RAIRO-Oper. Res. 56 (2022) 2367–2387
https://doi.org/10.1051/ro/2022114

RAIRO Operations Research
www.rairo-ro.org

ESTIMATION OF PORTFOLIO EFFICIENCY *VIA* STOCHASTIC DEA

Helu Xiao[1], Xin Liu[1], Tiantian Ren[2,3,*] and Zhongbao Zhou[2,3]

Abstract. Traditional data envelopment analysis (DEA) and diversification DEA are two common data-driven evaluation approaches, which have been widely used in the estimation of portfolio efficiency. The above two DEA approaches usually use the risk and expected return indicators to build the input-output process of portfolios. However, this input-output process derived from the risk and expected return is inconsistent with the actual investment process, since the real input should be the initial wealth, and the output should be the terminal wealth. To address this problem, we propose a novel input-output process based on the initial and terminal wealth of portfolios. We transform the terminal wealth into the rate of return and construct a stochastic attainable set by using portfolio returns. We provide three deterministic estimation approaches to deal with the stochastic attainable set, and then obtain three deterministic attainable sets. We further propose three stochastic DEA models to estimate the portfolio efficiency by using the above three deterministic attainable sets. Finally, we provide an empirical analysis to assess the portfolio efficiency of 50 open-ended funds in China. The results show that there are some differences in the portfolio efficiency and its ranking between the proposed DEA models and the existing DEA models, which further verify the rationality of the proposed DEA models.

Mathematics Subject Classification. 90B30, 90B50, 90C05, 90C30.

Received May 20, 2021. Accepted June 28, 2022.

1. Introduction

Nowadays, the estimation of portfolio efficiency has always been one of the research hotspots in the field of finance. How to estimate the portfolio efficiency has also been a concern of many researchers. DEA approach, as a data-driven evaluation approach, has been considered as an effective tool to solve the efficiency evaluation problem, because it does not need to specify the form of production function. This approach cannot only avoid the setting error of production function, but also provide a benchmark for decision-makers. In recent years, DEA has been widely used in the estimation of portfolio efficiency, among which the traditional DEA and diversification DEA models are two common methods to solve this problem. The above two kinds of evaluation methods need not make assumptions about the effectiveness of the capital market, but only need to use the multi-dimensional financial indicators (*e.g.*, transaction costs, and the expected return and risk indicators) to make a relative assessment of portfolios. To the best of our knowledge, the existing researches usually use the above

Keywords. Data envelopment analysis, portfolio efficiency, chance-constrained theory, mean-standard deviation criterion.

[1] School of Business, Hunan Normal University, Changsha 410081, P.R. China.

[2] School of Business Administration, Hunan University, Changsha 410082, P.R. China.

[3] Institute of Data Science and Decision Optimization, Hunan University, Changsha 410082, P.R. China.

[*]Corresponding author: `ttren@hnu.edu.cn` (T.Ren); `xiaohelu1986@163.com` (H.Xiao)

financial indicators to build the production possibility set of portfolios. However, the resulting input-output process is also questioned by many researchers, since they believe that these financial indicators cannot fully describe the actual input-output of portfolios (*e.g.*, [25,26,31]). Therefore, how to define the actual input-output of portfolios is an urgent problem to be solved in this paper.

In recent years, the traditional DEA from production theory has been gradually and partially transposed to the analysis of portfolio performance by regarding these financial indicators as the inputs and outputs. To the best of our knowledge, Murthi *et al.* [23] first applied traditional DEA to estimate the portfolio efficiency and then treated the standard deviation of fund return as an input and the expected return as an output to construct the corresponding DEA model. Basso and Funari [1] took different risk indicators as inputs and the expected return as an output, and also showed the relationship between the proposed DEA model and the traditional evaluation index. Chen and Lin [10] used VaR and CVaR instead of the traditional variance measurement to measure the portfolio's risk and then used them as the input indicators of the proposed DEA model. Basso and Funari [2] proposed a social responsibility measurement index for the social responsibility investment funds, combined the traditional expected return and risk indicators and constructed a DEA model suitable for the evaluation of social responsibility investment funds. Ding *et al.* [14] studied the applicability of DEA in the performance evaluation of portfolios with margin requirements. Liu *et al.* [20] proved the convergence of DEA frontier under the generalized mean-risk framework. Basso and Funari [3] studied the relationship between fund size and fund efficiency in combination with DEA and econometric models. Chen *et al.* [11] proposed three kinds of DEA models to evaluate the performance of portfolios with fuzzy returns. Zhou *et al.* [30] provided a DEA frontier improvement approach which can effectively improve the convergence speed of DEA frontier. Allevi *et al.* [4] used a DEA approach to evaluate the performance of green mutual funds, among which the environmental and financial indicators are both considered into the proposed DEA models.

Traditional DEA can be converted into a linear programming problem, in essence, and its calculation is relatively simple, but many researchers have questioned the ability of traditional DEA in diversifying risks. Since Markowitz [22] proposed that the portfolio can diversify risks, there is an ever-growing literature on the different diversification DEA models to estimate the portfolio efficiency. Briec *et al.* [8] constructed a diversification DEA model under the classical mean-variance criterion, and also distinguished the portfolio efficiency and the allocation efficiency. Joro and Na [16] proposed a diversification DEA model within the mean-variance-skewness framework, and then investigated the influence of skewness on the portfolio efficiency. Inspired by the previous work, Briec *et al.* [9] further developed a diversification DEA model with consideration of the high-order moment constraints, and also theoretically proved the global optimal solution availability of the proposed model. Lamb and Tee [17] introduced a general class of diversification DEA models based on the multi-return and multi-risk measures, and then systematically examined the relationship between the portfolio diversification, coherent risk measurement and stochastic dominance. Branda [7] constructed several kinds of diversification DEA models under the generalized mean-risk criterion, and also allowed that the negative values of input and output indicators exist in these models. Lin and Li [18] proposed a super-efficiency diversification DEA model based on the directional distance function measure to further distinguish the differences between the effective investment funds.

It is not difficult to find that the production possibility sets of the above DEA models are constructed by using the risk and return indicators. In the existing literature, the input-output property of the return and risk mainly includes the following two viewpoints. The first viewpoint treats the portfolio risk as an input indicator and the expected return as an output indicator. The second viewpoint considers that the expected return and risk should be both treated as output indicators, because the expected return and risk are two evaluation indicators derived from the portfolio return. However, the above two input-output assumptions are inconsistent with the actual investment process. In fact, it is more reasonable to construct the input-output process of portfolios by using the assumption that the initial wealth is an input and the terminal wealth is an output. This is also a novel input-output assumption which is different from the existing literature. Since the terminal wealth of a portfolio is a random variable on the portfolio return, thus the resulting production possibility set may also be random. It is also called as the stochastic production possibility set. The stochastic production possibility set can be treated as an extension of the traditional deterministic production possibility set. Obviously, it is unrealistic to estimate the portfolio efficiency directly based on the above stochastic production possibility set. Therefore, how to make a deterministic estimation of this stochastic set is also the fundamental problem to be solved.

To the best of our knowledge, the common method to deal with the stochastic production possibility set is using the classical chance-constrained theory to construct the corresponding chance-constrained DEA models. The related studies include Huang and Li [15], Cooper *et al.* [13], Tsionas and Papadakis [27], Olesen and Petersen [24] and so on. Obviously, the chance-constrained theory can be used to estimate the stochastic production possibility set of portfolios. In addition to this, the classical mean-risk portfolio optimization theory can also be used to derive the estimation of stochastic production possibility set. In this framework, we can use the return and risk measures (*e.g.*, the expectation and standard deviation measures) to estimate the stochastic constraint conditions in the stochastic set mentioned above. In this paper, we will investigate the deterministic estimation of the stochastic production possibility set of portfolios by using the chance-constrained and mean-standard deviation theories.

Under the proposed input-output process based on the initial and terminal wealth, the difficulty in acquiring the initial wealth of portfolios will make it difficult to estimate the portfolio efficiency. For the above reasons, we standardize the initial and terminal wealth of portfolios, and then the initial wealth can be normalized to one unit wealth and the terminal wealth can be converted into the portfolio return. Generally speaking, the portfolio return (*e.g.*, fund return) is usually observable. Therefore, it is more practical to estimate the portfolio efficiency based on portfolio returns. As the characteristics of data described in Liu *et al.* [19], the portfolio returns can be treated as the data without explicit inputs. And they proposed a DEA-WEI model to deal with this problem, in which the proposed model needs not know the specific values of inputs, but only needs to use the index data (*e.g.*, the data of fund returns). Yang *et al.* [28] further proposed a DEA-WEI model with quadratic utility terms, and also compared the difference between the benchmark obtained by the proposed DEA model and that obtained by the traditional DEA model. However, the above studies on the DEA-WEI models all assume that the outputs of decision-making units are the deterministic data, while the data of portfolio returns studied in this paper is still stochastic one. Therefore, how to integrate the DEA-WEI model and stochastic DEA is also the rest of our work.

Motivated by the aforementioned research line, we first define the input-output process of portfolios, and then construct the stochastic attainable set without explicit inputs. Under the framework of the chance-constrained and mean-standard deviation theories, we propose three deterministic estimation approaches to estimate the above stochastic attainable set. Then, we construct three kinds of stochastic DEA models based on the direction distance function measure, *i.e.*, chance-constrained stochastic DEA model, and nonlinear and linear stochastic DEA models under mean-standard deviation criterion. For the proposed chance-constrained DEA models, we assume that the portfolio returns follow a joint normal distribution, and then transform the chance-constrained DEA models into the deterministic quadratic programming problems. Further, we investigate the difference between the proposed stochastic DEA models under mean-standard deviation criterion and the existing DEA models. Finally, we apply the proposed stochastic DEA models to assess the portfolio efficiency of 50 growth open-ended funds in China. The empirical results show that the proposed stochastic DEA models are different from the traditional DEA models in the portfolio efficiency and its ranking. These further validate the rationality of the proposed stochastic DEA models.

The remainder of this paper is organized as follows. In Section 2, we redefine the input-output process of portfolios, and then construct a stochastic attainable set by using portfolio returns. In Section 3, we proposed three deterministic estimation approaches to estimate the stochastic attainable set of portfolios. In Section 4, we provide three stochastic DEA models to estimate the portfolio efficiency. In Section 5, we carry out an empirical analysis to verify the validity of the proposed stochastic DEA models, and then show the difference between the proposed DEA models and the existing DEA models. Finally, some concluding remarks are summarized here.

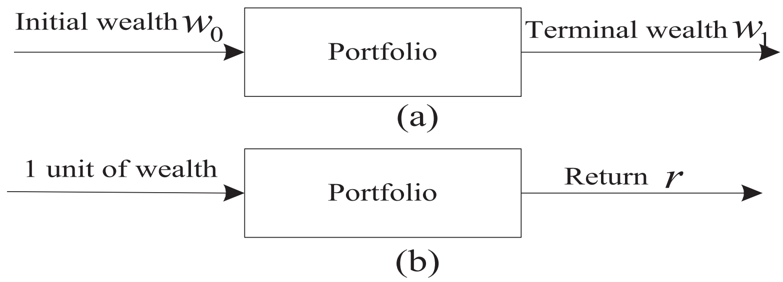

FIGURE 1. The input-output process of portfolios.

2. Revisiting the input-output process of portfolios

Suppose that there are m portfolios to be evaluated in the financial market. Let r_j denote the random return of the j-th portfolio, where $j = 1, \cdots, m$. Let $\bar{r}_j = E(r_j)$ and $\sigma(r_j) = \sqrt{\mathrm{Var}(r_j)}$ be the expectation and standard deviation of the j-th portfolio return, respectively, and $\Omega_{i,j} = \mathrm{Cov}(r_i, r_j)$ represent the covariance of portfolio returns r_i and r_j, where $i, j = 1, \cdots, m$. Assume that $\bar{\mathbf{r}} = [\bar{r}_1, \cdots, \bar{r}_m]'$ is the vector of expected returns and Ω is the covariance matrix of portfolio return vector $\mathbf{r} = [r_1, \cdots, r_m]'$. Based on the above m portfolios to be evaluated, this paper aims to estimate the portfolio efficiency based on production theory, and it is crucial to specify the input-output process of portfolios. Intuitively, for an investment activity in a financial asset, the input should be the initial wealth (*i.e.*, various costs) required to buy a certain amount of the asset and the output should be the gain received when the asset is sold. To this end, we let w_0 and w_1 be the initial and terminal wealth of a portfolio, respectively. Note that, w_0 is a deterministic value, while w_1 is a random value on the portfolio return. Therefore, for an actual investment process, we can conclude that the real input is the initial wealth and the output is the terminal wealth. In this case, the details about the input-output process of portfolios are shown in Figure 1a.

As shown in Figure 1a, for an investor, the input is the amount of money spent to purchase the portfolio (*i.e.*, initial wealth shown in Figure 1a, which includes transaction costs), and the output is the wealth that the investor will receive when the portfolio is sold (*i.e.*, the terminal wealth shown in Fig. 1a). In practice, the initial wealth of most portfolios (*e.g.*, investment funds) is not always observable, which is also inconvenient for evaluators. However, the initial wealth and the terminal wealth of a portfolio satisfy the following relationship: $w_1 = w_0 \times (1+r)$, where r denotes the portfolio return. Then, the initial wealth of a portfolio can be normalized, that is, its terminal wealth can be replaced by the portfolio return r. In this situation, the input of each portfolio can be unified as 1 unit wealth, and the output is the random return r. Thus, the input-output process of portfolios can be described by the portfolio return, and the detailed conversion process is shown in Figure 1b. With the input-output setup shown in Figure 1b, the evaluation of different portfolios can be understood as the investor spends 1 unit of wealth on each portfolio and ultimately compares which portfolio has a higher terminal return. Compared with the existing studies (*i.e.*, the DEA portfolio evaluation studies that use the return measure and risk measure of the portfolio return to construct the input-output process of portfolios), it is clear that the input-output process of portfolio defined in Figure 1 is a new understanding.

In this paper, we aims to build the DEA models based on the input-output process of portfolios in Figure 1b. Liu *et al.* [19] treated the production process shown in Figure 1b as a production process without explicit inputs, and then proposed a DEA-WEI model to deal with this problem. However, Liu *et al.* [19] assumed that the outputs of decision-making units are all deterministic values, while the portfolio's output r is a random variable. Inspired by the work of Cooper *et al.* [12], we use the idea of stochastic DEA to construct the stochastic attainable set without explicit inputs. Based on the above m portfolios to be evaluated and their portfolio returns (*i.e.*, r_j, $j = 1, \cdots, m$), we can construct the following stochastic attainable set by using the assumptions of convexity and strong free disposability.

$$\text{PPS} = \left\{ r \middle| \sum_{j=1}^{m} \lambda_j r_j \geq r,\ \sum_{j=1}^{m} \lambda_j = 1,\ \lambda_j \geq 0,\ j = 1, \cdots, m \right\}, \tag{2.1}$$

where λ_j denotes the intensity weight, which is used to construct the returns of some virtual portfolios based on the observation portfolios, and r denotes the virtual portfolio return that can be generated based on the observed portfolios.

From the expression of Set (2.1), we can find that Set (2.1) can be treated as the stochastic version of the attainable set provided by Liu *et al.* [19]. Due to the randomness of the portfolio return r_j, it is difficult for evaluators to directly assess the portfolio efficiency within the framework of the stochastic attainable set (2.1). Obviously, how to reasonably estimate the stochastic constraint in the stochastic attainable set (2.1) is the premise for building an effective DEA model. Under the framework of stochastic DEA, we propose three methods to construct the deterministic estimation of the above stochastic constraint, so as to create the corresponding stochastic DEA model to assess the portfolio efficiency. Further, we analyze the differences and connections between the proposed stochastic DEA models and the existing deterministic DEA models (*i.e.*, traditional DEA and diversification DEA models).

3. The deterministic estimation of stochastic attainable set

In this section, we first use the chance-constrained theory to deal with the above stochastic attainable set (2.1). In this case, the stochastic constraint shown in Set (2.1) only needs to be true at a given probability level $1-\alpha(0<\alpha\leq 1)$. Then, the corresponding deterministic estimation of the stochastic attainable set (2.1) can be expressed as

$$\text{PPS}^{(1)} = \left\{ r \middle| \begin{array}{l} P\left(\sum_{j=1}^{m} \lambda_j r_j \geq r\right) \geq 1-\alpha,\ 0<\alpha\leq 1 \\ \sum_{j=1}^{m} \lambda_j = 1,\ \lambda_j \geq 0,\ \ j = 1, \cdots, m. \end{array} \right\} \tag{3.1}$$

It is easy to find that $\text{PPS}^{(1)}$ denotes the set of the virtual portfolio returns (*i.e.*, r) that satisfy the above chance-constrained condition. Inspired by the classical mean-standard deviation portfolio optimization theory, we also provide a second method to deal with Set (2.1). Under the mean-standard deviation framework, we quantify the stochastic constraint in Set (2.1) by using the mean and standard deviation measures. Under the risk averse preference assumption, investors want the mean of portfolio return to be larger and the standard deviation of portfolio return to be smaller. Similar to the assumption using in Set (2.1), we consider that the mean and standard deviation of portfolio return satisfy the strong free disposability[1]. Then, the corresponding deterministic estimation can be described as

$$\text{PPS}^{(2)} = \left\{ r \middle| \begin{array}{l} E\left(\sum_{j=1}^{m} \lambda_j r_j\right) \geq E(r),\ \ \sigma\left(\sum_{j=1}^{m} \lambda_j r_j\right) \leq \sigma(r), \\ \sum_{j=1}^{m} \lambda_j = 1,\ \lambda_j \geq 0,\ \ j = 1, \cdots, m. \end{array} \right\} \tag{3.2}$$

Difference from $\text{PPS}^{(1)}$, $\text{PPS}^{(2)}$ denotes the set of the virtual portfolio returns (*i.e.*, r) that meet the above mean-standard deviation framework. Referring to the traditional DEA models, the third method is to deal with the stochastic constraint in Set (2.1) by linearizing the nonlinear constraints shown in $\text{PPS}^{(2)}$. In this situation, the corresponding deterministic estimation can be expressed as

$$\text{PPS}^{(3)} = \left\{ r \middle| \begin{array}{l} \sum_{j=1}^{m} \lambda_j E(r_j) \geq E(r),\ \ \sum_{j=1}^{m} \lambda_j \sigma(r_j) \leq \sigma(r), \\ \sum_{j=1}^{m} \lambda_j = 1,\ \lambda_j \geq 0,\ \ j = 1, \cdots, m. \end{array} \right\} \tag{3.3}$$

[1]The strong free disposability implies that virtual portfolios with smaller mean and larger variance than observed portfolios are achievable. The question of whether the strong free disposability of DEA is applicable to the performance evaluation of portfolios is more systematically analyzed by Lamb and Tee [17], who argue that the strong free disposability can still be retained in portfolio evaluation except for some non-convex risk measures, such as value-at-risk (VaR) and third-order moment measures.

The relationship between $\text{PPS}^{(2)}$ and $\text{PPS}^{(3)}$ is that the latter can be consider as a linear estimation of the former, since the conclusions of $E\left(\sum_{j=1}^{m}\lambda_j r_j\right) = \sum_{j=1}^{m}\lambda_j E(r_j)$ and $\sigma\left(\sum_{j=1}^{m}\lambda_j r_j\right) \le \sum_{j=1}^{m}\lambda_j \sigma(r_j)$ hold. Although $\text{PPS}^{(3)}$ ignores the risk diversification, the resulting linear model is more convenient to be solved, which is consistent with the characteristic of traditional DEA. Moreover, it follows from the conclusion of Liu *et al.* [20] that as the portfolio sample size increases, $\text{PPS}^{(3)}$ converges to $\text{PPS}^{(2)}$ with probability one. Therefore, it is also worth considering $\text{PPS}^{(3)}$ as a feasible estimation of Set (3.1).

In the following, this paper will construct the corresponding stochastic DEA model based on the above deterministic estimations, *i.e.*, $\text{PPS}^{(1)}$, $\text{PPS}^{(2)}$ and $\text{PPS}^{(3)}$.

4. Estimation of portfolio efficiency via stochastic DEA

Because of the negative value of portfolio returns, we will construct the corresponding stochastic DEA model based on the directional distance function. For any portfolio return r_d, we assume that, $r_d + \theta g_r \in \text{PPS}^{(1)}$, $r_d + \theta g_r \in \text{PPS}^{(2)}$ and $r_d + \theta g_r \in \text{PPS}^{(3)}$, where $g_r > 0$, and θ denotes the proportion of the portfolio return that can be expanded within different attainable sets. Under the three scenarios described above, the investor wants to find the maximum expansion proportion under a given attainable set and treats it as a measure of the scope for improvement in the portfolio to be evaluated. Then, we have the following evaluation models: $\theta_1^*(\alpha) = \max_\theta\left\{\theta | r_d + \theta g_r \in \text{PPS}^{(1)}\right\}$, $\theta_2^* = \max_\theta\left\{\theta | r_d + \theta g_r \in \text{PPS}^{(2)}\right\}$ and $\theta_3^* = \max_\theta\left\{\theta | r_d + \theta g_r \in \text{PPS}^{(3)}\right\}$. The detailed forms of the above stochastic DEA models are shown in Sections 4.1–4.3.

4.1. Chance-constrained stochastic DEA model

For the first evaluation model $\theta_1^*(\alpha) = \max_\theta\left\{\theta | r_d + \theta g_r \in \text{PPS}^{(1)}\right\}$, the details can be expressed as follows.

$$\begin{aligned} &\theta_1^*(\alpha) = \max \theta \\ &\text{s.t.} \begin{cases} P\left\{\sum_{j=1}^{m}\lambda_j r_j \ge r_d + \theta g_r\right\} \ge 1-\alpha, \\ \sum_{j=1}^{m}\lambda_j = 1,\ \lambda_j \ge 0,\ \theta \ge 0,\ j = 1, \cdots, m. \end{cases} \end{aligned} \tag{4.1}$$

Here α is a predetermined parameter, and it satisfies that $0 < \alpha \le 1$. Parameter α can be treated as the modeler's risk level, indicating the probability measure of the extent to which chance constraint violation is admitted. According to Model (4.1), we can find that, the lower the value of α, the higher the modeler's confidence about the portfolio being evaluated and the lower the modeler's risk, and *vice versa*. Referring to Zhou *et al.* [5], in most cases, the modeler's confidence is relatively high. In the following, we further assume that the predetermined parameter α satisfies $\alpha \le 0.5$.

Due to the existence of chance-constrained condition in Model (4.1), it is difficult for decision-makers to directly solve it. The traditional method of solving this kind of problem is to assume that the random output follows a given distribution function, and then transforms the chance constraint into the deterministic equivalent constraint. In this case, the above chance-constrained DEA model can be transformed into the corresponding equivalent model. Since the equivalent model only depends on the distribution of portfolio returns, we can estimate the distribution parameters of portfolio returns by using the historical data. Then, the portfolio efficiency can be assessed by using the equivalent model.

In the following, we assume that the portfolio returns follow a joint normal distribution, that is, $\mathbf{r} \sim N(\bar{\mathbf{r}},\ \Omega)$, where $\mathbf{r} = [r_1, \cdots, r_m]'$. Without loss of generality, this paper assumes that the direction g_r can be respectively valued in deterministic direction and exogenous random direction, and then Model (4.1) can be transformed into the corresponding deterministic equivalent model[2]. The relevant conclusions are shown in Theorems 4.1 and 4.2.

[2]In fact, whether in a bear market or a bull market, investors always hope that the more wealth, the better, so the proposed exogenous random direction can be considered as an idealized benchmark in the investors' heart, which is an improvement direction unrelated to the portfolios in the market. In addition, it is worth noting that the deterministic equivalent transformation of Model

Theorem 4.1. *Suppose that the portfolio returns satisfy $\mathbf{r} \sim N(\bar{\mathbf{r}},\ \Omega)$ and the direction g_r is a deterministic value, and then Model (4.1) can be equivalent to the following deterministic programming problem.*

$$\theta_{11}^*(\alpha) = \max\ \theta$$
$$\text{s.t.} \begin{cases} \sum_{j=1}^m \lambda_j E(r_j) + \Phi^{-1}(\alpha)\delta(\boldsymbol{\lambda}) \geq E(r_d) + \theta g_r, \\ \delta(\boldsymbol{\lambda}) = \sqrt{\mathrm{Var}\Big(\sum_{j=1}^m \lambda_j r_j\Big) + \mathrm{Var}(r_d) - 2\sum_{j=1}^n \lambda_j \mathrm{Cov}(r_j, r_d)}, \\ \sum_{j=1}^m \lambda_j = 1,\ \lambda_j \geq 0,\ \theta \geq 0,\ j = 1, \cdots, m \end{cases} \tag{4.2}$$

where $\boldsymbol{\lambda} = (\lambda_1, \cdots, \lambda_m)'$.

Proof. The chance constraint shown in Model (4.1), *i.e.*, $P\Big\{\sum_{j=1}^m \lambda_j r_j \geq r_d + \theta g_r\Big\} \geq 1-\alpha$, can be transformed as follows.

$$P\left\{\sum_{j=1}^m \lambda_j r_j \geq r_d + \theta g_r\right\} = P\left\{\begin{array}{l} \frac{\sum_{j=1}^m \lambda_j r_j - r_d - \theta g_r - \left[\sum_{j=1}^m \lambda_j E(r_j) - E(r_d) - \theta g_r\right]}{\sqrt{\mathrm{Var}\left(\sum_{j=1}^m \lambda_j r_j\right) + \mathrm{Var}(r_d) - 2\sum_{j=1}^m \lambda_j \mathrm{Cov}(r_j, r_d)}} \\ \geq \frac{-\left[\sum_{j=1}^m \lambda_j E(r_j) - E(r_d) - \theta g_r\right]}{\sqrt{\mathrm{Var}\left(\sum_{j=1}^m \lambda_j r_j\right) + \mathrm{Var}(r_d) - 2\sum_{j=1}^n \lambda_j \mathrm{Cov}(r_j, r_d)}} \end{array}\right\} \geq 1 - \alpha \tag{4.3}$$

when the portfolio returns meet that $\mathbf{r} \sim N(\bar{\mathbf{r}},\ \Omega)$, we can conclude that the random variable $\frac{\sum_{j=1}^m \lambda_j r_j - r_d - \theta g_r - \left[\sum_{j=1}^m \lambda_j E(r_j) - \theta E(r_d) - \theta g_r\right]}{\sqrt{\mathrm{Var}\left(\sum_{j=1}^m \lambda_j r_j\right) + \mathrm{Var}(r_d) - 2\sum_{j=1}^m \lambda_j \mathrm{Cov}(r_j, r_d)}}$ follows the standard normal distribution. Then, the chance constraint condition (4.3) can be equivalent to the following inequality condition.

$$\sum_{j=1}^m \lambda_j E(r_j) + \Phi^{-1}(\alpha)\sqrt{\mathrm{Var}\left(\sum_{j=1}^m \lambda_j r_j\right) + \mathrm{Var}(r_d) - 2\sum_{j=1}^n \lambda_j \mathrm{Cov}(r_j, r_d)} \geq E(r_d) + \theta g_r. \tag{4.4}$$

Further, let $\delta(\boldsymbol{\lambda}) = \sqrt{\mathrm{Var}\Big(\sum_{j=1}^m \lambda_j r_j\Big) + \mathrm{Var}(r_d) - 2\sum_{j=1}^n \lambda_j \mathrm{Cov}(r_j, r_d)}$, then Model (4.1) can be transformed as

$$\theta_{11}^*(\alpha) = \max\ \theta$$
$$\text{s.t.} \begin{cases} \sum_{j=1}^m \lambda_j E(r_j) + \Phi^{-1}(\alpha)\delta(\boldsymbol{\lambda}) \geq E(r_d) + \theta g_r, \\ \sum_{j=1}^m \lambda_j = 1,\ \lambda_j \geq 0,\ \theta \geq 0,\ j = 1, \cdots, m. \end{cases} \tag{4.5}$$

To sum up, we can conclude that Theorem 4.1 is true here. □

Theorem 4.2. *Assume that the portfolio return vector $\mathbf{r} \sim N(\bar{\mathbf{r}},\ \Omega)$ and the direction g_r is an exogenous random variable, and then Model (4.1) can also be equivalent to the following deterministic programming problem.*

$$\theta_{12}^*(\alpha) = \max\ \theta$$
$$\text{s.t.} \begin{cases} \sum_{j=1}^m \lambda_j E(r_j) + \Phi^{-1}(\alpha)\delta(\boldsymbol{\lambda},\ \theta) \geq E(r_d) + \theta E(g_r), \\ \delta(\boldsymbol{\lambda},\ \theta) = \sqrt{\mathrm{Var}\Big(\sum_{j=1}^m \lambda_j r_j\Big) + \mathrm{Var}(r_d) + \theta^2 \mathrm{Var}(g_r) - 2\sum_{j=1}^n \lambda_j \mathrm{Cov}(r_j, r_d)}, \\ \sum_{j=1}^m \lambda_j = 1,\ \lambda_j \geq 0,\ \theta \geq 0,\ j = 1, \cdots, m. \end{cases} \tag{4.6}$$

Here $\boldsymbol{\lambda} = (\lambda_1, \cdots, \lambda_m)'$.

(4.1) only needs to assume that the portfolio return vector follows a continuous distribution (see, [5]), but for comparison with the traditional DEA and diversification DEA models in the mean-standard deviation framework, we suppose that the portfolio return vector obeys a joint normal distribution.

Proof. Since the direction g_r is an exogenous variable, then we can derive the following conclusion: $\mathrm{Cov}(r_j, g_r) = 0,\ j = 1, \cdots, m$. Similar to Theorem 4.1, the chance constraint condition $P\left\{\sum_{j=1}^m \lambda_j r_j \geq r_d + \theta g_r\right\} \geq 1 - \alpha$ can be transformed as follows.

$$\sum_{j=1}^m \lambda_j E(r_j) + \Phi^{-1}(\alpha)\sqrt{\mathrm{Var}\left(\sum_{j=1}^m \lambda_j r_j\right) + \mathrm{Var}(r_d) + \theta^2 \mathrm{Var}(g_r) - 2\sum_{j=1}^n \lambda_j \mathrm{Cov}(r_j, r_d)} \geq E(r_d) + \theta E(g_r). \tag{4.7}$$

Let $\delta(\boldsymbol{\lambda},\ \theta) = \sqrt{\mathrm{Var}\left(\sum_{j=1}^m \lambda_j r_j\right) + \mathrm{Var}(r_d) + \theta^2 \mathrm{Var}(g_r) - 2\sum_{j=1}^n \lambda_j \mathrm{Cov}(r_j, r_d)}$, Model (4.1) can be transformed as

$$\begin{aligned} &\theta_{12}^*(\alpha) = \max\ \theta \\ &\text{s.t.} \begin{cases} \sum_{j=1}^m \lambda_j E(r_j) + \Phi^{-1}(\alpha)\delta(\boldsymbol{\lambda},\ \theta) \geq E(r_d) + \theta E(g_r), \\ \sum_{j=1}^m \lambda_j = 1,\ \lambda_j \geq 0,\ \theta \geq 0,\ j = 1, \cdots, m. \end{cases} \end{aligned} \tag{4.8}$$

Then, we can conclude that Theorem 4.2 holds here. □

In this section, we define $\widetilde{\mathbf{r}} = [r_d, g_r,\ r_1, \cdots, r_m]'$ and $\mathbf{Y} = (1, \theta,\ -\lambda_1, \cdots, -\lambda_m)'$. In addition, we let $\widetilde{\boldsymbol{\Omega}}$ denote the covariance matrix of the random vector $\widetilde{\mathbf{r}}$, where $\widetilde{\boldsymbol{\Omega}}$ is a semi-positive definite matrix. Due the fact that Model (4.2) is an extreme case of Model (4.6) (*i.e.*, $\mathrm{Var}(g_r) = 0$), then $\delta(\boldsymbol{\lambda})$ and $\delta(\boldsymbol{\lambda},\ \theta)$ can be uniformly represented as

$$\sqrt{\mathbf{Y}'\widetilde{\boldsymbol{\Omega}}\mathbf{Y}} = \left\|\widetilde{\boldsymbol{\Omega}}^{1/2}\mathbf{Y}\right\| \tag{4.9}$$

where $\|\cdot\|$ denotes the standard Euclidean norm, and $\widetilde{\boldsymbol{\Omega}}^{1/2}$ is the matrix that satisfies the condition $\widetilde{\boldsymbol{\Omega}}^{1/2}\widetilde{\boldsymbol{\Omega}}^{1/2} = \widetilde{\boldsymbol{\Omega}}$.

Therefore, Models (4.2) and (4.6) are respectively rewritten as the following second-order cone programming (SOCP).

$$\begin{aligned} &\theta_{11}^*(\alpha) = \max\ \theta \\ &\text{s.t.} \begin{cases} \sum_{j=1}^m \lambda_j E(r_j) + \Phi^{-1}(\alpha)\left\|\widetilde{\boldsymbol{\Omega}}^{1/2}\mathbf{Y}\right\| \geq E(r_d) + \theta g_r, \\ \sum_{j=1}^m \lambda_j = 1,\ \lambda_j \geq 0,\ \theta \geq 0,\ j = 1, \cdots, m. \end{cases} \end{aligned} \tag{4.10}$$

$$\begin{aligned} &\theta_{12}^*(\alpha) = \max\ \theta \\ &\text{s.t.} \begin{cases} \sum_{j=1}^m \lambda_j E(r_j) + \Phi^{-1}(\alpha)\left\|\widetilde{\boldsymbol{\Omega}}^{1/2}\mathbf{Y}\right\| \geq E(r_d) + \theta E(g_r), \\ \sum_{j=1}^m \lambda_j = 1,\ \lambda_j \geq 0,\ \theta \geq 0,\ j = 1, \cdots, m. \end{cases} \end{aligned} \tag{4.11}$$

Based on the findings in Lobo *et al.* [21] and Boyd and Vandenberghe [6] on SOCP and convex optimization, it is clear that Models (4.10) and (4.11) are both the convex programming when $\alpha \leq 0.5$.

Further, from the definition of stochastic effective shown in Cooper *et al.* [13] and Zhou *et al.* [29], for a given risk level α, when the optimal value of Model (4.2)/(4.6) is equal to 0, then the portfolio being evaluated is called as stochastic effective. Further, for any portfolio under evaluation, the portfolio efficiency can be expressed as $\eta_{11}^*(\alpha) = 1 - \theta_{11}^*(\alpha)$ or $\eta_{12}^*(\alpha) = 1 - \theta_{12}^*(\alpha)$, and the corresponding ranking is denoted as $R_{11}^*(\alpha)$ or $R_{12}^*(\alpha)$.

4.2. Nonlinear stochastic DEA model under mean-standard deviation criterion

Section 4.1 mainly uses the chance-constrained theory to deal with Set (3.1). In fact, in addition to this, the classical mean-standard deviation portfolio optimization theory can also be used to derive the deterministic estimation of Set (3.1). Under this framework, we have that $\theta_2^* = \max_\theta \Big\{\theta | r_d + \theta g_r \in \mathrm{PPS}^{(2)}\Big\}$. The detailed model can be described as

$$\theta_2^* = \max\ \theta$$
$$\text{s.t.} \begin{cases} E\Big(\sum_{j=1}^m \lambda_j r_j\Big) \geq E(r_d + \theta g_r), \\ \sigma\Big(\sum_{j=1}^m \lambda_j r_j\Big) \leq \sigma(r_d + \theta g_r), \\ \sum_{j=1}^m \lambda_j = 1,\ \lambda_j \geq 0,\ \theta \geq 0,\ j = 1, \cdots, m. \end{cases} \tag{4.12}$$

Similar to Models (4.2) and (4.6), we assume that the direction g_r can be both evaluated in the deterministic direction and exogenous random direction shown above. First, when g_r is a deterministic value, then we obtain the conclusion that $E(r_d + \theta g_r) = E(r_d) + \theta g_r$ and $\sigma(r_d + \theta g_r) = \sigma(r_d)$. Further, Model (4.12) can be simplified to the following DEA model.

$$\theta_{21}^* = \max\ \theta$$
$$\text{s.t.} \begin{cases} E\Big(\sum_{j=1}^m \lambda_j r_j\Big) \geq E(r_d) + \theta g_r, \\ \sigma\Big(\sum_{j=1}^m \lambda_j r_j\Big) \leq \sigma(r_d), \\ \sum_{j=1}^m \lambda_j = 1,\ \lambda_j \geq 0,\ \theta \geq 0,\ j = 1, \cdots, m. \end{cases} \tag{4.13}$$

Similarly, when the direction g_r is an exogenous random variable, we have that $E(r_d + \theta g_r) = E(r_d) + \theta E(g_r)$ and $\sigma(r_d + \theta g_r) = \sigma(r_d) + \theta\sigma(g_r)$. Then, Model (4.12) can be simplified as follows.

$$\theta_{22}^* = \max\ \theta$$
$$\text{s.t.} \begin{cases} E\Big(\sum_{j=1}^m \lambda_j r_j\Big) \geq E(r_d) + \theta E(g_r), \\ \sigma\Big(\sum_{j=1}^m \lambda_j r_j\Big) \leq \sigma(r_d) + \theta\sigma(g_r), \\ \sum_{j=1}^m \lambda_j = 1,\ \lambda_j \geq 0,\ \theta \geq 0,\ j = 1, \cdots, m. \end{cases} \tag{4.14}$$

Under the mean-standard deviation criterion, when the optimal value of Model (4.13)/(4.14) is equal to 0, the portfolio being evaluated is said to be stochastic effective. In addition to this, the portfolio efficiency can be defined as $\eta_{21}^* = 1 - \theta_{21}^*$ and $\eta_{22}^* = 1 - \theta_{22}^*$. Then, the corresponding portfolio rankings are denoted as R_{21}^* and R_{22}^*, respectively.

4.3. Linear stochastic DEA model under mean-standard deviation criterion

The above DEA evaluation models are all nonlinear in nature. Although they all consider the diversification role of portfolios, they may not have advantages in solving these models. Referring to the traditional linear DEA model, the stochastic DEA model with linear constraints can be constructed under the mean-standard deviation criterion. For any portfolio being evaluated, the evaluation model can be expressed as $\theta_3^* = \max_\theta \Big\{\theta | r_d + \theta g_r \in \mathrm{PPS}^{(3)}\Big\}$. The details are shown as follows.

$$\theta_3^* = \max\ \theta$$
$$\text{s.t.} \begin{cases} \sum_{j=1}^m \lambda_j E(r_j) \geq E(r_d + \theta g_r), \\ \sum_{j=1}^m \lambda_j \sigma(r_j) \leq \sigma(r_d + \theta g_r), \\ \sum_{j=1}^m \lambda_j = 1,\ \lambda_j \geq 0,\ \theta \geq 0,\ j = 1, \cdots, m. \end{cases} \tag{4.15}$$

Similarly, when g_r is a deterministic value, Model (4.15) can be simplified as follows.

$$\theta^*_{31} = \max\ \theta$$
$$\text{s.t.} \begin{cases} \sum_{j=1}^m \lambda_j E(r_j) \geq E(r_d) + \theta g_r, \\ \sum_{j=1}^m \lambda_j \sigma(r_j) \leq \sigma(r_d), \\ \sum_{j=1}^m \lambda_j = 1,\ \lambda_j \geq 0,\ \theta \geq 0,\ j = 1, \cdots, m \end{cases} \tag{4.16}$$

when g_r is an exogenous random variable, Model (4.15) can be rewritten as

$$\theta^*_{32} = \max\ \theta$$
$$\text{s.t.} \begin{cases} \sum_{j=1}^m \lambda_j E(r_j) \geq E(r_d) + \theta E(g_r), \\ \sum_{j=1}^m \lambda_j \sigma(r_j) \leq \sigma(r_d) + \theta\sigma(g_r), \\ \sum_{j=1}^m \lambda_j = 1,\ \lambda_j \geq 0,\ \theta \geq 0,\ j = 1, \cdots, m. \end{cases} \tag{4.17}$$

In this case, for any portfolio being evaluated, its efficiency is defined as $\eta^*_{31} = 1 - \theta^*_{31}$ or $\eta^*_{32} = 1 - \theta^*_{32}$, and the corresponding portfolio ranking is denoted as R^*_{31} or R^*_{32}.

4.4. The proposed stochastic DEA models *vs.* the existing DEA models

In order to further investigate the difference between the proposed stochastic DEA models and the existing DEA models, this paper will revisit the existing DEA models and their production possibility set. To the best of our knowledge, the diversification DEA and traditional DEA are two common approaches used in estimating the portfolio efficiency. However, different from the proposed models, the input-output process shown in the existing studies is constructed by using the return and risk indicators of portfolio returns. Under the mean-standard deviation criterion, the production possibility set of the diversification DEA and traditional DEA can be respectively expressed as follows.

$$\text{PPS}^{(4)} = \left\{ (\widetilde{\sigma},\ \widetilde{r}) \middle| \begin{array}{l} E\left(\sum_{j=1}^m \lambda_j r_j\right) \geq \widetilde{r},\ \ \sigma\left(\sum_{j=1}^m \lambda_j r_j\right) \leq \widetilde{\sigma}, \\ \sum_{j=1}^m \lambda_j = 1,\ \lambda_j \geq 0,\ \ j = 1, \cdots, m \end{array} \right\} \tag{4.18}$$

$$\text{PPS}^{(5)} = \left\{ (\widetilde{\sigma},\ \widetilde{r}) \middle| \begin{array}{l} \sum_{j=1}^m \lambda_j E(r_j) \geq\ \widetilde{r},\ \ \sum_{j=1}^m \lambda_j \sigma(r_j) \leq \widetilde{\sigma}, \\ \sum_{j=1}^m \lambda_j = 1,\ \lambda_j \geq 0,\ \ j = 1, \cdots, m. \end{array} \right\} \tag{4.19}$$

Based on Sets $\text{PPS}^{(4)}$ and $\text{PPS}^{(5)}$, we assume that Point $(\sigma_d,\ r_d)$ is made up of the standard deviation and expected return of the portfolio under evaluation. From the existing literature, we find that the corresponding DEA models are usually constructed by using Sets $\text{PPS}^{(4)}$ and $\text{PPS}^{(5)}$, and the input and output indicators are also characterized by the portfolio's returns and risks. At present, there are two different viewpoints in the selection of input and output indicators. The first viewpoint is that, the portfolio risk is regarded as an input indicator, while the expected return of a portfolio is regarded as an output indicator. The second viewpoint is that, the expected return and risk should be both treated as output indicators, because the expected return and risk are two evaluation indicators derived from the portfolio return. In particular, the expected return is treated as a desirable output and the risk is treated as an undesirable output. To further show the difference between the proposed DEA models and the existing DEA models, we will compare the proposed models and the existing DEA models under the different input-output assumptions. Since the proposed DEA models in this paper are all output-oriented ones, thus we mainly focus on the comparative analysis under the output-oriented measure.

First, when the portfolio risk is regarded as an input indicator, and the expected return is regarded as an output indicator, using Sets $\mathrm{PPS}^{(4)}$ and $\mathrm{PPS}^{(5)}$, the following DEA models can be derived under the output-oriented measure.

$$\theta_{41}^* = \max\ \theta$$
$$\text{s.t.} \begin{cases} E\Big(\sum_{j=1}^m \lambda_j r_j\Big) \geq E(r_d) + \theta g_O, \\ \sigma\Big(\sum_{j=1}^m \lambda_j r_j\Big) \leq \sigma(r_d), \\ \sum_{j=1}^m \lambda_j = 1,\ \lambda_j \geq 0,\ \theta \geq 0,\ j = 1, \cdots, m \end{cases} \tag{4.20}$$
$$\theta_{51}^* = \max\ \theta$$
$$\text{s.t.} \begin{cases} \sum_{j=1}^m \lambda_j E(r_j) \geq E(r_d) + \theta g_O, \\ \sum_{j=1}^m \lambda_j \sigma(r_j) \leq \sigma(r_d), \\ \sum_{j=1}^m \lambda_j = 1,\ \lambda_j \geq 0,\ \theta \geq 0,\ j = 1, \cdots, m \end{cases} \tag{4.21}$$

when the direction of Models (4.20) and (4.21) is taken as $g_O = g_r$, where g_r is the deterministic direction described in Sections 4.1–4.3, then we find that the Models (4.20) and (4.21) are consistent with the stochastic DEA models (4.13) and (4.16) proposed in this paper.

On the other hand, we assume that the expected return and risk of a portfolio are both regarded as output indicators. Referring to Zhou *et al.* [31], for a given direction $\vec{g} = (g_O^U,\ g_O^D)'$, then we have that $\theta_{42}^* = \max_\theta \Big\{\theta | (\sigma(r_d) + \theta g_O^U,\ E(r_d) + \theta g_O^D) \in \mathrm{PPS}^{(4)}\Big\}$ and $\theta_{52}^* = \max_\theta \Big\{\theta | (\sigma(r_d) + \theta g_O^U,\ E(r_d) + \theta g_O^D) \in \mathrm{PPS}^{(5)}\Big\}$. Specifically speaking, θ_{42}^* and θ_{52}^* can be expressed as follows.

$$\theta_{42}^* = \max\ \theta$$
$$\text{s.t.} \begin{cases} E\Big(\sum_{j=1}^m \lambda_j r_j\Big) \geq E(r_d) + \theta g_O^D, \\ \sigma\Big(\sum_{j=1}^m \lambda_j r_j\Big) \leq \sigma(r_d) + \theta g_O^U, \\ \sum_{j=1}^m \lambda_j = 1,\ \lambda_j \geq 0,\ \theta \geq 0,\ j = 1, \cdots, m \end{cases} \tag{4.22}$$
$$\theta_{52}^* = \max\ \theta$$
$$\text{s.t.} \begin{cases} \sum_{j=1}^m \lambda_j E(r_j) \geq E(r_d) + \theta g_O^D, \\ \sum_{j=1}^m \lambda_j \sigma(r_j) \leq \sigma(r_d) + \theta g_O^U, \\ \sum_{j=1}^m \lambda_j = 1,\ \lambda_j \geq 0,\ \theta \geq 0,\ j = 1, \cdots, m. \end{cases} \tag{4.23}$$

In order to further investigate the relationship between the proposed DEA models and the existing DEA models, we assume that g_r is an exogenous random direction and also let $g_O^U = \sigma(g_r)$ and $g_O^D = E(g_r)$. In this case, we can find that Models (4.22) and (4.23) are consistent with the stochastic DEA models (4.14) and (4.17) proposed in this paper. This further verifies that the proposed DEA models have good uniformity.

In the existing literature, diversification DEA and traditional DEA models (4.22) and (4.23) are generally assumed that $g_O^U < 0$ and $g_O^D > 0$, because decision-makers are willing to find a benchmark with higher expected return and less risk as the goal for improvement. However, in the stochastic DEA models (4.13) and (4.16), the exogenous random direction g_r satisfies the following conditions: $\sigma(g_r) > 0$ and $E(g_r) > 0$. At this time, the benchmark provided by the proposed DEA models has the characteristics of higher return and higher risk compared with the portfolio under evaluation. To analyze the difference between the proposed DEA models and the existing DEA models more intuitively, the direction of Models (4.22) and (4.23) might as well be $\vec{g} = (-\sigma(g_r),\ E(g_r))'$, and the improvement direction of risk and return derived from the exogenous random

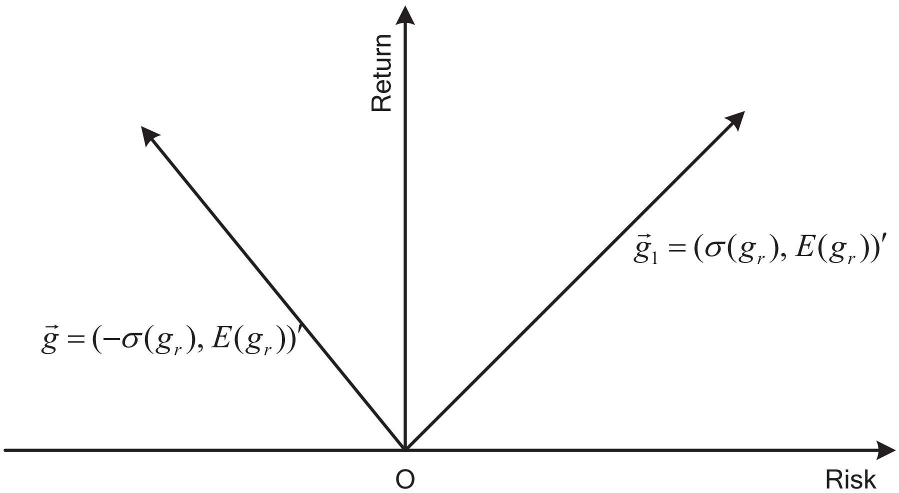

FIGURE 2. The difference between the projection directions of different DEA models.

direction g_r in Models (4.13) and (4.16) is $\vec{g}_1 = (\sigma(g_r),\ E(g_r))'$. In this case, the detailed relationship between $\vec{g}$ and $\vec{g}_1$ is shown in Figure 2. As can be seen from Figure 2, we can find that, the existing diversification DEA and traditional DEA models usually adopt a more conservative improvement direction $\vec{g} = (-\sigma(g_r),\ E(g_r))'$, because the benchmark of an ineffective portfolio has the characteristics of higher return and lower risk. However, the benchmark sought by the proposed models (4.13) and (4.16) is characterized by higher return and higher risk. Therefore, compared with the existing DEA models, the stochastic DEA models (4.13) and (4.16) are more optimistic evaluation methods.

5. Empirical analysis

In order to verify the effectiveness of the above DEA models and discuss the difference between the proposed DEA models and the existing DEA models, this paper selects 50 growing open-ended funds from China's fund market, whose fund codes are shown in Table 1. The range of weekly historical returns is from January 1, 2017 to December 31, 2019, and the data is downloaded from RESSET database (`http://db.resset.com/`). Before beginning the following empirical analysis, this paper still needs to check the normality of the fund returns selected here. As far as we know, Jarque–Bera test is a common test method to check the normality. Here, Jarque–Bera test (JB test for short) is conducted at the significance level of 5%. If the JB statistic is greater than 5.99, the test result rejects the null hypothesis of normal distribution. The specific test results are shown in Table 1.

According to the results in Table 1, we find that the historical weekly returns of the above funds are all subject to normal distribution. This also indicates that the 50 funds selected here can be used to check the effectiveness of the proposed DEA models.

5.1. Fund efficiency analysis under the deterministic direction

When the improvement direction g_r is a deterministic value, this paper will discuss the difference between the three proposed stochastic DEA models in the portfolio efficiency and its ranking. Since the chance-constrained stochastic DEA model (4.2) has taken into account the risk preference of decision-makers, that is, the portfolio efficiency obtained from Model (4.2) is dependent on the risk preference α. In this paper, it is assumed that α can be respectively evaluated at 0.3, 0.35, 0.4, 0.45 and 0.5. In addition this, we might as well assume that the

TABLE 1. The normality test of the return rates of funds being evaluated.

Fund code	JB test		Fund code	JB test	
	JB statistic	p-value		JB statistic	p-value
000534	0.2900**	0.5000	002783	2.7493**	0.1899
000535	1.7380**	0.3547	002980	1.1725**	0.5000
000729	3.2161**	0.1445	020003	1.8535**	0.3318
000756	5.4371**	0.0528	110013	2.3025**	0.2504
000880	1.7582**	0.3506	160216	3.2412**	0.1425
000913	4.7620**	0.0690	163409	1.8909**	0.3246
001018	2.3446**	0.2433	163804	2.7026**	0.1954
001060	4.1191**	0.0907	166801	3.0399**	0.1597
001069	3.4310**	0.1282	168103	5.4568**	0.0525
001097	2.6440**	0.2025	210010	1.9788**	0.3081
001272	3.2691**	0.1403	210011	2.0224**	0.3001
001305	1.6082**	0.3822	233009	4.8008**	0.0679
001479	2.0594**	0.2934	360011	3.3404**	0.1348
001513	5.4835**	0.0520	375010	5.1723**	0.0581
001543	5.4558**	0.0525	378546	2.5896**	0.2092
001667	3.6884**	0.1119	379010	1.3702**	0.4419
001676	4.0922**	0.0919	460001	5.2850**	0.0557
001735	0.6131**	0.5000	519056	1.8950**	0.3238
001838	1.0733**	0.5000	519173	5.5303**	0.0512
001933	0.9167**	0.5000	519918	1.5981**	0.3845
001986	0.4307**	0.5000	519975	2.9280**	0.1705
002251	1.0494**	0.5000	519987	4.0989**	0.0916
002317	3.8162**	0.1049	530019	3.4660**	0.1258
002450	5.0220**	0.0619	630008	3.2529**	0.1416
002577	3.7593**	0.1079	660010	2.8393**	0.1798

Notes. **5% significance level.

direction satisfies $g_r = \max_j(E(r_j)) - \min_j(E(r_j))$. Based on Models (4.2), (4.13) and (4.16), the evaluation results can be obtained as shown in Table 2.

It can be seen from Table 2 that when the risk preference of decision-makers α is lower (*e.g.*, 0.3 and 0.35), most of the funds being evaluated are stochastic effective. It also suggests that at a lower level of risk preference, it is harder for decision-makers to judge the efficiency differences between the 50 funds selected above. This is because the lower the level of risk preference, the more cautious decision-makers are, so that decision-makers cannot find an absolute dominant investment fund among the 50 funds selected here. When the risk preference of decision-makers is higher (*e.g.*, 0.4, 0.45 and 0.5), the ability of chance-constrained DEA model (4.2) to identify funds also increases. In addition, different from Model (4.2), the stochastic DEA models (4.13) and (4.16) do not consider the risk preference of decision-makers, but only alone measure the expected return and the risk of a portfolio under the mean-standard deviation framework. That is to say, Models (4.13) and (4.16) do not consider the return and risk of a portfolio as a whole, and the evaluation results are more relaxed, therefore they have a higher degree of differentiation of funds.

Next, we further quantify the differences between the efficiencies (or rankings) obtained from different DEA models. Referring to the existing literature (*e.g.*, [7,20]), we use the Pearson correlation test to analyze the fund efficiencies and rankings in Table 2. Based on the efficiency series and ranking series in Table 2, the correlation coefficients between the efficiency (or ranking) series obtained from different DEA models can be obtained, and the results are shown in Tables 3 and 4.

Table 2. Stochastic DEA efficiencies and rankings of funds being evaluated.

Fund code	Model (4.2) ($\alpha = 0.3$)		Model (4.2) ($\alpha = 0.35$)		Model (4.2) ($\alpha = 0.4$)		Model (4.2) ($\alpha = 0.45$)		Model (4.2) ($\alpha = 0.5$)		Model (4.13)		Model (4.16)	
	Efficiency $\eta^*_{11}(0.3)$	Ranking $R^*_{11}(0.3)$	Efficiency $\eta^*_{11}(0.35)$	Ranking $R^*_{11}(0.35)$	Efficiency $\eta^*_{11}(0.4)$	Ranking $R^*_{11}(0.4)$	Efficiency $\eta^*_{11}(0.45)$	Ranking $R^*_{11}(0.45)$	Efficiency $\eta^*_{11}(0.5)$	Ranking $R^*_{11}(0.5)$	Efficiency η^*_{21}	Ranking R^*_{21}	Efficiency η^*_{31}	Ranking R^*_{31}
000534	1.0000	1	1.0000	1	0.9814	5	0.7559	19	0.4103	26	0.4103	30	0.4103	31
000535	1.0000	1	1.0000	1	0.8589	18	0.5558	39	0.1794	47	0.2038	44	0.2136	44
000729	1.0000	1	1.0000	1	0.8148	20	0.6078	37	0.2910	43	0.2910	43	0.2910	43
000756	1.0000	1	1.0000	1	0.8929	14	0.6826	27	0.3856	30	0.3856	33	0.3856	34
000880	1.0000	1	1.0000	1	0.9093	12	0.6990	25	0.4147	25	0.4199	29	0.4222	29
000913	1.0000	1	1.0000	1	1.0000	1	1.0000	1	0.6596	8	0.6596	10	0.6596	10
001018	1.0000	1	1.0000	1	1.0000	1	1.0000	1	0.7053	6	0.7053	8	0.7053	8
001060	1.0000	1	1.0000	1	1.0000	1	1.0000	1	0.7083	5	0.7533	6	0.7684	6
001069	1.0000	1	1.0000	1	1.0000	1	0.9046	5	0.5630	12	0.5630	17	0.5630	17
001097	1.0000	1	1.0000	1	1.0000	1	0.7255	20	0.3774	31	0.3774	34	0.3774	35
001272	1.0000	1	1.0000	1	0.9293	11	0.6654	30	0.3402	36	0.4267	28	0.4462	27
001305	1.0000	1	1.0000	1	0.9397	10	0.6632	31	0.3063	41	0.3063	41	0.3063	41
001479	1.0000	1	1.0000	1	0.9592	8	0.7242	21	0.4036	28	0.5226	19	0.5446	18
001513	1.0000	1	1.0000	1	1.0000	1	1.0000	1	0.7963	3	0.7963	5	0.7963	5
001543	1.0000	1	1.0000	1	1.0000	1	1.0000	1	0.7999	2	0.7999	4	0.7999	4
001667	1.0000	1	1.0000	1	0.9455	9	0.6690	29	0.3453	35	0.3656	37	0.3740	36
001676	1.0000	1	0.9388	5	0.7440	21	0.5313	40	0.2917	42	0.3940	32	0.4147	30
001735	1.0000	1	1.0000	1	0.8792	17	0.6410	35	0.3290	37	0.3594	39	0.3710	37
001838	1.0000	1	1.0000	1	1.0000	1	0.6583	32	0.0912	48	0.0912	46	0.0912	46
001933	1.0000	1	1.0000	1	0.8178	19	0.5948	38	0.2903	44	0.3715	35	0.3905	33
001986	1.0000	1	1.0000	1	1.0000	1	0.8879	7	0.5185	16	0.5185	20	0.5185	20
002251	1.0000	1	1.0000	1	1.0000	1	0.7009	24	0.2016	46	0.2016	45	0.2016	45
002317	1.0000	1	1.0000	1	0.8898	16	0.6422	34	0.3199	40	0.7118	7	0.7457	7
002450	1.0000	1	1.0000	1	1.0000	1	1.0000	1	0.7849	4	0.9015	3	0.9233	3
002577	1.0000	1	1.0000	1	0.9070	13	0.6427	33	0.2890	45	0.2993	42	0.3039	42
002783	1.0000	1	1.0000	1	1.0000	1	0.8274	9	0.4859	19	0.9265	2	0.9580	2
002980	1.0000	1	1.0000	1	1.0000	1	0.9044	6	0.5675	11	0.5675	16	0.5675	16
020003	1.0000	1	1.0000	1	0.8914	15	0.6722	28	0.3641	34	0.3641	38	0.3641	39
110013	1.0000	1	1.0000	1	0.9871	3	0.9421	3	0.6449	9	0.6449	11	0.6449	11
160216	1.0000	1	1.0000	1	1.0000	1	1.0000	1	0.4064	27	0.4064	31	0.4064	32
163409	1.0000	1	1.0000	1	1.0000	1	0.7959	15	0.4882	18	0.5040	22	0.5107	22
163804	1.0000	1	1.0000	1	1.0000	1	0.8103	12	0.5470	13	0.6024	13	0.6192	13
166801	1.0000	1	1.0000	1	1.0000	1	0.7911	18	0.4834	20	0.6035	12	0.6256	12
168103	1.0000	1	1.0000	1	0.9596	7	0.7019	23	0.3698	32	0.4524	25	0.4715	25
210010	1.0000	1	1.0000	1	1.0000	1	0.8151	10	0.4388	22	1.0000	1	1.0000	1
210011	0.9871	2	0.9848	3	0.9822	4	0.7948	16	0.4183	24	1.0000	1	1.0000	1
233009	1.0000	1	0.8938	6	0.6358	23	0.3579	42	0.0000	50	0.0000	48	0.0000	48
360011	1.0000	1	1.0000	1	0.9926	2	0.6908	26	0.3234	38	0.4274	27	0.4482	26
375010	1.0000	1	1.0000	1	1.0000	1	0.9303	4	0.5122	17	0.5122	21	0.5122	21
378546	1.0000	1	1.0000	1	1.0000	1	0.8108	11	0.3951	29	0.4689	24	0.4873	24
379010	1.0000	1	1.0000	1	1.0000	1	1.0000	1	0.5879	10	0.5879	14	0.5879	15
460001	1.0000	1	1.0000	1	0.9782	6	0.7938	17	0.5392	14	0.5394	18	0.5396	19
519056	1.0000	1	1.0000	1	1.0000	1	1.0000	1	1.0000	1	1.0000	1	1.0000	1
519173	1.0000	1	1.0000	1	1.0000	1	0.7066	22	0.3667	33	0.3667	36	0.3667	38
519918	1.0000	1	1.0000	1	1.0000	1	0.8024	14	0.4347	23	0.4347	26	0.4347	28
519975	1.0000	1	0.9445	4	0.6998	22	0.4361	41	0.0890	49	0.0890	47	0.0890	47
519987	1.0000	1	1.0000	1	1.0000	1	0.8031	13	0.4633	21	0.4847	23	0.4934	23
530019	1.0000	1	0.9888	2	0.8148	20	0.6108	36	0.3213	39	0.3460	40	0.3559	40
630008	1.0000	1	1.0000	1	1.0000	1	0.8483	8	0.5212	15	0.5830	15	0.6004	14
660010	1.0000	1	1.0000	1	1.0000	1	0.9734	2	0.6629	7	0.6777	9	0.6840	9

As shown in Table 3, under the framework of chance-constrained DEA model (4.2), the correlation coefficients of fund efficiencies obtained at different risk levels are still quite different, especially when the risk preferences of decision-makers are quite different (*e.g.*, $\alpha = 0.3$ and $\alpha = 0.5$). This indicates that the decision-makers' risk preference will have a greater impact on the fund efficiency. On the other hand, under the mean-standard deviation criterion, we find that the fund efficiencies obtained by using the nonlinear and linear stochastic DEA models (*i.e.*, Models (4.13) and (4.16)) have a higher correlation (*i.e.*, the correlation coefficient is 0.9992). This also shows that, although Model (4.16) does not consider the diversification effect between portfolios, it can be still better approximate the fund efficiency obtained from Model (4.13). Most importantly, as the risk preference of decision-makers increases, the correlation coefficient between the fund efficiency from Model (4.2) and that from Model (4.13) or Model (4.16) also increases.

TABLE 3. The correlation of fund efficiencies from different stochastic DEA models.

Correlation coefficient	$\eta_{11}^*(0.3)$	$\eta_{11}^*(0.35)$	$\eta_{11}^*(0.4)$	$\eta_{11}^*(0.45)$	$\eta_{11}^*(0.5)$	η_{21}^*	η_{31}^*
$\eta_{11}^*(0.3)$	1.0000						
$\eta_{11}^*(0.35)$	0.0786	1.0000					
$\eta_{11}^*(0.4)$	−0.0573	0.7687	1.0000				
$\eta_{11}^*(0.45)$	−0.0175	0.5436	0.8020	1.0000			
$\eta_{11}^*(0.5)$	0.0196	0.4385	0.6245	0.8952	1.0000		
η_{21}^*	−0.3066	0.3701	0.5662	0.7325	0.8110	1.0000	
η_{31}^*	−0.2995	0.3691	0.5590	0.7199	0.8014	0.9992	1.0000

TABLE 4. The correlation of efficiency rankings from different stochastic DEA models.

Correlation coefficient	$R_{11}^*(0.3)$	$R_{11}^*(0.35)$	$R_{11}^*(0.4)$	$R_{11}^*(0.45)$	$R_{11}^*(0.5)$	R_{21}^*	R_{31}^*
$R_{11}^*(0.3)$	1.0000						
$R_{11}^*(0.35)$	0.2417	1.0000					
$R_{11}^*(0.4)$	−0.0476	0.5342	1.0000				
$R_{11}^*(0.45)$	−0.0236	0.4269	0.8528	1.0000			
$R_{11}^*(0.5)$	−0.0148	0.3717	0.7111	0.9289	1.0000		
R_{21}^*	−0.2248	0.2744	0.6165	0.7903	0.8528	1.0000	
R_{31}^*	−0.2248	0.2632	0.5978	0.7742	0.8418	0.9984	1.0000

Table 4 shows that, for Model (4.2), the risk preference of decision-makers will also have a great impact on the fund ranking. In addition, Models (4.13) and (4.16) have a good consistency in the fund ranking. Further, the correlation coefficient between the fund ranking of Model (4.2) and that of Model (4.13) or Model (4.16) is increasing with the risk preference of decision-makers. Obviously, the conclusion in Table 4 is also in line with that in Table 3.

5.2. Fund efficiency analysis under the exogenous random direction

In this section, we assume that the improvement direction g_r is an exogenous random variable. Similar to Section 5.1, we will discuss the difference between the fund efficiencies/rankings of the three proposed DEA models. In addition, we assume that α can be respectively evaluated at 0.3, 0.35, 0.4, 0.45 and 0.5. Further, we also suppose that the random direction g_r satisfies the conditions that $E(g_r) = \max_j(E(r_j)) - \min_j(E(r_j))$ and $\sigma(g_r) = \max_j(\sigma(r_j)) - \min_j(\sigma(r_j))$[3]. Using Models (4.6), (4.14) and (4.17), we can obtain the evaluation results as shown in Table 5.

Similar to Section 5.1, this paper will make a correlation test of the fund efficiency and its ranking, and then analyze the differences among different stochastic DEA models. The detailed results are shown in Tables 6 and 7.

For Model (4.6), Tables 6 and 7 show that decision-makers' risk preference will also have a great impact on both the fund efficiency and the fund ranking. In addition, from the results shown in Tables 6 and 7, we find that Models (4.14) and (4.17) also have good consistency in the fund efficiency and its ranking. Further, under the framework of Models (4.6), (4.14) and (4.17), the corresponding correlation coefficient of between fund efficiencies/rankings increases with the increase of decision-makers' risk preference. It is not difficult to

[3]It is worth noting that we only take the mean and variance of the exogenous random direction at the values given above and are not trying to show that there is a relationship between the random direction and the portfolio returns. In fact, under our assumption we can arbitrarily give the mean and variance of the random direction.

TABLE 5. Stochastic DEA efficiencies and rankings of funds being evaluated.

Fund code	Model (4.6) ($\alpha = 0.3$)		Model (4.6) ($\alpha = 0.35$)		Model (4.6) ($\alpha = 0.4$)		Model (4.6) ($\alpha = 0.45$)		Model (4.6) ($\alpha = 0.5$)		Model (4.14)		Model (4.17)	
	Efficiency $\eta^*_{12}(0.3)$	Ranking $R^*_{12}(0.3)$	Efficiency $\eta^*_{12}(0.35)$	Ranking $R^*_{12}(0.35)$	Efficiency $\eta^*_{12}(0.4)$	Ranking $R^*_{12}(0.4)$	Efficiency $\eta^*_{12}(0.45)$	Ranking $R^*_{12}(0.45)$	Efficiency $\eta^*_{12}(0.5)$	Ranking $R^*_{12}(0.5)$	Efficiency η^*_{22}	Ranking R^*_{22}	Efficiency η^*_{32}	Ranking R^*_{32}
000534	1.0000	1	1.0000	1	0.9820	5	0.7869	19	0.4103	26	0.4103	26	0.4103	25
000535	1.0000	1	1.0000	1	0.8772	18	0.6244	39	0.1794	47	0.1794	47	0.1794	46
000729	1.0000	1	1.0000	1	0.8527	21	0.6773	37	0.2910	43	0.2910	43	0.2910	42
000756	1.0000	1	1.0000	1	0.9078	15	0.7324	25	0.3856	30	0.3856	30	0.3856	29
000880	1.0000	1	1.0000	1	0.9205	12	0.7447	22	0.4147	25	0.4147	25	0.4147	24
000913	1.0000	1	1.0000	1	1.0000	1	1.0000	1	0.6596	8	0.6596	9	0.6596	10
001018	1.0000	1	1.0000	1	1.0000	1	1.0000	1	0.7053	6	0.7053	7	0.7053	8
001060	1.0000	1	1.0000	1	1.0000	1	1.0000	1	0.7083	5	0.7083	6	0.7083	6
001069	1.0000	1	1.0000	1	1.0000	1	0.9099	6	0.5630	12	0.5630	15	0.5630	14
001097	1.0000	1	1.0000	1	1.0000	1	0.7581	21	0.3774	31	0.3774	31	0.3774	30
001272	1.0000	1	1.0000	1	0.9354	11	0.7124	30	0.3402	36	0.3402	36	0.3402	35
001305	1.0000	1	1.0000	1	0.9440	10	0.7095	31	0.3063	41	0.3063	41	0.3063	40
001479	1.0000	1	1.0000	1	0.9616	8	0.7598	20	0.4036	28	0.4036	28	0.4036	27
001513	1.0000	1	1.0000	1	1.0000	1	1.0000	1	0.7963	3	0.7963	3	0.7963	4
001543	1.0000	1	1.0000	1	1.0000	1	1.0000	1	0.7999	2	0.7999	2	0.7999	3
001667	1.0000	1	1.0000	1	0.9490	9	0.7140	29	0.3453	35	0.3453	35	0.3453	34
001676	1.0000	1	0.9471	5	0.8055	22	0.6210	40	0.2917	42	0.2917	42	0.2917	41
001735	1.0000	1	1.0000	1	0.8960	17	0.6972	32	0.3290	37	0.3290	37	0.3290	36
001838	1.0000	1	1.0000	1	1.0000	1	0.6903	35	0.0912	48	0.0912	48	0.0912	47
001933	1.0000	1	1.0000	1	0.8531	19	0.6637	38	0.2903	44	0.2903	44	0.2903	43
001986	1.0000	1	1.0000	1	1.0000	1	0.8958	7	0.5185	16	0.5185	19	0.5185	18
002251	1.0000	1	1.0000	1	1.0000	1	0.7303	26	0.2016	46	0.2016	46	0.2016	45
002317	1.0000	1	1.0000	1	0.9082	14	0.6911	34	0.3199	40	0.3199	40	0.3199	39
002450	1.0000	1	1.0000	1	1.0000	1	1.0000	1	0.7849	4	0.7849	4	0.7849	5
002577	1.0000	1	1.0000	1	0.9168	13	0.6949	33	0.2890	45	0.2890	45	0.2890	44
002783	1.0000	1	1.0000	1	1.0000	1	0.8409	9	0.4859	19	0.5710	13	0.7063	7
002980	1.0000	1	1.0000	1	1.0000	1	0.9106	5	0.5675	11	0.5675	14	0.5675	13
020003	1.0000	1	1.0000	1	0.9063	16	0.7228	28	0.3641	34	0.3641	34	0.3641	33
110013	1.0000	1	1.0000	1	0.9884	4	0.9476	3	0.6449	9	0.6449	11	0.6449	11
160216	1.0000	1	1.0000	1	1.0000	1	1.0000	1	0.4064	27	0.4064	27	0.4064	26
163409	1.0000	1	1.0000	1	1.0000	1	0.8180	16	0.4882	18	0.4882	21	0.4882	20
163804	1.0000	1	1.0000	1	1.0000	1	0.8321	10	0.5470	13	0.5470	16	0.5470	15
166801	1.0000	1	1.0000	1	1.0000	1	0.8129	17	0.4834	20	0.4834	22	0.4834	21
168103	1.0000	1	1.0000	1	0.9618	7	0.7408	24	0.3698	32	0.3698	32	0.3698	31
210010	1.0000	1	1.0000	1	1.0000	1	0.8284	11	0.4388	22	0.7643	5	1.0000	1
210011	0.9944	2	0.9925	2	0.9888	3	0.8109	18	0.4183	24	0.6507	10	0.8588	2
233009	1.0000	1	0.9117	6	0.7269	24	0.4841	42	0.0000	50	0.0000	50	0	49
360011	1.0000	1	1.0000	1	0.9926	2	0.7286	27	0.3234	38	0.3234	38	0.3234	37
375010	1.0000	1	1.0000	1	1.0000	1	0.9326	4	0.5122	17	0.5122	20	0.5122	19
378546	1.0000	1	1.0000	1	1.0000	1	0.8236	12	0.3951	29	0.3951	29	0.3951	28
379010	1.0000	1	1.0000	1	1.0000	1	1.0000	1	0.5879	10	0.5879	12	0.5879	12
460001	1.0000	1	1.0000	1	0.9791	6	0.8198	15	0.5392	14	0.5392	17	0.5392	16
519056	1.0000	1	1.0000	1	1.0000	1	1.0000	1	1.0000	1	1.0000	1	1.0000	1
519173	1.0000	1	1.0000	1	1.0000	1	0.7414	23	0.3667	33	0.3667	33	0.3667	32
519918	1.0000	1	1.0000	1	1.0000	1	0.8202	14	0.4347	23	0.4347	24	0.4347	23
519975	1.0000	1	0.9503	4	0.7697	23	0.5428	41	0.0890	49	0.0890	49	0.0890	48
519987	1.0000	1	1.0000	1	1.0000	1	0.8231	13	0.4633	21	0.4633	23	0.4633	22
530019	1.0000	1	0.9893	3	0.8529	20	0.6794	36	0.3213	39	0.3213	39	0.3213	38
630008	1.0000	1	1.0000	1	1.0000	1	0.8607	8	0.5212	15	0.5212	18	0.5212	17
660010	1.0000	1	1.0000	1	1.0000	1	0.9738	2	0.6629	7	0.6629	8	0.6629	9

find that the conclusions in Tables 6 and 7 are consistent with those in Tables 3 and 4. This also indicates that the different improvement directions will not change the properties of the stochastic DEA models.

5.3. Comparative analysis of different DEA models

Sections 5.1 and 5.2 mainly focus on the analysis of the differences between the proposed stochastic DEA models. In the following, we will investigate the difference between the stochastic DEA models and the existing DEA models. To keep consistent with the directions selected in Sections 5.1 and 5.2, we assume that the directions of traditional DEA models (4.20)–(4.23) satisfy that $g_O = g_O^D = \max_j(E(r_j)) - \min_j(E(r_j))$ and $g_O^U = \max_j(\sigma(r_j)) - \min_j(\sigma(r_j))$. Based on Models (4.20)–(4.23), the empirical results can be obtained as shown in Table 8.

Table 6. The correlation of fund efficiencies from different stochastic DEA models.

Correlation coefficient	$\eta_{12}^*(0.3)$	$\eta_{12}^*(0.35)$	$\eta_{12}^*(0.4)$	$\eta_{12}^*(0.45)$	$\eta_{12}^*(0.5)$	η_{22}^*	η_{32}^*
$\eta_{12}^*(0.3)$	1.0000						
$\eta_{12}^*(0.35)$	0.0301	1.0000					
$\eta_{12}^*(0.4)$	−0.0675	0.7425	1.0000				
$\eta_{12}^*(0.45)$	−0.0052	0.5134	0.7776	1.0000			
$\eta_{12}^*(0.5)$	0.0196	0.4409	0.6336	0.9055	1.0000		
η_{22}^*	−0.1382	0.4301	0.6467	0.8794	0.9593	1.0000	
η_{32}^*	−0.2563	0.4003	0.6254	0.8160	0.8811	0.9781	1.0000

Table 7. The correlation of efficiency rankings from different stochastic DEA models.

Correlation coefficient	$R_{12}^*(0.3)$	$R_{12}^*(0.35)$	$R_{12}^*(0.4)$	$R_{12}^*(0.45)$	$R_{12}^*(0.5)$	R_{22}^*	R_{32}^*
$R_{12}^*(0.3)$	1.0000						
$R_{12}^*(0.35)$	0.0995	1.0000					
$R_{12}^*(0.4)$	−0.0672	0.5909	1.0000				
$R_{12}^*(0.45)$	−0.0023	0.4595	0.8229	1.0000			
$R_{12}^*(0.5)$	−0.0148	0.3923	0.7046	0.9434	1.0000		
R_{22}^*	−0.1534	0.3730	0.7033	0.9234	0.9705	1.0000	
R_{32}^*	−0.2235	0.3638	0.7019	0.9046	0.9468	0.9947	1.0000

It can be seen from Table 8 that both Models (4.20) and (4.21) take the standard deviation of the fund return as an input and the expected return as an output. Under this input-output assumption, diversification DEA model (4.20) and traditional DEA model (4.21) are consistent in both the fund efficiency and the fund ranking. On the other hand, Models (4.22) and (4.23) treat the expectation and standard deviation of fund returns as output indicators. From Table 8, we find that Models (4.22) and (4.23) are also consistent in the fund efficiency and its ranking. For the above reasons, we only select the diversification DEA models (4.20) and (4.22) to compare with the proposed stochastic DEA models.

Additionally, since the above chance-constrained DEA models with the lower risk preference has a poor ability to identify the above 50 funds, this paper only selects $\alpha = 0.45$ as the unique risk preference of decision-makers. Under this given risk preference, we also take the corresponding chance-constrained DEA models (*i.e.*, Models (4.2) and (4.6)) as the comparison with the traditional DEA models. Furthermore, as shown in Sections 5.1 and 5.2, we can find that Models (4.13) and (4.16), Models (4.14) and (4.17) are both consistent in the fund efficiency and its ranking. For this reason, Models (4.13) and (4.14) are also included in the comparison with the traditional DEA models. Based on the DEA models selected above, we can obtain the test results as shown in Tables 9 and 10.

As shown in Table 9, by comparing the chance-constrained DEA models (4.2) and (4.6) and the traditional DEA models (4.16) and (4.17), it is not difficult to find that the correlation coefficients between the efficiencies obtained by Models (4.2) and (4.6) and the efficiencies obtained by Models (4.16) and (4.17) are not very high. These indicate that there are differences between Models (4.2) and (4.6) and Models (4.16) and (4.17). In particular, the difference between Models (4.2) and (4.6) and Model (4.17) is more obvious. In addition, compared with Models (4.2) and (4.6), the correlations between the stochastic DEA models (4.13) and (4.14) and the existing DEA models (4.16) and (4.17) are higher, but there are still some differences between them.

Table 8. The fund efficiencies and rankings derived from traditional DEA models.

Fund code	Model (4.20)		Model (4.21)		Model (4.22)		Model (4.23)	
	Efficiency η_{41}^*	Ranking R_{41}^*	Efficiency η_{51}^*	Ranking R_{51}^*	Efficiency η_{42}^*	Ranking R_{42}^*	Efficiency η_{52}^*	Ranking R_{52}^*
000534	0.4103	30	0.4103	31	0.6573	27	0.6737	27
000535	0.2038	44	0.2136	44	0.5581	43	0.5765	43
000729	0.2910	43	0.2910	43	0.5755	41	0.5929	41
000756	0.3856	33	0.3856	34	0.6259	35	0.6421	35
000880	0.4199	29	0.4222	29	0.6721	25	0.6888	24
000913	0.6596	10	0.6596	10	0.6808	24	0.6854	25
001018	0.7053	8	0.7053	8	0.7857	9	0.7960	10
001060	0.7533	6	0.7684	6	0.8620	4	0.8753	4
001069	0.5630	17	0.5630	17	0.6882	21	0.7003	23
001097	0.3774	34	0.3774	35	0.6321	33	0.6488	33
001272	0.4267	28	0.4462	27	0.6834	23	0.7018	22
001305	0.3063	41	0.3063	41	0.5745	42	0.5915	42
001479	0.5226	19	0.5446	18	0.7366	15	0.7548	14
001513	0.7963	5	0.7963	5	0.7963	8	0.7963	9
001543	0.7999	4	0.7999	4	0.8267	6	0.8323	6
001667	0.3656	37	0.3740	36	0.6451	29	0.6629	29
001676	0.3940	32	0.4147	30	0.6665	26	0.6848	26
001735	0.3594	39	0.3710	37	0.6433	30	0.6613	30
001838	0.0912	46	0.0912	46	0.3352	48	0.3515	48
001933	0.3715	35	0.3905	33	0.6534	28	0.6718	28
001986	0.5185	20	0.5185	20	0.6111	37	0.6218	38
002251	0.2016	45	0.2016	45	0.4463	46	0.4627	46
002317	0.7118	7	0.7457	7	0.8569	5	0.8630	5
002450	0.9015	3	0.9233	3	0.9452	3	0.9587	3
002577	0.2993	42	0.3039	42	0.6071	38	0.6251	37
002783	0.9265	2	0.9580	2	0.9627	2	0.9774	2
002980	0.5675	16	0.5675	16	0.7076	18	0.7203	19
020003	0.3641	38	0.3641	39	0.6307	34	0.6477	34
110013	0.6449	11	0.6449	11	0.7507	13	0.7620	13
160216	0.4064	31	0.4064	32	0.5235	44	0.5353	44
163409	0.5040	22	0.5107	22	0.7205	16	0.7365	16
163804	0.6024	13	0.6192	13	0.7789	11	0.7949	11
166801	0.6035	12	0.6256	12	0.7807	10	0.7984	8
168103	0.4524	25	0.4715	25	0.6973	20	0.7154	20
210010	1.0000	1	1.0000	1	1.0000	1	1.0000	1
210011	1.0000	1	1.0000	1	1.0000	1	1.0000	1
233009	0.0000	48	0.0000	48	0.4299	47	0.4473	47
360011	0.4274	27	0.4482	26	0.6845	22	0.7029	21
375010	0.5122	21	0.5122	21	0.5998	39	0.6103	40
378546	0.4689	24	0.4873	24	0.7060	19	0.7239	18
379010	0.5879	14	0.5879	15	0.6215	36	0.6283	36
460001	0.5394	18	0.5396	19	0.7372	14	0.7521	15
519056	1.0000	1	1.0000	1	1.0000	1	1.0000	1
519173	0.3667	36	0.3667	38	0.5983	40	0.6143	39
519918	0.4347	26	0.4347	28	0.6406	31	0.6558	31
519975	0.0890	47	0.0890	47	0.4594	45	0.4778	45
519987	0.4847	23	0.4934	23	0.7107	17	0.7272	17
530019	0.3460	40	0.3559	40	0.6352	32	0.6531	32
630008	0.5830	15	0.6004	14	0.7684	12	0.7848	12
660010	0.6777	9	0.6840	9	0.8166	7	0.8299	7

TABLE 9. The correlation of fund efficiencies from different DEA models.

Correlation coefficient	$\eta^*_{11}(0.45)$	$\eta^*_{12}(0.45)$	η^*_{21}	η^*_{22}	η^*_{41}	η^*_{42}
$\eta^*_{11}(0.45)$	1.0000					
$\eta^*_{12}(0.45)$	0.9979	1.0000				
η^*_{21}	0.7325	0.7288	1.0000			
η^*_{22}	0.8747	0.8794	0.9154	1.0000		
η^*_{41}	0.7325	0.7288	1.0000	0.9154	1.0000	
η^*_{42}	0.5312	0.5275	0.9433	0.8098	0.9433	1.0000

TABLE 10. The correlation of efficiency rankings from different DEA models.

Correlation coefficient	$R^*_{11}(0.45)$	$R^*_{12}(0.45)$	R^*_{21}	R^*_{22}	R^*_{41}	R^*_{42}
$R^*_{11}(0.45)$	1.0000					
$R^*_{12}(0.45)$	0.9968	1.0000				
R^*_{21}	0.7903	0.7942	1.0000			
R^*_{22}	0.9147	0.9234	0.9072	1.0000		
R^*_{41}	0.7903	0.7942	1.0000	0.9072	1.0000	
R^*_{42}	0.5498	0.5583	0.8955	0.7525	0.8955	1.0000

Table 10 shows the difference between the efficiency rankings of stochastic DEA models and those of the existing DEA models. It can be seen from Table 10 that there are also certain differences between the above stochastic DEA models and the existing DEA models in the efficiency ranking. This further validates the conclusions in Table 9. Since the proposed stochastic DEA models are all derived from the actual input-output process of portfolios, they can better reflect the real portfolio efficiency. Therefore, the results in Tables 9 and 10 further indicate the rationality of the stochastic DEA models proposed in this paper.

6. Conclusion

Portfolio efficiency evaluation has always been one of the research hotspots in the financial field. To the best of our knowledge, traditional DEA and diversification DEA models are two common nonparametric evaluation methods, which have been widely used in the portfolio efficiency evaluation. The existing researches usually use the expected return and risk indicators to build the input-output process of portfolios. At present, there are mainly two viewpoints on this issue. The first viewpoint considers the portfolio's risk as an input and the expected return as an output. The second viewpoint is that the expected return and risk of a portfolio should be both regarded as output indicators, because they are two evaluation indicators derived from the portfolio return. It is not difficult to find that these existing input-output assumptions are inconsistent with the actual investment process, because the real input should be the initial wealth and the output be the terminal wealth. In this input-output framework, this paper standardizes the initial wealth of a portfolio and converts the terminal wealth value into the form of return rate, which is usually observable.

In this paper, we first clarify the actual input-output process of portfolios, and then construct the corresponding stochastic production possibility set. In the following, we propose three estimation methods to deal with the stochastic production possibility sets constructed above, and then propose three deterministic sets as the estimations of the above stochastic production possibility set. Using the above three kinds of deterministic estimation, three kinds of stochastic DEA models are directly constructed, including the chance-constrained DEA model, and the nonlinear and linear stochastic DEA models under the mean-standard deviation criterion.

Finally, the proposed stochastic DEA models are used to assess the efficiency of 50 open growth models in China's fund market. By comparing with the existing DEA models, the effectiveness of the proposed stochastic DEA models is further verified.

Acknowledgements. This research is supported by the National Natural Science Foundation of China (Nos. 71771082 and 71801091), Hunan Provincial Natural Science Foundation of China (No. 2020JJ5377), China Postdoctoral Science Foundation funded project (No. 2020M682577), and Hunan Key Laboratory of Macroeconomic Big Data Mining and its Application.

References

[1] A. Basso and S. Funari, A data envelopment analysis approach to measure the mutual fund performance. *Eur. J. Oper. Res.* **135** (2001) 477–492.

[2] A. Basso and S. Funari, Constant and variable returns to scale DEA models for socially responsible investment funds. *Eur. J. Oper. Res.* **235** (2014) 775–783.

[3] A. Basso and S. Funari, The role of fund size in the performance of mutual funds assessed with DEA models. *Eur. J. Finan.* **23** (2017) 457–473.

[4] E. Basso, A. Allevi, F. Bonenti, G. Oggioni and R. Riccardi, Measuring the environmental performance of green SRI funds: a DEA approach. *Energy Econ.* **79** (2019) 32–44.

[5] P. Bonami and M.A. Lejeune, An exact solution approach for portfolio optimization problems under stochastic and integer constraints. *Oper. Res.* **57** (2009) 650–670.

[6] S. Boyd and L. Vandenberghe, Convex Optimization. Cambridge University Press, Cambridge (2004).

[7] M. Branda, Diversification-consistent data envelopment analysis based on directional-distance measures. *Omega* **52** (2015) 65–76.

[8] W. Briec, K. Kerstens and J.B. Lesourd, Single-period Markowitz portfolio selection, performance gauging, and duality: a variation on the Luenberger shortage function. *J. Optim. Theory App.* **120** (2004) 1–27.

[9] W. Briec, K. Kerstens and O. Jokung, Mean-variance-skewness portfolio performance gauging: a general shortage function and dual approach. *Manage. Sci.* **53** (2007) 135–149.

[10] Z. Chen and R. Lin, Mutual fund performance evaluation using data envelopment analysis with new risk measures. *OR Spect.* **28** (2006) 375–398.

[11] W. Chen, Y. Gai and P. Gupta, Efficiency evaluation of fuzzy portfolio in different risk measures via DEA. *Ann. Oper. Res.* **269** (2018) 103–127.

[12] W.W. Cooper, Z. Huang, V. Lelas, S.X. Li and O.B. Olesen, Chance constrained programming formulations for stochastic characterizations of efficiency and dominance in DEA. *J. Prod. Anal.* **9** (1998) 53–79.

[13] W.W. Cooper, H. Deng, Z. Huang and S.X. Li, Chance constrained programming approaches to technical efficiencies and inefficiencies in stochastic data envelopment analysis. *J. Oper. Res. Soc.* **53** (2002) 1347–1356.

[14] H. Ding, Z. Zhou, H. Xiao, C. Ma and W. Liu, Performance evaluation of portfolios with margin requirements. *Math. Prob. Eng.* **2014** (2014) 1–8.

[15] Z. Huang and S.X. Li, Stochastic DEA models with different types of input-output disturbances. *J. Prod. Anal.* **15** (2001) 95–113.

[16] T. Joro and P. Na, Portfolio performance evaluation in a mean-variance-skewness framework. *Eur. J. Oper. Res.* **175** (2006) 446–461.

[17] J.D. Lamb and K.H. Tee, Data envelopment analysis models of investment funds. *Eur. J. Oper. Res.* **216** (2012) 687–696.

[18] R. Lin and Z. Li, Directional distance based diversification super-efficiency DEA models for mutual funds. *Omega* **97** (2019) 102096.

[19] W.B. Liu, D.Q. Zhang, W. Meng, X.X. Li and F. Xu, A study of DEA models without explicit inputs. *Omega* **39** (2011) 472–480.

[20] W. Liu, Z. Zhou, D. Liu and H. Xiao, Estimation of portfolio efficiency via DEA. *Omega* **52** (2015) 107–118.

[21] M.S. Lobo, L. Vandenberghe, S. Boyd and H. Lebret, Applications of second-order cone programming. *Linear Algebra App.* **284** (1998) 193–228.

[22] H. Markowitz, Portfolio selection. *J. Finan.* **7** (1952) 77–91.

[23] B.P.S. Murthi, Y.K. Choi and P. Desai, Efficiency of mutual funds and portfolio performance measurement: a non-parametric approach. *Eur. J. Oper. Res.* **98** (1997) 408–418.

[24] O.B. Olesen and N.C. Petersen, Stochastic data envelopment analysis – a review. *Eur. J. Oper. Res.* **251** (2016) 2–21.

[25] T. Ren, Z. Zhou and H. Xiao, Estimation of portfolio efficiency considering social responsibility: evidence from the multi-horizon diversification DEA. *RAIRO: Oper. Res.* **55** (2021) 611–637.

[26] A.C. Tarnaud and H. Leleu, Portfolio analysis with DEA: prior to choosing a model. *Omega* **75** (2018) 57–76.

[27] E.G. Tsionas and E.N. Papadakis, A Bayesian approach to statistical inference in stochastic DEA. *Omega* **38** (2010) 309–314.

[28] G. Yang, W. Shen, D. Zhang and W. Liu, Extended utility and DEA models without explicit input. *J. Oper. Res. Soc.* **65** (2014) 1212–1220.

[29] Z. Zhou, L. Lin, H. Xiao, C. Ma and S. Wu, Stochastic network DEA models for two-stage systems under the centralized control organization mechanism. *Comput. Ind. Eng.* **110** (2017) 404–412.

[30] Z. Zhou, H. Xiao, Q. Jin and W. Liu, DEA frontier improvement and portfolio rebalancing: an application of China mutual funds on considering sustainability information disclosure. *Eur. J. Oper. Res.* **269** (2018) 111–131.

[31] Z. Zhou, E. Chen, H. Xiao, T. Ren and Q. Jin, Performance evaluation of portfolios with fuzzy returns. *RAIRO: Oper. Res.* **53** (2019) 1581–1600.