral. Then $x \in P$ is a vertex of P if and only if there is no way to express x as a convex combination $(1-\lambda)y+\lambda z$ such that $y \in P$, $z \in P$ and $0 < \lambda < 1$, except by taking $y = z = x$. Meanwhile, an extreme direction of P is defined to be a direction that cannot be expressed as a strictly positive combination of two linearly independent recession vectors of P. For an index set I, $|I|$ represents the number of elements in I.

A vector v is said to be a subgradient of a convex function f at $\bar{x}$ with $f(\bar{x})$ finite, written $v \in \partial f(\bar{x})$, if

$$f(x) \geq f(\bar{x}) + \langle v, x - \bar{x} \rangle \quad \forall x.$$

A random variable will be an element of $\mathcal{L}^2(\Omega) = \mathcal{L}^2(\Omega, \mathcal{M}, \mathcal{P})$, where the elements ω of Ω represent future states, $\mathcal{M}$ is the filed of measurable subsets of Ω, and $\mathcal{P}$ is a probability measure on $\mathcal{M}$. In particular, the space $\mathcal{L}^2(\Omega)$ contains all constant random variables, $\mathbb{R} \equiv C$. The letter C will always stand for a constant in the real numbers $\mathcal{R}$, and any (in)equalities between random variables are to be viewed in the sense of holding almost surely. The essential infimum and supremum of $\mathbb{R}$ will be denoted simply by $\inf(\mathbb{R})$ and $\sup(\mathbb{R})$. We adopt the notion that

$$\mathbb{R} = \mathbb{R}_+ - \mathbb{R}_- \text{ with } \mathbb{R}_+ = \max\{0, \mathbb{R}\},\ \mathbb{R}_- = \max\{0, -\mathbb{R}\}.$$

2 The portfolio optimization model

In this section, we introduce our bi-criteria portfolio optimization model. Throughout the paper, for a random variable $\mathbb{R}$, we denote by $\mathrm{E}(\mathbb{R})$ the mathematical expectation of $\mathbb{R}$, and by $\mathcal{D}(\mathbb{R})$ the general deviation measure of $\mathbb{R}$ defined as follows.

Definition 2.1. *(General deviation measures, [12]). By a deviation measure will be meant any functional* $\mathcal{D} : \mathcal{L}^2(\Omega) \to [0, \infty]$ *satisfying*

(D1) $\mathcal{D}(\mathbb{R} + C) = \mathcal{D}(\mathbb{R})$ *for all* $\mathbb{R}$ *and constants* C*,*

(D2) $\mathcal{D}(0) = 0$ *and* $\mathcal{D}(\lambda\mathbb{R}) = \lambda\mathcal{D}(\mathbb{R})$ *for all* $\mathbb{R}$ *and* $\lambda > 0$*,*

(D3) $\mathcal{D}(\mathbb{R} + \mathbb{R}') \leq \mathcal{D}(\mathbb{R}) + \mathcal{D}(\mathbb{R}')$ *for all* $\mathbb{R}$ *and* $\mathbb{R}'$*,*

(D4) $\mathcal{D}(\mathbb{R}) \geq 0$ *for all* $\mathbb{R}$*, with* $\mathcal{D}(\mathbb{R}) > 0$ *for nonconstant* $\mathbb{R}$*.*

Axiom **(D1)** is equivalent to $\mathcal{D}(\mathbb{R}) = \mathcal{D}(\mathbb{R} - \mathrm{E}(\mathbb{R}))$ for all $\mathbb{R}$. Axiom **(D2)** is positive homogeneity. The combination of **(D2)** and **(D3)** is the property known as sublinearity. It implies that $\mathcal{D}$ is a convex functional on $\mathcal{L}^2(\Omega)$. Axiom **(D4)** means $\mathcal{D}(\mathbb{R}) = 0$ for constant $\mathbb{R}$, whereas $\mathcal{D}(\mathbb{R}) > 0$ for nonconstant $\mathbb{R}$.

Assume that an investor has one unit initial wealth, which is to be invested in n possible assets S_j, $j = 1, \cdots, n$. Let $\mathbb{R}_j$ be the return rate of the asset S_j, which is a random variable. Let x_j be the allocation for investment to S_j. Note that by allowing $x_j < 0$ we are concerned with the situation where short selling is permitted. Thus, in our setting, any vector $x \in \mathcal{R}^n$ can be called a feasible portfolio provided that

$$\sum_{j=1}^{n} x_j = 1$$

is satisfied. For any given feasible portfolio x, its total return rate $\sum_{j=1}^{n} x_j\mathbb{R}_j$ is a random variable, whose expectation is denoted by

$$\pi(x) := \mathrm{E}\left(\sum_{j=1}^{n} x_j\mathbb{R}_j\right),$$

which is a criterion that investors wish to maximize. Instead of measuring the uncertainty in the random variable $\sum_{j=1}^{n} x_j\mathbb{R}_j$ in a direct way, we prefer to do it in a two-step way: first measure the uncertainty in each random variable $x_j\mathbb{R}_j$ by virtue of some deviation measure, and then take the maximum of these deviation measures as the degree of uncertainty in $\sum_{j=1}^{n} x_j\mathbb{R}_j$, i.e.,

$$\omega(x) := \max_{1 \leq j \leq n} \mathcal{D}(x_j\mathbb{R}_j),$$

is a convex criterion that investors wish to minimize (the convexity of ω comes from that of $\mathcal{D}$). Due to the nonnegativity and the positive homogeneity of $\mathcal{D}$, the latter criterion can be reformulated as

$$\omega(x) = \max_{1\leq j\leq n} \max \left\{x_j\mathcal{D}(\mathbb{R}_j), -x_j\mathcal{D}(-\mathbb{R}_j)\right\}.$$

We assume that investors wish to maximize $\pi(x)$ while minimizing $\omega(x)$, which are two criteria in conflict. Our portfolio optimization model can be formulated as a bi-criteria piecewise linear convex program

$$\begin{array}{ll} \text{minimize} & (\omega(x), -\pi(x)) \\ \text{subject to} & \displaystyle\sum_{j=1}^{n} x_j = 1, \end{array} \tag{1}$$

or explicitly,

$$\begin{array}{ll} \text{minimize} & \left(\displaystyle\max_{1\leq j\leq n} \max \left\{\underline{q}_j x_j, -\overline{q}_j x_j\right\}, -\sum_{j=1}^{n} x_j r_j\right) \\ \text{subject to} & \displaystyle\sum_{j=1}^{n} x_j = 1, \end{array}$$

where the following notations are used throughout the paper:

$$r_j = \mathrm{E}(\mathbb{R}_j), \quad \underline{q}_j = \mathcal{D}(\mathbb{R}_j), \quad \overline{q}_j = \mathcal{D}(-\mathbb{R}_j).$$

Remark 2.1. *For an asset in a long position, the uncertainty below the expectation is undesirable, while the uncertainty above the expectation is desirable. In contrast, for an asset in a short position, the uncertainty below the expectation is desirable, while the uncertainty above the expectation is undesirable. In this regard, some deviation measures may fit better for our model than the others, e.g.,* $\sqrt{\mathrm{E}(\mathrm{E}(\mathbb{R}) - \mathbb{R})_+^2}$ *is more suitable than* $\sqrt{\mathrm{E}(\mathbb{R} - \mathrm{E}(\mathbb{R}))_+^2}$ *in measuring the downside uncertainty in* $\mathbb{R}$*. We list some deviation measures that fit our model and often appear in the literature as follows:*

(a) *The standard deviation* $\mathcal{D}(\mathbb{R}) := \sqrt{\mathrm{E}(\mathrm{E}(\mathbb{R}) - \mathbb{R})^2}$*;*

(b) *The absolute deviation* $\mathcal{D}(\mathbb{R}) := \mathrm{E}(|\mathbb{R} - \mathrm{E}(\mathbb{R})|)$*;*

(c) *The lower semi-absolute deviation* $\mathcal{D}(\mathbb{R}) := \mathrm{E}(\mathrm{E}(\mathbb{R}) - \mathbb{R})_+$*;*

(d) *The standard lower semideviation* $\mathcal{D}(\mathbb{R}) := \sqrt{\mathrm{E}(\mathrm{E}(\mathbb{R}) - \mathbb{R})_+^2}$*;*

(e) *The lower range deviation* $\mathcal{D}(\mathbb{R}) := \mathrm{E}(\mathbb{R}) - \inf(\mathbb{R})$*;*

(f) *The CVaR-deviation* $\mathcal{D}(\mathbb{R}) := \mathrm{CVaR}_\alpha(\mathbb{R} - \mathrm{E}(\mathbb{R}))$ *for some* $\alpha \in (0, 1)$.

We refer the reader to [12] for more derivation measures and their properties. When the deviation measure $\mathcal{D}$ *is symmetric with respect to ups and downs such as the ones listed in (a)-(c), we have* $\underline{q}_j = \overline{q}_j$ *for all* j*. However, when the deviation measure* $\mathcal{D}$ *is not symmetric with respect to ups* *and downs such as the ones listed in (d)-(f), we may not have* $\underline{q}_j = \overline{q}_j$ *for all* j*. When* $\mathcal{D}$ *is as listed in (b) and* $x_j \geq 0$ *for all* j*, we have*

$$\omega(x) = \max_{1\leq j\leq n} x_j \,\mathrm{E}(|\mathbb{R}_j - \mathrm{E}(\mathbb{R}_j)|),$$

which is exactly the same as the l_∞ *risk function studied in [4, Definition 2.1].*

Remark 2.2. *For any* $\xi > 0$*, we have*

$$\begin{aligned}
\Pr\left(\pi(x) - \sum_{j=1}^{n} \mathbb{R}_j x_j \geq \xi\right) &= \Pr\left(\left[\sum_{j=1}^{n} \Big(\mathrm{E}(x_j\mathbb{R}_j) - x_j\mathbb{R}_j\Big)\right]_+ \geq \xi\right) \\
&\leq \mathrm{E}\left[\sum_{j=1}^{n} \Big(\mathrm{E}(x_j\mathbb{R}_j) - x_j\mathbb{R}_j\Big)\right]_+ /\xi \\
&\leq \mathrm{E}\left[\sum_{j=1}^{n} \Big(\mathrm{E}(x_j\mathbb{R}_j) - x_j\mathbb{R}_j\Big)_+\right] /\xi \\
&\leq n \max_{1\leq j\leq n} \mathrm{E}\Big(\mathrm{E}(x_j\mathbb{R}_j) - x_j\mathbb{R}_j\Big)_+ /\xi,
\end{aligned}$$

where the first inequality follows from the Markov inequality. That is, the probability that the downside deviation of the random variable $\sum_{j=1}^{n} x_j\mathbb{R}_j$ *from its expected return* $\pi(x)$ *is greater than a pre-specified level is bounded by a constant (independent of the choice of* x*) multiplied by*

$$\max_{1\leq j\leq n} \mathrm{E}\Big(\mathrm{E}(x_j\mathbb{R}_j) - x_j\mathbb{R}_j\Big)_+ . \tag{2}$$

This probability will be small if (2) is kept small, suggesting that we can use (2) to measure the degree of the downside uncertainty in $\sum_{j=1}^{n} x_j\mathbb{R}_j$*. Note that (2) is equal to* $\omega(x)$ *when the deviation measure* $\mathcal{D}$ *is given as in Remark 2.1 (c), and that (2) is less than* $\omega(x)$ *when the deviation measure* $\mathcal{D}$ *is given as in Remark 2.1 (b) and (e). Roughly, the degree of the downside uncertainty in* $\sum_{j=1}^{n} x_j\mathbb{R}_j$ *can be 'measured' in some sense by* $\omega(x)$*. This is the reason why we employ* $\omega(x)$ *as a criterion to be minimized in our model.*

Recall that a feasible portfolio x is said to be efficient if there exists no feasible portfolio x' such that

$$\omega(x') \leq \omega(x), \qquad \pi(x') \geq \pi(x),$$

and at least one of the inequality holds strictly. Accordingly, the function value $(\omega(x), \pi(x))$ is said to be an efficient point in the risk-return (i.e., $\omega - \pi$) plane. Throughout this paper, we denote by **EP** the set of all efficient portfolios, and by **EF** the efficient frontier which consists of all efficient points in the $\omega - \pi$ plane. In section 3 below, we will provide an analytic derivation of **EP** and **EF**.

3 Analytic Derivation of Efficient Portfolios and Efficient Frontier

Throughout the paper, we assume that all assets are risky (i.e., $\overline{q}_j$, $\underline{q}_j > 0$ for all j), that

$$r_1 \le r_2 \le \cdots \le r_{n-1} \le r_n,$$

and that there exist some j and k such that $r_j \neq r_k$. To seek **EP** and **EF** for the bi-criteria piecewise linear convex program (1), one way is to convert(1) into a parametric optimization problem

$$\begin{aligned} &\text{minimize} && \lambda\omega(x) - \pi(x) \\ &\text{subject to} && \sum_{j=1}^{n} x_j = 1, \end{aligned} \tag{3}$$

where $\lambda > 0$ can be considered as an investor's risk tolerance parameter-the larger the λ, the less risk the investor is to tolerate.

In what follows, we denote by $E(\lambda)$ the optimal solution set of (3). As (1) is a bi-criteria convex program, x is an efficient portfolio if and only if there is some $\lambda > 0$ such that $x \in E(\lambda)$, see [5]. This entails that

$$\mathbf{EP} = \bigcup_{\lambda\in(0,+\infty)} E(\lambda). \tag{4}$$

The aim of this section is to derive **EP** and **EF** by exploring (3) and providing some analytical expressions for $E(\lambda)$.

To begin with, we introduce two functions which play crucial roles in our analytic derivation. In terms of the following notations:

$$\begin{aligned} &I := \{1, \cdots, n\}, \\ &I^+(t) := \{j \in I \mid r_j > t\}, \\ &I^0(t) := \{j \in I \mid r_j = t\}, \\ &I^-(t) := \{j \in I \mid r_j < t\}, \\ &I^{0+}(t) := \{j \in I \mid r_j \ge t\}, \end{aligned}$$

we define $g : \mathcal{R} \to \mathcal{R}$ by

$$g(t) := \sum_{j\in I^{0+}(t)} \frac{r_j - t}{\underline{q}_j} + \sum_{j\in I^-(t)} \frac{t - r_j}{\overline{q}_j},$$

and $h : \mathcal{R} \to \mathcal{R}$ by

$$h(t) := \sum_{j\in I^{0+}(t)} \frac{1}{\underline{q}_j} - \sum_{j\in I^-(t)} \frac{1}{\overline{q}_j},$$

and set

$$m := \max\{j \in I \mid h(r_j) > 0\}, \quad e := (1, \cdots, 1)^T \in \mathcal{R}^n.$$

Moreover, we define

$$x^{0,1} := \left(1/\underline{q}_1, \cdots, 1/\underline{q}_n\right)^T / h(r_1),$$

and for each $2 \le k \le m$ with $r_{k-1} < r_k$, define $x^{k-1,k}$ by

$$x_j^{k-1,k} := \begin{cases} -\dfrac{1}{h(r_k)}\dfrac{1}{\overline{q}_j} & \text{if } j \in \{1, \cdots, k-1\}, \\ \dfrac{1}{h(r_k)}\dfrac{1}{\underline{q}_j} & \text{if } j \in \{k, \cdots, n\}. \end{cases}$$

The following lemma, which is very helpful for our analytic derivation, follows readily from the definitions of g and h.

Lemma 3.1. *The following properties of g and h hold:*

(a) *h is a decreasing upper semi-continuous step function with $h(t) > 0$ for all $t \le r_m$ and $h(t) \le 0$ for all $t > r_m$. Explicitly, we have*

$$h(t) = \begin{cases} \displaystyle\sum_{j=1}^{n} \frac{1}{\underline{q}_j} & \text{if } t \le r_1, \\ \displaystyle\sum_{j=k}^{n} \frac{1}{\underline{q}_j} - \sum_{j=1}^{k-1} \frac{1}{\overline{q}_j} & \text{if } r_{k-1} < t \le r_k, \\ \displaystyle -\sum_{j=1}^{n} \frac{1}{\overline{q}_j} & \text{if } t > r_n, \end{cases}$$

which yields that

$$h(r_{k-1}) - h(t) = \sum_{r_j = r_{k-1}} \left(\frac{1}{\underline{q}_j} + \frac{1}{\overline{q}_j}\right) \quad \forall t \in (r_{k-1}, r_k],$$

and

$$h(r_n) - h(t) = \sum_{r_j = r_n} \left(\frac{1}{\underline{q}_j} + \frac{1}{\overline{q}_j}\right) \quad \forall t \in (r_n, +\infty).$$

(b) *g is a continuous piecewise linear convex function. Explicitly, we have*

$$g(t) = \begin{cases} g(r_1) + h(r_1)(r_1 - t) & \text{if } t \le r_1, \\ g(r_k) + h(r_k)(r_k - t) & \text{if } r_{k-1} \le t \le r_k, \\ \displaystyle\sum_{j=1}^{n} \frac{t - r_j}{\overline{q}_j} & \text{if } t > r_n. \end{cases}$$

(c) *$g(t)$ is strictly decreasing on $(-\infty, r_m]$, and $\min_{t \in \mathcal{R}} g(t) = g(r_m) > 0$.*

The following lemma, which follows directly from [1, Exercise 8.31], is helpful for deriving optimality conditions for (3).

Lemma 3.2. *Let $x \in \mathcal{R}^n \backslash \{0\}$. Then $v \in \partial\omega(x)$ (i.e., v is a subgradient of ω at x) if and only if*

$$x_j = \begin{cases} -\frac{\omega(x)}{\bar{q}_j} & \text{if } v_j < 0, \\ \frac{\omega(x)}{\underline{q}_j} & \text{if } v_j > 0, \end{cases}$$

and

$$\sum_{v_j>0} \frac{v_j}{\underline{q}_j} + \sum_{v_j<0} -\frac{v_j}{\bar{q}_j} = 1.$$

Our first finding is that the portfolio having the global minimum l_∞ downside risk is the $1/\underline{q}_j$ portfolio strategy, which recovers the so-called $1/N$ portfolio strategy [3] when all $\underline{q}_j$ are equal.

Proposition 3.1 (global minimum l_∞ downside risk portfolio)**.** *The unique optimal solution to*

$$\begin{aligned} &\text{minimize} && \omega(x) \\ &\text{subject to} && \sum_{j=1}^{n} x_j = 1, \end{aligned} \tag{5}$$

is $x^{0,1}$, for which $\omega(x^{0,1}) = 1/h(r_1)$ and $\pi(x^{0,1}) = r_1 + g(r_1)/h(r_1)$.

Proof. According to the optimality condition for convex optimization (cf. 8.15 Theorem of Rockafellar and Wets (1998)), $x \in \mathcal{R}^n$ is a solution to (5) if and only if there exists some $\tau \in \mathcal{R}$ such that

$$\sum_{j=1}^{n} x_j = 1, \tag{6}$$

and

$$-\tau e \in \partial\omega(x). \tag{7}$$

We have $x \neq 0$, for otherwise (6) cannot be fulfilled. Applying Lemma 3.2, we confirm that $\tau = -1/h(r_1) < 0$, and that there is a unique vector $x = x^{0,1}$ satisfying both (6) and (7). This completes the proof. □

Now we are ready for providing an analytic derivation of $E(\lambda)$, the optimal solution set of the parametric optimization problem (3).

Theorem 3.1. *The parametric optimization problem (3) has a nonempty optimal solution set $E(\lambda)$ if and only if $\lambda \geq g(r_m)$, in which case, $x \in E(\lambda)$ if and only if each component x_j can be written as*

$$x_j = \left(\frac{1-\theta_j}{\underline{q}_j} - \frac{\theta_j}{\bar{q}_j} \right) \Big/ \sum_{i=1}^{n} \left(\frac{1-\theta_i}{\underline{q}_i} - \frac{\theta_i}{\bar{q}_i} \right), \tag{8}$$

where $\theta \in \mathcal{R}^n$ is any vector satisfying

(a) *in the case of $g(r_1) < \lambda < +\infty$:*

$$\theta_j = 0 \quad \text{for all } j,$$

i.e., $E(\lambda) = \{x^{0,1}\}$ in this case;

(b) *in the case of* $g(r_k) < \lambda < g(r_{k-1})$ *for some* $k \in \{2, \cdots, m\}$*:*

$$\theta_j = \begin{cases} 1 & \textit{if } j \in \{1, \cdots, k-1\}, \\ 0 & \textit{if } j \in \{k, \cdots, n\}, \end{cases}$$

i.e., $E(\lambda) = \{x^{k-1,k}\}$ *in this case;*

(c) *in the case of* $\lambda = g(r_k)$ *with* $r_k < r_m$*:*

$$\theta_j = \begin{cases} 1 & \textit{if } r_j < r_k, \\ 0 & \textit{if } r_j > r_k, \end{cases}$$

and

$$0 \le \theta_j \le 1 \textit{ if } r_j = r_k;$$

(d) *in the case of* $\lambda = g(r_m)$*:*

$$\theta_j = \begin{cases} 1 & \textit{if } r_j < r_m, \\ 0 & \textit{if } r_j > r_m, \end{cases}$$

$$0 \le \theta_j \le 1 \textit{ if } r_j = r_m,$$

and

$$\sum_{i=1}^{n} \left(\frac{1-\theta_i}{\underline{q}_i} - \frac{\theta_i}{\overline{q}_i} \right) > 0.$$

Proof. In view of Lemma 3.1 (c), we have $0 < g(r_m) \le g(r_{m-1}) \le \cdots \le g(r_2) \le g(r_1) < +\infty$, and $g(r_k) < g(r_{k-1})$ whenever $r_{k-1} < r_k \le r_m$. To fully characterize the optimal solution set $E(\lambda)$ of the parametric optimization problem (3) for all $\lambda > 0$, we only need to consider cases (a)-(d) and the case that $0 < \lambda < g(r_m)$.

According to the optimality condition for convex optimization (cf. 8.15 Theorem of Rockafellar and Wets (1998)), $x \in E(\lambda)$, i.e., x is an optimal solution to (3) if and only if there exists some $\tau \in \mathcal{R}$ such that

$$\langle e, x \rangle = 1, \tag{9}$$

and

$$r - \tau e \in \lambda \partial \omega(x). \tag{10}$$

Observing that $x \neq 0$ (otherwise contradicting to (9)), we have $\omega(x) > 0$. Depending on the value of τ, we can divide $\{1, \cdots, n\}$ into three distinct subindex sets as follows:

$$I^- := I^-(\tau),\ I^0 := I^0(\tau),\ I^+ := I^+(\tau).$$

By Lemma 3.2, (10) holds if and only if

$$x_j = \begin{cases} -\frac{\omega(x)}{\overline{q}_j} & \text{if } j \in I^-, \\ \frac{\omega(x)}{\underline{q}_j} & \text{if } j \in I^+, \end{cases} \tag{11}$$

and

$$\sum_{j\in I^+}(r_j-\tau)\frac{1}{\underline{q}_j}+\sum_{j\in I^-}(\tau-r_j)\frac{1}{\overline{q}_j}=\lambda.$$

The latter equation can be written in terms of the function g as

$$g(\tau)=\lambda. \tag{12}$$

By the definition of $\omega(x)$, we have

$$-\frac{\omega(x)}{\overline{q}_j}\le x_j\le\frac{\omega(x)}{\underline{q}_j}\quad\forall j\in I^0. \tag{13}$$

In view of (11), we deduce that (9) holds if and only if there are some $\theta_j\in[0,1]$ with $j\in I^0$ such that

$$x_j=\left(-\frac{\theta_j}{\overline{q}_j}+\frac{1-\theta_j}{\underline{q}_j}\right)\omega(x)\quad\forall j\in I^0, \tag{14}$$

and

$$\sum_{j\in I^+}\frac{1}{\underline{q}_j}-\sum_{j\in I^-}\frac{1}{\overline{q}_j}+\sum_{j\in I^0}\left(-\frac{\theta_j}{\overline{q}_j}+\frac{1-\theta_j}{\underline{q}_j}\right)=\frac{1}{\omega(x)}. \tag{15}$$

By the definition of h, we have

$$h(\tau)-\sum_{j\in I^0}\theta_j\left(\frac{1}{\overline{q}_j}+\frac{1}{\underline{q}_j}\right)=\sum_{j\in I^+}\frac{1}{\underline{q}_j}-\sum_{j\in I^-}\frac{1}{\overline{q}_j}+\sum_{j\in I^0}\left(-\frac{\theta_j}{\overline{q}_j}+\frac{1-\theta_j}{\underline{q}_j}\right)>0, \tag{16}$$

where the inequality follows from (15) and the fact that $\omega(x)>0$. According to Lemma 3.1 (a), the inequality in (16) holds for some $\theta_j\in[0,1]$ with $j\in I^0$, if and only if, $\tau\le r_m$. By Lemma 3.1 (c), we conclude that the equation (12) has a (unique) solution in $\tau\in(-\infty,r_m]$ if and only if $\lambda\ge g(r_m)$. This indicates that $E(\lambda)\neq\emptyset$ if and only if $\lambda\ge g(r_m)$.

It remains to show (a)-(d) by first identifying the unique $\tau\in(-\infty,r_m]$ satisfying (12), and then using the relations (11) and (14-16) to describe components x_j of each $x\in E(\lambda)$ as follows:

$$x_j=\left(\frac{1-\theta_j}{\underline{q}_j}-\frac{\theta_j}{\overline{q}_j}\right)\Bigg/\sum_{i=1}^{n}\left(\frac{1-\theta_i}{\underline{q}_i}-\frac{\theta_i}{\overline{q}_i}\right), \tag{17}$$

where $\theta\in\mathcal{R}^n$ is any vector satisfying

$$\theta_j=\begin{cases}1 & \text{if } j\in I^-,\\ 0 & \text{if } j\in I^+,\end{cases}\qquad\text{and}\quad 0\le\theta_j\le 1 \text{ if } j\in I^0, \tag{18}$$

and

$$\sum_{i=1}^{n}\left(\frac{1-\theta_i}{\underline{q}_i}-\frac{\theta_i}{\overline{q}_i}\right)>0. \tag{19}$$

So the key point to verify the optimality of x is to determine the value of the unique $\tau \in (-\infty, r_m]$ via the piecewise linear equation (12) for all particular cases (a)-(d). This can be done in an analytical way as in the following, where the fact pointed out in Lemma 3.1 (c) that g is strict decreasing on the interval $(-\infty, r_m]$ plays a key role.

In case (a), we have $g(\tau) = \lambda > g(r_1)$ and hence $\tau < r_1$. More precisely, we have by Lemma 3.1 (b),

$$\tau = \frac{g(r_1) - \lambda}{h(r_1)} + r_1. \tag{20}$$

This implies that $I^+ = \{1, \cdots, n\}$ and $I^- = I^0 = \emptyset$. In this case, the inequality (19) holds for the unique θ satisfying (18), because we have

$$\sum_{i=1}^{n} \left(\frac{1-\theta_i}{\underline{q}_i} - \frac{\theta_i}{\overline{q}_i} \right) = h(r_1) > 0.$$

In case (b), we have $g(r_k) < g(\tau) = \lambda < g(r_{k-1})$ and hence $r_{k-1} < \tau < r_k$. More precisely, we have by Lemma 3.1 (b),

$$\tau = \frac{g(r_k) - \lambda}{h(r_k)} + r_k. \tag{21}$$

This implies that $I^+ = \{j \mid r_j \geq r_k\}$, $I^- = \{j \mid r_j < r_k\}$ and $I^0 = \emptyset$. In this case, the inequality (19) holds for the unique θ satisfying (18), because we have

$$\sum_{i=1}^{n} \left(\frac{1-\theta_i}{\underline{q}_i} - \frac{\theta_i}{\overline{q}_i} \right) = h(r_k) \geq h(r_m) > 0.$$

In case (c), we have $g(\tau) = \lambda = g(r_k)$ and hence $\tau = r_k < r_m$. This implies that $I^+ = \{j \mid r_j > r_k\}$, $I^- = \{j \mid r_j < r_k\}$, $I^0 = \{j \mid r_j = r_k\}$. In this case, the inequality (19) holds for all the θ satisfying (18), because we have

$$\sum_{i=1}^{n} \left(\frac{1-\theta_i}{\underline{q}_i} - \frac{\theta_i}{\overline{q}_i} \right) \geq h(r_k) - \sum_{j \in I^0} \left(\frac{1}{\overline{q}_j} + \frac{1}{\underline{q}_j} \right) \geq h(r_m) > 0,$$

where the first inequality follows by setting $\theta_j = 1$ for all $j \in I^0$, and the second inequality follows from Lemma 3.1 (a).

In case (d), we have $g(\tau) = \lambda = g(r_m)$ and hence $\tau = r_m$. This implies that $I^+ = \{j \mid r_j > r_m\}$, $I^- = \{j \mid r_j < r_m\}$, $I^0 = \{j \mid r_j = r_m\}$. Note that in this case, the inequality (19) cannot hold for all the θ satisfying (16), as in particular when $\theta_j = 1$ for all $j \in I^0$, we get from Lemma 3.1 (a)

$$\sum_{i=1}^{n} \left(\frac{1-\theta_i}{\underline{q}_i} - \frac{\theta_i}{\overline{q}_i} \right) = h(r_m) - \sum_{j \in I^0} \left(\frac{1}{\overline{q}_j} + \frac{1}{\underline{q}_j} \right) \leq 0.$$

This completes the proof. □

Remark 3.1. *For each efficient portfolio x in the form of (8), we have*

$$\mathcal{D}(x_j\mathbb{R}_j) = \omega(x)$$

for most j's such that $\theta_j \in \{0, 1\}$, *and*

$$\mathcal{D}(x_j\mathbb{R}_j) < \omega(x)$$

for only a few j's such that $0 < \theta_j < 1$. *This suggests that our efficient portfolios are almost the risk parity ones in the sense that the risks are allocated equally across the investments. See [25] for more details on the risk parity.*

Remark 3.2. *When* λ *moves inside the open intervals* $(g(r_1), +\infty)$ *and* $(g(r_k), g(r_{k-1}))$ *for some* $k \in \{2, \cdots, m\}$ *with* $r_{k-1} < r_k$, *the parametric optimization problem (3) shares the same unique optimal solution, which remains as an optimal solution to the parametric optimization problem (3) whenever* λ *moves on to the left or the right endpoint of the corresponding open interval. In other words, all the solutions to the parametric optimization problem (3) can be found by checking finitely many optimal solution sets* $E(\lambda)$ *for* $\lambda = g(r_k)$ *with* $k \in \{1, \cdots, m\}$, *and each* $E(g(r_{k-1}))$ *with* $k \in \{2, \cdots, m\}$ *has a common point* $x^{k-1,k}$ *with* $E(g(r_k))$ *as long as* $r_{k-1} < r_k$. *Some elaborate properties of* $E(g(r_k))$ *with* $k \in \{1, \cdots, m\}$ *are presented in the following corollary.*

Corollary 3.1. *Fix some* $k \in \{1, \cdots, m\}$. *In terms of*

$$I^- := I^-(r_k), \quad I^0 := I^0(r_k), \quad I^+ := I^+(r_k), \quad \xi(k) := \min I^0, \quad \eta(k) = \max I^0,$$

and

$$\Theta =: \left\{\theta \in \mathcal{R}^n \mid \theta_j = 1 \ \forall j \in I^-, \ 0 \le \theta_j \le 1 \ \forall j \in I^0, \ \theta_j = 0 \ \forall j \in I^+\right\},$$

the following statements on the optimal solution set $E(\lambda)$ *of (3) for* $\lambda = g(r_k)$ *are true:*

(i) $x \in E(\lambda)$ *if and only if* $\pi(x) = g(r_k)\omega(x) + r_k$ *and* $\langle e, x\rangle = 1$.

(ii) $E(\lambda)$ *is the solution set of the linear system*

$$\begin{cases} -\overline{q}_j x_j &= \dfrac{\pi(x) - r_k}{g(r_k)} \quad \forall j \in I^-, \\ -\overline{q}_j x_j &\le \dfrac{\pi(x) - r_k}{g(r_k)} \quad \forall j \in I^0, \\ \underline{q}_j x_j &\le \dfrac{\pi(x) - r_k}{g(r_k)} \quad \forall j \in I^0, \\ \underline{q}_j x_j &= \dfrac{\pi(x) - r_k}{g(r_k)} \quad \forall j \in I^+, \\ \langle e, x\rangle &= 1. \end{cases} \tag{22}$$

(iii) *In the case of* $r_k < r_m$, $E(\lambda)$ *is a bounded polyhedron satisfying*

$$\{\omega(x) \mid x \in E(\lambda)\} = \left[\frac{1}{h(r_k)}, \ \frac{1}{h(r_{\eta(k)+1})}\right]. \tag{23}$$

In this case, x is a vertex of $E(\lambda)$ if and only if

$$x_j = \left(\frac{1-\theta_j}{\underline{q}_j} - \frac{\theta_j}{\overline{q}_j} \right) \Bigg/ \sum_{i=1}^{n} \left(\frac{1-\theta_i}{\underline{q}_i} - \frac{\theta_i}{\overline{q}_i} \right), \tag{24}$$

where $\theta \in \Theta$ is any vector satisfying $\theta_j \in \{0,1\}$ for all $j \in I^0$, entailing that the number of vertices of $E(\lambda)$ is $2^{|I^0|}$. In particular when $I_0 = \{k\}$,

$$E(\lambda) = [x^{k-1,k}, x^{k,k+1}],$$

i.e., $E(\lambda)$ is a closed line segment joining $x^{k-1,k}$ and $x^{k,k+1}$.

(iv) *In the case of $r_k = r_m$, $E(\lambda)$ is an unbounded polyhedron satisfying*

$$\{\omega(x) \mid x \in E(\lambda)\} = \left[\frac{1}{h(r_m)}, +\infty \right). \tag{25}$$

In this case, x is a vertex of $E(\lambda)$ if and only if (24) holds for any $\theta \in \Theta$ satisfying $\theta_j \in \{0,1\}$ for all $j \in I^0$ and

$$\sum_{i=1}^{n} \left(\frac{1-\theta_i}{\underline{q}_i} - \frac{\theta_i}{\overline{q}_i} \right) > 0. \tag{26}$$

Moreover, $w \in \mathcal{R}^n$ is an extreme direction of $E(\lambda)$ if and only if there are some $\tau > 0$ and $j_0 \in I^0$ such that $w = \tau \bar{w}$, and for all j,

$$\bar{w}_j = \frac{1-\theta_j}{\underline{q}_j} - \frac{\theta_j}{\overline{q}_j}, \tag{27}$$

where $\theta \in \Theta$ is any vector such that $\theta_j \in \{0,1\}$ for all $j \in I^0 \backslash \{j_0\}$ and

$$\sum_{i=1}^{n} \left(\frac{1-\theta_i}{\underline{q}_i} - \frac{\theta_i}{\overline{q}_i} \right) = 0.$$

In particular when $I_0 = \{m\}$,

$$E(\lambda) = \left\{ x^{m-1,m} + t(x^{m-1,m} - e_m) \mid t \geq 0 \right\},$$

i.e., $E(\lambda)$ is a half-line emanating from $x^{m-1,m}$, whose reverse extension passes through e^m.

(v) *The dimension of $E(\lambda)$ is $|I^0|$ and*

$$E(\lambda) \subset x^{\xi(k)-1,\xi(k)} + \text{cone}\left\{ x^{\xi(k)-1,\xi(k)} - e_j, j \in I^0 \right\},$$

where $x^{\xi(k)-1,\xi(k)}$ is a vertex of $E(\lambda)$ and $\text{cone}\left\{ x^{\xi(k)-1,\xi(k)} - e_j, j \in I^0 \right\}$ is the smallest convex cone containing $E(\lambda) - x^{\xi(k)-1,\xi(k)}$.

Proof. For the sake of simplicity, we denote by $x(\theta)$ the vector whose components are in the form of (8). According to Theorem 3.1, $x \in E(\lambda)$ if and only if $x = x(\theta)$ for some $\theta \in \Theta$ in the case of $r_k < r_m$, while in the case of $r_k = r_m$, $x \in E(\lambda)$ if and only if $x = x(\theta)$ for some $\theta \in \Theta$ satisfying

$$\sum_{i=1}^{n} \left(\frac{1-\theta_i}{\underline{q}_i} - \frac{\theta_i}{\overline{q}_i} \right) > 0. \tag{28}$$

Let $\tilde{x} = x(\theta)$ with $\theta \in \Theta$ satisfying $\theta_j = 0$ for all $j \in I^0$. Clearly, we have $\tilde{x} \in E(\lambda)$ and $\tilde{x} = x^{\xi(k)-1,\xi(k)}$. By definition, we have

$$\omega(\tilde{x}) = \frac{1}{h(r_k)},$$

and

$$\begin{aligned} \pi(\tilde{x}) &= \frac{1}{h(r_k)} \left(\sum_{j \in I^-} -\frac{r_j}{\overline{q}_j} + \sum_{j \in I^0 \cup I^+} \frac{r_j}{\underline{q}_j} \right) \\ &= \frac{1}{h(r_k)} \left(\sum_{j \in I^-} \frac{r_k - r_j}{\overline{q}_j} + \sum_{j \in I^0 \cup I^+} \frac{r_j - r_k}{\underline{q}_j} \right) + r_k \\ &= g(r_k)\omega(\tilde{x}) + r_k. \end{aligned}$$

So we have $\lambda\omega(\tilde{x}) - \pi(\tilde{x}) = g(r_k)\omega(\tilde{x}) - \pi(\tilde{x}) = r_k$, meaning that the optimal value of (3) is r_k. Therefore, $x \in E(\lambda)$ if and only if $\langle e, x \rangle = 1$ and $\lambda\omega(x) - \pi(x) = r_k$, i.e., statement (i) is true. By statement (i), Theorem 3.1 and the definition of ω, we can easily show that statement (ii) is also true, implying that $E(\lambda)$ is a polyhedral set.

To show statement (iii), assume that $\lambda = g(r_k)$ with $r_k < r_m$. In view of the equality

$$\omega(x(\theta)) = \frac{1}{h(r_k) - \sum_{j \in I^0} \theta_j (\frac{1}{\overline{q}_j} + \frac{1}{\underline{q}_j})} \quad \forall \theta \in \Theta,$$

we get (23) immediately from Lemma 3.1 (a) and the assumption that $r_k < r_m$. It then follows that for all $x \in E(\lambda)$ and all j,

$$|x_j| \le \omega(x) \max_{1 \le j \le n} \left\{ \frac{1}{\overline{q}_j}, \frac{1}{\underline{q}_j} \right\} \le \frac{1}{h(r_{\eta(k)+1})} \max_{1 \le j \le n} \left\{ \frac{1}{\overline{q}_j}, \frac{1}{\underline{q}_j} \right\},$$

which implies the boundedness of $E(\lambda)$. Let $x \in E(\lambda)$ or equivalently let $x =: x(\theta)$ for some $\theta \in \Theta$. From statements (i-ii), it follows that $x(\theta)$ is a solution to the following system of linear equations:

$$\begin{cases} -\bar{q}_j x_j &= \dfrac{\pi(x) - r_k}{g(r_k)} \quad \forall j \in I^-, \\ -\bar{q}_j x_j &= \dfrac{\pi(x) - r_k}{g(r_k)} \quad \forall j \in I^{0-}, \\ \underline{q}_j x_j &= \dfrac{\pi(x) - r_k}{g(r_k)} \quad \forall j \in I^{0+}, \\ \underline{q}_j x_j &= \dfrac{\pi(x) - r_k}{g(r_k)} \quad \forall j \in I^+, \\ \langle e, x\rangle &= 1, \end{cases} \tag{29}$$

where

$$I^{0-} =: \{j \in I^0 \mid \theta_j = 1\} \quad \text{and} \quad I^{0+} =: \{j \in I^0 \mid \theta_j = 0\}.$$

In the case of $\theta_j \in \{0,1\}$ for all $j \in I^0$, it follows from statements (i-ii) that each $x(\tilde{\theta})$ with $\tilde{\theta} \in \Theta\backslash\{\theta\}$ violates at least one equation in (29), implying that $x(\theta)$ is the unique solution to the system of linear equations (29) and hence a vertex of $E(\lambda)$. While in the case of $0 < \theta_{j_0} < 1$ for some $j_0 \in I^0$, the vector $x(\tilde{\theta})$ with $\tilde{\theta}_j = \theta_j$ for all $j \in I^0\backslash\{j_0\}$ and $\tilde{\theta}_{j_0} = \frac{1}{2}\theta_{j_0}$ is another distinct solution to the system of linear equations (29), implying that $x(\theta)$ is not a vertex of $E(\lambda)$. So, x is a vertex of $E(\lambda)$ if and only if $x = x(\theta)$ for any $\theta \in \Theta$ satisfying $\theta_j \in \{0,1\}$ for all $j \in I^0$. Clearly, the number of vertices of $E(\lambda)$ is $2^{|I^0|}$, and in particular $x^{\xi(k)-1,\xi(k)}$ is a vertex of $E(\lambda)$. In the case of $I_0 = \{k\}$, $E(\lambda)$ has exactly two vertices: $x^{k-1,k}$ and $x^{k,k+1}$ (corresponding to $\theta_k = 0$ and $\theta_k = 1$, respectively).

To show statement (iv), assume that $\lambda = g(r_m)$. From Lemma 3.1 (a) and the fact that $m = \max\{j \mid h(r_j) > 0\}$, it follows that

$$h(r_m) - \sum_{j\in I^0}\Big(\frac{1}{\bar{q}_j} + \frac{1}{\underline{q}_j}\Big) \le 0,$$

which, together with the fact that for all $\theta \in \Theta$ satisfying (28) the following hold:

$$\omega(x(\theta)) = \frac{1}{h(r_m) - \sum_{j\in I^0}\theta_j\Big(\frac{1}{\bar{q}_j} + \frac{1}{\underline{q}_j}\Big)} \ge \frac{1}{h(r_m)},$$

implies (25) and hence the unboundedness of $E(\lambda)$. By the same argument as in statement (iii), the (nonempty) set of vertices of $E(\lambda)$ can be described as required, and in particular $x^{\xi(k)-1,\xi(k)}$ is a vertex of $E(\lambda)$. Furthermore, $w \in \mathcal{R}^n$ is a direction of $E(\lambda)$ if and only if w is a nonzero solution to the following homogeneous linear system:

$$\begin{cases} -\overline{q}_j w_j &= \dfrac{\pi(w)}{g(r_m)} \quad \forall j \in I^-, \\ -\overline{q}_j w_j &\leq \dfrac{\pi(w)}{g(r_m)} \quad \forall j \in I^0, \\ \underline{q}_j w_j &\leq \dfrac{\pi(w)}{g(r_m)} \quad \forall j \in I^0, \\ \underline{q}_j w_j &= \dfrac{\pi(w)}{g(r_m)} \quad \forall j \in I^+, \\ \langle e, w\rangle &= 0. \end{cases} \tag{30}$$

Denote by W the solution set of (30). In view of $I^0 \neq \emptyset$, we have

$$\frac{\pi(w)}{g(r_m)} > 0 \qquad w \in W \backslash \{0\}.$$

This suggests, by homogeneity, that each direction of $E(\lambda)$ has a unique representation on the hyperplane $\{w \in \mathcal{R}^n \mid \pi(w) = g(r_m)\}$, and each extreme direction of $E(\lambda)$ corresponds to a vertex of the bounded polyhedron defined as the solution set of the linear system

$$\begin{cases} -\overline{q}_j w_j &= 1 \quad \forall j \in I^-, \\ -\overline{q}_j w_j &\leq 1 \quad \forall j \in I^0, \\ \underline{q}_j w_j &\leq 1 \quad \forall j \in I^0, \\ \underline{q}_j w_j &= 1 \quad \forall j \in I^+, \\ \langle e, w\rangle &= 0, \\ \pi(w) &= g(r_m), \end{cases} \tag{31}$$

for which, the last equality can be removed as it is, by the definitions of π and g, a consequence of all the other equalities. It is rather straightforward to verify that, the set of vectors $\bar{w}$ described in statement (v) is nothing else but the set of vertices of the above bounded polyhedron. In particular when $I^0 = \{m\}$, the unique vertex of the above bounded polyhedron is $\bar{w} = h(r_m)(x^{m-1,m} - e_m)$ and the unique vertex of $E(\lambda)$ is $x^{m-1,m}$.

To show statement (v), we note that the vectors $x^{\xi(k)-1,\xi(k)} - e_j$ with $j \in I^0$ are linearly independent due to $|I^0| < n$, and that each $x \in E(\lambda)$ can be decomposed as

$$x = x^{\xi(k)-1,\xi(k)} + \sum_{j \in I^0} \omega(x) \left[\frac{1}{\omega(x^{\xi(k)-1,\xi(k)})} x^{\xi(k)-1,\xi(k)} - \frac{1}{\omega(x)} x \right]_j \left(x^{\xi(k)-1,\xi(k)} - e_j \right),$$

implying that $E(\lambda) - x^{\xi(k)-1,\xi(k)}$ is included in the cone

$$C_k =: \text{cone}\left\{ x^{k_{\min}-1,k_{\min}} - e_j,\, j \in I^0 \right\},$$

whose dimension is $|I^0|$. This entails that the dimension of $E(\lambda)$ is no more than $|I^0|$.

Now assuming that C' is a convex cone that contains $E(\lambda) - x^{\xi(k)-1,\,\xi(k)}$, we will show that $C_k \subset C'$. Let $j_0 \in I^0$. If $h(r_k) > \frac{1}{\bar{q}_{j_0}} + \frac{1}{\underline{q}_{j_0}}$, we have by definition

$$e_{j_0} = x(\theta) + \frac{h(r_k)}{\frac{1}{\bar{q}_{j_0}} + \frac{1}{\underline{q}_{j_0}}}\left(x^{\xi(k)-1,\,\xi(k)} - x(\theta)\right),$$

or equivalently,

$$\left[\frac{h(r_k)}{\frac{1}{\bar{q}_{j_0}} + \frac{1}{\underline{q}_{j_0}}} - 1\right]^{-1}\left(x^{\xi(k)-1,\,\xi(k)} - e_{j_0}\right) = x(\theta) - x^{\xi(k)-1,\,\xi(k)} \in E(\lambda) - x^{\xi(k)-1,\,\xi(k)} \subset C',$$

where $x(\theta)$ is a vertex of $E(\lambda)$ with $\theta \in \Theta$ satisfying $\theta_{j_0} = 1$ for some $j_0 \in I^0$ and $\theta_j = 0$ for all $j \in I^0 \backslash \{j_0\}$. Alternatively if $h(r_k) \le \frac{1}{\bar{q}_{j_0}} + \frac{1}{\underline{q}_{j_0}}$ (this is possible only when $r_k = r_m$), we have by definition

$$e_{j_0} = x^{\xi(k)-1,\,\xi(k)} - \frac{\bar{w}(\theta)}{h(r_m)},$$

or equivalently,

$$x^{\xi(k)-1,\,\xi(k)} - e_{j_0} = e_{j_0} + 2\frac{\bar{w}(\theta)}{h(r_m)} - x^{\xi(k)-1,\,\xi(k)} \in E(\lambda) - x^{\xi(k)-1,\,\xi(k)} \subset C',$$

where $\bar{w}(\theta)$ is an extreme direction of $E(\lambda)$ defined by (27) with $\theta \in \Theta$ satisfying $\theta_{j_0} = \frac{h(r_m)}{\frac{1}{\bar{q}_{j_0}} + \frac{1}{\underline{q}_{j_0}}}$ and $\theta_j = 0$ for all $j \in I^0 \backslash \{j_0\}$. This suggests that $e_j \in \mathrm{aff}(E(\lambda))$ and $x^{\xi(k)-1,\,\xi(k)} - e_j \in C'$ for all $j \in I^0$. Therefore, we have $C_k \subset C'$, implying that the smallest cone containing $E(\lambda) - x^{\xi(k)-1,\,\xi(k)}$ is C_k. Moreover, as the simplex $\mathrm{conv}\{x^{\xi(k)-1,\,\xi(k)}, e_j (j \in I^0)\}$ is included in $\mathrm{aff}(E(\lambda))$, the dimension of $E(\lambda)$ is clearly no less than $|I^0|$. This completes the proof. □

In view of (4), Theorem 3.1, Remark 3.2 and Corollary 3.1, we can summarize our main result on **EP** and **EF** in the following theorem whose proof are rather straightforward and thus omitted.

Theorem 3.2. *Let m' be the number of distinct r_j's with $r_j < r_m$. Then* **EP** *consists of m' bounded and one unbounded polyhedra, and accordingly,* **EF** *consists of m' closed line segments and one closed half-line. Explicitly, we have*

$$\mathbf{EP} = \bigcup_{k=1}^{m} E(g(r_k)),$$

and

$$\mathbf{EF} = \left\{(t,\ f(t))^T \mid t \ge \frac{1}{h(r_1)}\right\},$$

where various descriptions of the polyhedra $E(g(r_k))$ can be found in Corollary 3.1, and $x^{k-1,k}$ is the only common point of $E(g(r_{k-1}))$ and $E(g(r_k))$ whenever $r_{k-1} < r_k$, and

$$f(t) = \begin{cases} g(r_k)t + r_k & \text{if } t \in [\frac{1}{h(r_k)},\ \frac{1}{h(r_{k+1})}] \text{ with } k < m, \\ g(r_m)t + r_m & \text{if } t \in [\frac{1}{h(r_m)},\ +\infty). \end{cases} \tag{32}$$

In the case of $r_1 < r_2 < \cdots < r_m$, **EP** *consists of* $m-1$ *closed line segments and one closed half-line, i.e.,*

$$\mathbf{EP} = \bigcup_{k=1}^{m-1}\left[x^{k-1,k}, x^{k,k+1}\right] \bigcup \left\{x^{m-1,m} + t(x^{m-1,m} - e_m) \mid t \geq 0\right\}, \tag{33}$$

and **EF** *consists of* $m-1$ *closed line segments and one closed half-line, i.e.,*

$$\mathbf{EF} = \bigcup_{k=1}^{m-1}\left[\left(\frac{1}{h(r_k)}, \frac{g(r_k)}{h(r_k)} + r_k\right)^T, \quad \left(\frac{1}{h(r_{k+1})}, \frac{g(r_k)}{h(r_{k+1})} + r_k\right)^T\right]$$
$$\bigcup \left\{\left(\frac{1}{h(r_m)}, \frac{g(r_m)}{h(r_m)} + r_m\right)^T + t\,(1,\ g(r_m))^T \mid t \geq 0\right\}.$$

As noted in the previous theorem that **EP** consists of finitely many polyhedra, it is interesting to know how these polyhedra distribute in the hyperplane $\{x \in \mathcal{R}^n \mid x_1 + \cdots + x_n = 1\}$ and how to identify **EP** or its estimates by some sets having simple structures. This is done in the following theorem.

Theorem 3.3. *The dimension of* **EP** *is* $\min\{m, n-1\}$*, the smallest closed and convex set containing* **EP** *is the polyhedral set*

$$\operatorname{conv} V + \operatorname{cone} W,$$

and the smallest closed and convex cone containing $\mathbf{EP} - x^{0,1}$ *is the polyhedral cone*

$$\operatorname{cone}(V - x^{0,1}) + \operatorname{cone} W,$$

where V *consists of the distinct vertices of* $E(g(r_k))$ *for all* $k \in \{1, \cdots, m\}$ *and* W *consists of the extreme directions of* $E(g(r_m))$*. Moreover, we have*

$$\begin{aligned}
\mathbf{EP} &\subset \operatorname{conv} V + \operatorname{cone} W \\
&\subset x^{0,1} + \operatorname{cone}(V - x^{0,1}) + \operatorname{cone} W \\
&\subset x^{0,1} + \operatorname{cone}\{x^{\xi(1)-1,\xi(1)} - e_1, x^{\xi(2)-1,\xi(2)} - e_2, \cdots, x^{\xi(m)-1,\xi(m)} - e_m\} \\
&\subset x^{0,1} + \operatorname{cone}\{x^{0,1} - e_1, x^{0,1} - e_2, \cdots, x^{0,1} - e_m\},
\end{aligned} \tag{34}$$

where $\xi(k) =: \min\left\{j \mid r_j = r_k\right\}$.

Proof. To begin, we note that our proof in Corollary 3.1 shows that $e_j \in \operatorname{aff}(\mathbf{EP})$ for all $j = 1, \cdots, m$, and that each $x \in \mathbf{EP}$ can be decomposed as

$$x = x^{0,1} + \sum_{j=1}^{m} \omega(x)\left[\frac{1}{\omega(x^{0,1})}x^{0,1} - \frac{1}{\omega(x)}x\right]_j \left(x^{0,1} - e_j\right).$$

Thus, $\mathbf{EP} - x^{0,1}$ is included in the polyhedral cone

$$C^2 =: \operatorname{cone}\{x^{0,1} - e_1, \cdots, x^{0,1} - e_m\},$$

whose dimension is $\min\{m, n-1\}$. So the dimension of **EP** is no more than $\min\{m, n-1\}$. To show the reverse, we observe that aff(**EP**) contains, in the case of $m = n$, the simplex $\operatorname{conv}\{e_1, \cdots, e_n\}$ with dimension $n-1 = \min\{m, n-1\}$, while in the case of $m < n$, the simplex $\operatorname{conv}\{e_1, \cdots, e_m, x^{0,1}\}$ with dimension $m = \min\{m, n-1\}$. So the dimension of **EP** is $\min\{m, n-1\}$.

In view of Theorem 3.2 and Corollary 3.1, we have

$$\mathbf{EP} \subset \operatorname{conv} V + \operatorname{cone} W.$$

Let A be any closed and convex set containing **EP**. As $V \subset \mathbf{EP} \subset A$, we have $\operatorname{conv} V \subset A$. In view of Corollary 3.1, we have

$$x^{\xi(m)-1,\xi(m)} + \operatorname{cone} W \subset \mathbf{EP} \subset A,$$

and hence by [1, Theorem 3.6], $\operatorname{cone} W$ is included in the recession cone of A. Therefore, we have $\operatorname{conv} V + \operatorname{cone} W \subset A$. That is, the polyhedral set $\operatorname{conv} V + \operatorname{cone} W$ is the smallest closed and convex set containing **EP**. As $\operatorname{conv} V - x^{0,1} = \operatorname{conv}(V - x^{0,1}) \subset \operatorname{cone}(V - x^{0,1})$, we have

$$\mathbf{EP} - x^{0,1} \subset \operatorname{cone}(V - x^{0,1}) + \operatorname{cone} W.$$

Let B be any closed and convex cone containing $\mathbf{EP} - x^{0,1}$. As $V - x^{0,1} \subset \mathbf{EP} - x^{0,1} \subset B$, we have $\operatorname{cone}(V - x^{0,1}) \subset B$. In view of Corollary 3.1, we have

$$x^{\xi(m)-1,\xi(m)} - x^{0,1} + \operatorname{cone} W \subset \mathbf{EP} - x^{0,1} \subset B,$$

and hence by [1, Theorem 3.6], $\operatorname{cone} W$ is included in the recession cone of B, which is nothing else but B itself. So we actually have $\operatorname{cone}(V - x^{0,1}) + \operatorname{cone} W \subset B$. This implies that $\operatorname{cone}(V - x^{0,1}) + \operatorname{cone} W$ is the smallest closed and convex cone containing $\mathbf{EP} - x^{0,1}$.

To show the third inclusion in (34), let $\eta(k) =: \max\{j \mid r_j = r_k\}$ for each $k = 1, \cdots, m$. Let $v \in V$. Then there is some $k \in \{1, \cdots, m\}$ such that $v \in E(g(r_k))$. By Corollary 3.1, we have

$$v - x^{\xi(k)-1,\xi(k)} \in \operatorname{cone}\left\{x^{\xi(i)-1,\xi(i)} - e_i,\ i = \xi(k), \cdots, \eta(k)\right\},$$

and for all $j = 2, \cdots, k$,

$$x^{\xi(j)-1,\xi(j)} - x^{\xi(j-1)-1,\xi(j-1)} \in \operatorname{cone}\left\{x^{\xi(i)-1,\xi(i)} - e_i,\ i = \xi(j-1), \cdots, \eta(j-1)\right\}.$$

This implies that $v - x^{0,1} \in \operatorname{cone}\left\{x^{\xi(i)-1,\xi(i)} - e_i,\ i = 1, \cdots, \eta(k)\right\}$ because it holds that

$$\begin{aligned} v - x^{0,1} &= v - \sum_{j=2}^{k}\left(x^{\xi(j)-1,\xi(j)} - x^{\xi(j)-1,\xi(j)}\right) - x^{\xi(1)-1,\xi(1)} \\ &= v - x^{\xi(k)-1,\xi(k)} + \sum_{j=2}^{k}\left(x^{\xi(j)-1,\xi(j)} - x^{\xi(j-1)-1,\xi(j-1)}\right). \end{aligned}$$

Thus, we have $\operatorname{cone}(V - x^{0,1}) \subset \operatorname{cone}\left\{x^{\xi(1)-1,\xi(1)} - e_1, \cdots, x^{\xi(m)-1,\xi(m)} - e_m\right\}$. By Corollary 3.1 again, we have $\operatorname{cone} W \subset \operatorname{cone}\{x^{\xi(i)-1,\xi(i)} - e_i,\ i = \xi(m), \cdots, m\}$. That is, the third inclusion in (34) holds.

It remains to show the last inclusion in (34). Let $k \in \{1, \cdots, m\}$ and let $j \in \{\xi(k), \cdots, \eta(k)\}$. We have

$$x^{\xi(j)-1,\xi(j)} - x^{0,1} = \sum_{i=1}^{\xi(k)-1} \omega(x^{\xi(i)-1,\xi(i)}) \left[\frac{1}{\omega(x^{0,1})} x^{0,1} - \frac{1}{\omega(x^{\xi(i)-1,\xi(i)})} x^{\xi(i)-1,\xi(i)} \right]_i (x^{0,1} - e_i),$$

and thus

$$x^{\xi(j)-1,\xi(j)} - e_j \in \text{cone}\{x^{0,1} - e_1, \cdots, x^{0,1} - e_{\xi(k)-1}, x^{0,1} - e_j\}.$$

This verifies the last inclusion in (34). This completes the proof. □

Remark 3.3. *From Corollary 3.1 and Theorems 3.2 and 3.3, it can be seen that* **EP** *can be generated or estimated from some collections of finitely many efficient portfolios. This can be considered as a generalization of the well-known Two-fund Theorem in Merton [26].*

j	1	2	3	4	5	6	7
$\hat{r}_j$	0.0500	0.1000	0.1500	0.2000	0.2300	0.2500	0.2800
$\underline{\hat{q}}_j$	0.0268	0.0339	0.0395	0.0215	0.0692	0.0803	0.0848
$\overline{\hat{q}}_j$	0.0400	0.0550	0.0600	0.0742	0.0827	0.0900	0.0950
$h(\hat{r}_j)$	177.3367	115.0232	67.3429	25.3598	-34.6290	-61.1717	-84.7361
$g(\hat{r}_j)$	18.7874	13.0362	9.6691	8.4011	9.4400	10.6634	13.2055
$x_j^{0,1}$	0.2104	0.1663	0.1428	0.2623	0.0815	0.0702	0.0665
$x_j^{1,2}$	-0.2173	0.2565	0.2201	0.4044	0.1256	0.1083	0.1025
$x_j^{2,3}$	-0.3712	-0.2700	0.3759	0.6907	0.2146	0.1849	0.1751
$x_j^{3,4}$	-0.9858	-0.7170	-0.6572	1.8341	0.5698	0.4911	0.4650

Table 1: Estimated parameters, function values of g and h, and some key portfolios

4 Performance of Efficient and Inefficient Portfolios

In this section, by doing some Monte Carlo simulations, we will test the performance of some efficient portfolios versus some inefficient ones, all of which can be calculated using the formulas presented in last section.

As investment is a matter of standing in the present and looking at the future, it is impossible for an investor to know the true distribution information on the random return rates $\mathbb{R}_j$ so that the parameters $r_j := \mathrm{E}(\mathbb{R}_j)$, $\underline{q}_j := \mathcal{D}(\mathbb{R}_j)$, $\overline{q}_j := \mathcal{D}(-\mathbb{R}_j)$ required in our model cannot be calculated. However, these parameters can be estimated in one way or another, denoted respectively by $\hat{r}_j$, $\underline{\hat{q}}_j$ and $\overline{\hat{q}}_j$. Now assume that there are $n = 7$ possible assets to be invested, and that values of the estimated parameters $\hat{r}_j$, $\underline{\hat{q}_j}$ and $\overline{\hat{q}}_j$ are listed in Table 1. In applying the results and the formulas presented in last section, we use $\hat{r}_j$, $\underline{\hat{q}_j}$ and $\overline{\hat{q}}_j$ in place of r_j, $\underline{q}_j$ and $\overline{q}_j$, respectively. The values of $g(\hat{r}_j)$, $h(\hat{r}_j)$ and $x_j^{k-1,k}$ with $j = 1, \cdots, 7$ and $k = 1, \cdots, 4$ are also listed in Table 1. Then we have $m := \{j \mid h(\hat{r}_j) > 0\} = 4$. According to Theorem 3.2, we have

$$\mathbf{EP} = \bigcup_{k=1}^{3} \left[x^{k-1,k}, x^{k,k+1}\right] \bigcup \left\{x^{3,4} + t(x^{3,4} - e_4) \mid t \geq 0\right\},$$

and we can plot **EF** in the $\omega - \pi$ plane as shown in Figure 1.

Monte Carlo simulations will be done based on the assumption that the realization value of $\mathbb{R}_j$ is a random sampling on the interval $[\hat{r}_j - a\underline{\hat{q}}_j, \hat{r}_j + a\overline{\hat{q}}_j]$ for every $j = 1, 2, \ldots, n$, where a is a positive integer. We note that this assumption holds if $\mathbb{R}_j$ is an essentially bounded random variable, and its realization value falls into an interval based on $\hat{r}_j$ (point estimation). In these simulations, the smaller the value of a is, the more accurate the uncertainty is captured by the estimated $\underline{\hat{q}}_j$ and $\overline{\hat{q}}_j$. To evaluate the performance for a given portfolio x, we introduce the following performance indicators:

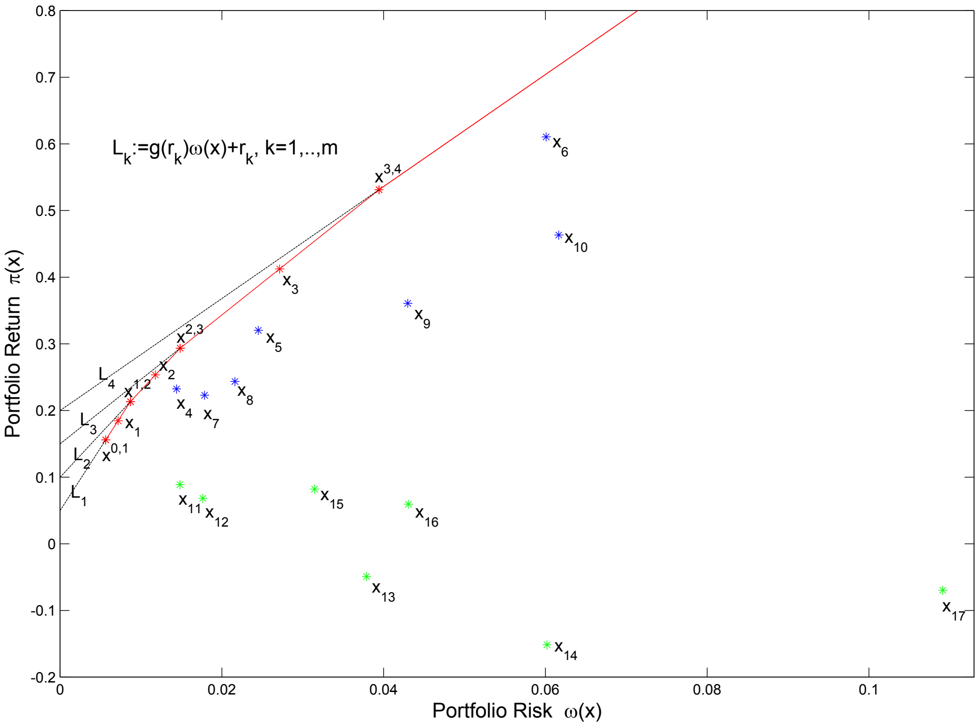

Figure 1: The efficient frontier and the portfolios for testing in the $\omega - \pi$ plane.

(1) $\text{rate}_{\pi(x)} := N_{\pi(x)}/N$, where $N := 100,000$ is the number of simulations, and $N_{\pi(x)}$ is the number of times that $\sum_{j=1}^{n} \mathbb{R}_j^t x_j$ is no less than $\pi(x)$,

(2) Sample mean: $\mu_x := (1/N) \sum_{t=1}^{N} (\sum_{j=1}^{n} \mathbb{R}_j^t x_j)$, where $\mathbb{R}_j^t$ $(t = 1, 2, \ldots, N)$ is the tth random sampling realization value of $\mathbb{R}_j$,

(3) Sample standard deviation: $\sigma_x := \sqrt{(1/(N-1)) \sum_{t=1}^{N} (\sum_{j=1}^{n} \mathbb{R}_j^t x_j - \mu_x)^2}$,

(4) μ_x/σ_x, the ratio between sample mean and sample standard deviation (following Park et al. [10]).

Thanks to Theorems 3.2 and 3.3, we are able to systematically select 7 efficient portfolios and 14 inefficient ones to joint the comparison group. The details are as follows.

(I) $\{x^{0,1}, x^{1,2}, x^{2,3}, x^{3,4}, x_1, x_2, x_3\} \subset$ **EP**. These efficient portfolios are marked by red dots in Figure 1;

(II) $\{x_4, \cdots, x_{10}\} \subset \left(x^{0,1} + \text{cone}\{x^{0,1} - e_1, \cdots, x^{0,1} - e_4\}\right) \backslash$**EP**. These inefficient portfolios are marked by blue dots in Figure 1;

(III) $\{x_{11}, \cdots, x_{17}\} \subset \{x \in \mathcal{R}^7 \mid x_1 + \cdots + x_7 = 1\} \backslash \left(x^{0,1} + \text{cone}\{x^{0,1} - e_1, x^{0,1} - e_2, \cdots, x^{0,1} - e_4\}\right)$. These inefficient portfolios are marked by green dots in Figure 1.

We show the performance of the selected portfolios in Table 2 ∼ Table 4. It can be seen from these tables that the lower the value of $\pi(x)$, the more frequently the portfolio can reach it, and this frequency is relatively less affected by the value of a. On the other hand, the greater the value of $\pi(x)$, the greater the μ_x and σ_x. Besides, as the value of a increases, μ_x and σ_x also increase, but whether μ_x/σ_x increases or decreases is uncertain. Moreover, these data also show that the performance of efficient portfolios is relatively stable, while the portfolios in the cone without in **EP** perform well when a is small. But as a increases, some of them perform worse. In addition, it is noted that the performance of the portfolios outside the cone are the most volatile. The most typical one is x_{17}. When the realization value of $\mathbb{R}_j$ appears in the interval $[\hat{r}_j - 5\underline{\hat{q}}_j, \hat{r}_j + 5\hat{\bar{q}}_j]$, its performance is particularly poor. But when the realization value of $\mathbb{R}_j$ appears in the interval $[\hat{r}_j - 30\underline{\hat{q}}_j, \hat{r}_j + 30\hat{\bar{q}}_j]$, its performance is quite impressive.

portfolio	$x^{0,1}$	x_1	$x^{1,2}$	x_2	$x^{2,3}$	x_3	$x^{3,4}$
$\pi(x)$	0.1559	0.1846	0.2133	0.2535	0.2936	0.4124	0.5313
$\text{rate}_{\pi(x)}$	0.8520	0.8377	0.8095	0.7788	0.7448	0.7029	0.6772
μ_x	0.2197	0.2567	0.2937	0.3435	0.3932	0.5521	0.7110
σ_x	0.0137	0.0150	0.0168	0.0188	0.0214	0.0280	0.0342
μ_x/σ_x	16.0061	17.1015	17.5067	18.2739	18.3648	19.7528	20.7845
portfolio	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}
$\pi(x)$	0.2325	0.3203	0.6105	0.2230	0.2436	0.3608	0.4633
$\text{rate}_{\pi(x)}$	0.7909	0.7254	0.6665	0.6959	0.6593	0.5623	0.5314
μ_x	0.3183	0.4264	0.8169	0.2792	0.2975	0.4016	0.4925
σ_x	0.0180	0.0232	0.0380	0.0183	0.0200	0.0284	0.0343
μ_x/σ_x	17.6648	18.3866	21.5249	15.2583	14.8741	14.1371	14.3590
portfolio	x_{11}	x_{12}	x_{13}	x_{14}	x_{15}	x_{16}	x_{17}
$\pi(x)$	0.0889	0.0683	-0.0490	-0.1514	0.0821	0.0595	-0.0696
$\text{rate}_{\pi(x)}$	0.7797	0.7446	0.6405	0.6066	0.7690	0.7518	0.7123
μ_x	0.1602	0.1420	0.0379	-0.0530	0.2388	0.2447	0.2781
σ_x	0.0168	0.0185	0.0273	0.0334	0.0257	0.0291	0.0438
μ_x/σ_x	9.5192	7.6538	1.3888	-1.5902	9.2972	8.4209	6.3487

Table 2: Performance of portfolios with realization values of $\mathbb{R}_j$ in $[\hat{r}_j - a\underline{\hat{q}}_j, \hat{r}_j + a\hat{\hat{q}}_j]$ with $a = 5$

portfolio	$x^{0,1}$	x_1	$x^{1,2}$	x_2	$x^{2,3}$	x_3	$x^{3,4}$
$\pi(x)$	0.1559	0.1846	0.2133	0.2535	0.2936	0.4124	0.5313
$\text{rate}_{\pi(x)}$	0.8538	0.8385	0.8117	0.7816	0.7479	0.7064	0.6806
μ_x	0.2836	0.3289	0.3743	0.4342	0.4941	0.6943	0.8945
σ_x	0.0194	0.0212	0.0237	0.0265	0.0302	0.0394	0.0482
μ_x/σ_x	14.6289	15.5263	15.8123	16.3687	16.3468	17.6134	18.5401
portfolio	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}
$\pi(x)$	0.2325	0.3203	0.6105	0.2230	0.2436	0.3608	0.4633
$\text{rate}_{\pi(x)}$	0.7932	0.7291	0.6698	0.6979	0.6619	0.5625	0.5314
μ_x	0.4045	0.5340	1.0280	0.3358	0.3519	0.4433	0.5232
σ_x	0.0254	0.0327	0.0535	0.0258	0.0282	0.0401	0.0484
μ_x/σ_x	15.9096	16.3109	19.2027	13.0046	12.4657	11.0527	10.8010
portfolio	x_{11}	x_{12}	x_{13}	x_{14}	x_{15}	x_{16}	x_{17}
$\pi(x)$	0.0889	0.0683	-0.0490	-0.1514	0.0821	0.0595	-0.0696
$\text{rate}_{\pi(x)}$	0.7798	0.7452	0.6400	0.6061	0.7692	0.7528	0.7134
μ_x	0.2313	0.2153	0.1238	0.0440	0.3959	0.4304	0.6270
σ_x	0.0238	0.0262	0.0386	0.0471	0.0363	0.0411	0.0619
μ_x/σ_x	9.7167	8.2069	3.2121	0.9330	10.8967	10.4728	10.1214

Table 3: Performance of portfolios with realization values of $\mathbb{R}_j$ in $[\hat{r}_j - a\underline{\hat{q}}_j, \hat{r}_j + a\hat{\hat{q}}_j]$ with $a = 10$

portfolio	$x^{0,1}$	x_1	$x^{1,2}$	x_2	$x^{2,3}$	x_3	$x^{3,4}$
$\pi(x)$	0.1559	0.1846	0.2133	0.2535	0.2936	0.4124	0.5313
$\text{rate}_{\pi(x)}$	0.8538	0.8385	0.8117	0.7816	0.7479	0.7064	0.6806
μ_x	0.5388	0.6175	0.6961	0.7956	0.8950	1.2580	1.6210
σ_x	0.0336	0.0367	0.0410	0.0459	0.0523	0.0683	0.0836
μ_x/σ_x	16.0489	16.8282	16.9798	17.3172	17.0973	18.4258	19.3974
portfolio	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}
$\pi(x)$	0.2325	0.3203	0.6105	0.2230	0.2436	0.3608	0.4633
$\text{rate}_{\pi(x)}$	0.7932	0.7291	0.6698	0.6979	0.6619	0.5625	0.5314
μ_x	0.7485	0.9614	1.8630	0.5615	0.5685	0.6082	0.6429
σ_x	0.0440	0.0567	0.0927	0.0447	0.0489	0.0695	0.0839
μ_x/σ_x	16.9979	16.9535	20.0917	12.5541	11.6278	8.7554	7.6635
portfolio	x_{11}	x_{12}	x_{13}	x_{14}	x_{15}	x_{16}	x_{17}
$\pi(x)$	0.0889	0.0683	-0.0490	-0.1514	0.0821	0.0595	-0.0696
$\text{rate}_{\pi(x)}$	0.7798	0.7452	0.6400	0.6061	0.7692	0.7528	0.7134
μ_x	0.5161	0.5092	0.4694	0.4348	1.0236	1.1724	2.0201
σ_x	0.0412	0.0454	0.0668	0.0816	0.0629	0.0712	0.1073
μ_x/σ_x	12.5180	11.2075	7.0297	5.3249	16.2635	16.4683	18.8280

Table 4: Performance of portfolios with realization values of $\mathbb{R}_j$ in $[\hat{r}_j - a\underline{\hat{q}}_j, \hat{r}_j + a\hat{\overline{q}}_j]$ with $a = 30$

5 Conclusions

In this paper, we revisited the bi-criteria portfolio optimization model where the short selling was permitted, and a trade-off was sought between the expected return rate of a portfolio and the maximum of the uncertainty measured by a general deviation measure for all the investments comprising a portfolio. We provided not only explicit analytical formulas for all the efficient portfolios, but also explored as a whole the set of all the efficient portfolios and its structure such as dimensionality and distribution. In particular, we generalized the classical Two-fund Theorem by providing some collections of finitely many efficient portfolios to generate or estimate the set of all the efficient portfolios. We also noticed that our efficient portfolios are almost the risk parity ones in the sense that the risks are allocated equally across the investments. In order to test the performance of the efficient portfolios versus the inefficient ones, we carried out Monte Carlo simulations based on the assumptions that the return rate is an essentially bounded random variable, and that its realization value falls into some interval based on the estimated expected return. It was shown by the simulation results that the performance of the efficient portfolios were more reliable than that of the inefficient ones. Finally, it should be noted that our efficient portfolios do not have sparsity, which can be improved in future work by building some new models.

Acknowledgements. The authors are grateful to the editor and the anonymous reviewer for their valuable comments and suggestions toward the improvement of this paper. Kaiwen Meng's work was supported in part by the National Natural Science Foundation of China (11671329, 12001445), the Fundamental Research Funds for the Central Universities (JBK1805001) and the Scientific

Research Fund of Sichuan Provincial Science and Technology Department (2021YJ0524). Xiaoqi Yang's work was supported in part by the Research Grants Council of Hong Kong (PolyU 152165/18E). Carisa Kwok Wai Yu's work was supported in part by grants from the Research Grants Council of the Hong Kong Special Administrative Region, China (UGC/FDS14/P02/17).

References

1. Rockafellar, R. T., Wets, R. J.-B: Variational Analysis. Springer-Verlag, Berlin (1998)
2. Rockafellar, R. T.: Convex Analysis. Princeton University Press, Princeton (1970)
3. Victor DeMiguel, Lorenzo Garlappi, and Raman Uppal: Optimal Versus Naive Diversification: How Inefficient is the $1/N$ Portfolio Strategy? Review of Financial Studies 22 (5), 1915-1953 (2009)
4. Cai, X. Q., K. L. Teo, X. Q. Yang, X. Y. Zhou: Portfolio optimization under a minimax rule. Management Science 46(7), 957-972 (2000)
5. Ehrgott, M.: Multicriteria Optimization. 2nd ed. Springer, Berlin (2005)
6. Markowitz, H. M.: Portfolio Selection. Journal of Finance. 7, 77-91 (1952)
7. Markowitz, H. M.: Portfolio selection, efficient diversification of investments. New York: Wiley. 7, 77-91 (1959)
8. Konno, H.: Piecewise Linear Risk Function and Portfolio Optimization, Journal of the Operations Research Society of Japan. 33, 139-156 (1990)
9. Konno, H., Yamazaki, H.: Mean-Absolute Deviation Portfolio Optimization Model and Its Application to Tokyo Stock Market, Management Science. 37, 519-531 (1991)
10. Park, S., Song, H., Lee, S.: Linear programing models for portfolio optimization using a benchmark, The European Journal of Finance. 25(4-6), 435-457 (2019)
11. Rockafellar R T, Uryasev S: Optimization of conditional value-at-risk, Journal of Risk. 2, 21-42 (2000)
12. R. Tyrrell Rockafellar, Stan Uryasev, Michael Zabarankin: Generalized deviations in risk analysis, Finance and Stochastics. 10, 51-74 (2006)
13. Rockafellar R T, Uryasev S, Zabarankin M: Master funds in portfolio analysis with general deviation measures. Journal of Banking & Finance. 30(2), 743-778 (2006)
14. Caçador S, Dias J M, Godinho P: Portfolio selection under uncertainty: a new methodology for computing relative-robust solutions, International Transactions in Operational Research. 28(3), 1296-1329 (2021)
15. Guo S, Ching W K: High-order Markov-switching portfolio selection with capital gain tax, Expert Systems with Applications. 165: 113915
16. Shadabfar M, Cheng L: Probabilistic approach for optimal portfolio selection using a hybrid Monte Carlo simulation and Markowitz model, Alexandria Engineering Journal. 59(5), 3381-3393 (2020)

17. Mavrotas G, Makryvelios E: Combining multiple criteria analysis, mathematical programming and Monte Carlo simulation to tackle uncertainty in Research and Development project portfolio selection: A case study from Greece, European Journal of Operational Research. 291(2), 794-806 (2021)

18. Kizys R, Juan A A, Sawik B, et al: A biased-randomized iterated local search algorithm for rich portfolio optimization, Applied Sciences. 9(17), 3509 (2019)

19. Sawik B T: Conditional value-at-risk vs. value-at-risk to multi-objective portfolio optimization, Applications of management science. Emerald Group Publishing Limited. (2012)

20. Sawik B: Downside risk approach for multi-objective portfolio optimization, Operations research proceedings. Springer, Berlin, Heidelberg. 191-196 (2012)

21. Ogryczak W, Ruszczyski A: From stochastic dominance to mean-risk models: Semideviations as risk measures, European Journal of Operational Research. 116(1), 33-50 (1999)

22. Artzner P, Delbaen F, Eber J M, et al: Coherent measures of risk, Mathematical Finance. 9(3), 203-228 (1999)

23. De Giorgi E: RewardCrisk portfolio selection and stochastic dominance, Journal of Banking & Finance. 29(4), 895-926 (2005)

24. Young M R: A minimax portfolio selection rule with linear programming solution, Management Science. 44(5), 673-683 (1998)

25. Kolm P N, Ttnc R, Fabozzi F J: 60 Years of portfolio optimization: Practical challenges and current trends, European Journal of Operational Research. 234(2), 356-371 (2014)

26. Merton R C: An analytic derivation of the efficient portfolio frontier, Journal of Financial and Quantitative Analysis. 1851-1872 (1972)

27. Bertsimas D, Pachamanova D: Robust multiperiod portfolio management in the presence of transaction costs, Computers & Operations Research. 35(1), 3-17 (2008)

28. Qi Y, Steuer R E, Wimmer M: An analytical derivation of the efficient surface in portfolio selection with three criteria, Annals of Operations Research. 251(1-2), 161-177 (2017)