

# Estimation of portfolio efficient frontier by different measures of risk via DEA

M. Sanei <sup>\*</sup>, S. Banihashemi <sup>†‡</sup>, M. Kaveh <sup>§</sup>

Received Date: 2014-11-27    Revised Date: 2015-05-22    Accepted Date: 2016-03-11

## Abstract

In this paper, linear Data Envelopment Analysis models are used to estimate Markowitz efficient frontier. Conventional DEA models assume non-negative values for inputs and outputs. However, variance is the only variable in these models that takes non-negative values. Therefore, negative data models which the risk of the assets had been used as an input and expected return was the output are utilized. At the beginning variance was considered as a risk measure. However, both theories and practices indicate that variance is not a good measure of risk. Then value at risk is introduced as new risk measure. In this paper, we should prove that with increasing sample size, the frontiers of the linear models with both variance and value at risk, as risk measure, gradually approximate the frontiers of the mean-variance and mean-value at risk models and non-linear model with negative data. Finally, we present a numerical example with variance and value at risk that obtained via historical simulation and variance-covariance method as risk measures to demonstrate the usefulness and effectiveness of our claim.

*Keywords* : Portfolio; Data Envelopment Analysis (DEA); Value at Risk (VaR); Negative data.

## 1 Introduction

In financial literature, a portfolio is an appropriate mix investments held by an institution or private individuals. For investors, best portfolios or assets selection and risks management are always challenging topics. Investors typically try to find portfolios or assets offering less risk and more return. Evaluation of portfolio performance has created a large interest among employees also

academic researchers because of huge amount of money are being invested in financial markets. One important idea in portfolio evaluation is the portfolio frontier approach, which measures performance of a portfolio by some its distances to the efficient portfolio frontier. In 1952 Markowitz [20] work, laid the base of the frontier approach under the mean-variance (MV) framework. This model was due to the nature of the variance in quadratic form, and tries to decrease variance as a risk parameter in all levels of mean. This model results in an area with a frontier called efficient frontier. Data Envelopment Analysis has proved the efficiency for assessing the relative efficiency of Decision Making Units (DMUs) that employs multiple inputs to produce multiple outputs (Charnes et al. 1978 [9]). Mean-variance idea has been much extended afterwards, and the

<sup>\*</sup>Department of Applied Mathematics, Islamic Azad University of Central Tehran branch, Tehran, Iran.

<sup>†</sup>Corresponding author. [shbanihashemi@atu.ac.ir](mailto:shbanihashemi@atu.ac.ir)

<sup>‡</sup>Department of Mathematics, Faculty of Mathematics and Computer Science, Allameh Tabataba'i University, Tehran Iran.

<sup>§</sup>Department of Applied Mathematics, Islamic Azad University of Central Tehran branch, Tehran, Iran.

models further being developed along this idea are often referred as nonlinear DEA models.

In 1999 Morey and Morey [23] proposed mean-variance framework based on Data Envelopment Analysis, in which variance of the portfolio is used as an input to DEA models and expected return is used as an output. In 2004 Briec et al. [6] tried to project points in a preferred direction on efficient frontier and evaluate points' efficiencies by their distances. Demonstrated model by Briec et al. [7] which is also known as a shortage function, has some advantages. For example optimization can be done in any direction of a mean-variance space according to the investors' ideal. Furthermore, in shortage function, efficiency of each security is defined as the distance between the asset and its projection in a pre-assumed direction. As an instance in variance direction optimization it is equal to the ratio between variance of projection point and variance of asset. Based on this definition if distance equals to zero, that security is on the frontier area and its efficiency equals to 1. This number, in fact, is the result of shortage function which tries to summarize value of efficiency by a number. Similar to any other model, mean-variance model has its own assumptions. Normality is one of its important assumptions. In mean-variance model, distribution of mean of securities in a particular time horizon should be normal. In contrast Mandelbrot [19] showed, not only empirical distributions are widely skewed, but they also have thicker tails than normal. Ariditti [2] and Kraus and Litzenberger [17] also showed that expected return in respect of third moment is positive. Ariditti [2], Kane [15], Ho and Chang [12] showed that most investors prefer positive skewed assets or portfolios, which means that skewness is an output parameter and same as mean or expected return, should be increased. Based on Mitton and Vorkink [22] most investors scarify mean-variance model efficiencies for higher skewed portfolios. In this way Joro and Na [14] introduced mean-variance-skewness framework, in which skewness of returns considered as outputs. Also, Joro and Na reported that the linear DEA estimation of portfolio efficiency is not consistent with the results from their non-linear model. Briec et al. [7] introduced a new shortage function which obtains an efficiency measure which looks to improve both mean and skewness and decreases variance. Kers-

tence et al. [16] introduced a geometric representation of the MVS frontier related to new tools introduced in their paper. In the new models instead of estimating the whole efficient frontier, only the projection points of the assets are computed. In these models a non-linear DEA-type framework is used where the correlation structure among the units is taken into account. Nowadays, most investors think consideration of skewness and kurtosis in models are critical. Mhiri and Prigent [21] analyzed the portfolio optimization problem by introducing higher moments of return – the main financial index. However, using this approach needs variety of assumptions hold, there is not a general willingness to incorporate higher order moments. Up to this point the assumption is that variance is a parameter that evaluates risk and it is preferred to be decreased, although, not everybody wants this. For example a venture capitalist prefers risky portfolios or assets, followed by more return than normal. In mean-variance models evaluation, such situations are considered as undesirable situations. But they are not really undesirable for those who are interested in risk for higher returns. There are some approaches, trying to address such ambiguities by introducing other parameters, such as semi variance. However, each approach has its own disadvantage which makes it less desirable. A new approach to manage and control risk is value at risk (VaR) approach. This new approach focuses on the left hand side of the range of normal distribution where negative returns come with high risk. Value at risk was first proposed by Baumol [4]. The goal is to measure loss of return on left side of the portfolio's return distribution by reporting a number. Based on VaR definition, it is assumed that securities have a multivariate normal distribution but they also work on non-normal securities. Silvapulle and Granger [28] estimated VaR by using ordered statistics and nonparametric kernel estimation of density function. Chen and Tang [10] investigated another nonparametric estimation of VaR for dependent financial returns. Bingham et al. [5] studied VaR by using semi-parametric estimation of VaR based on normal mean-variance mixtures framework. A fully nonparametric estimation of dynamic VaR is also developed by Jeong and Kang [13] based on the adaptive volatility estimation and the nonparametric quantiles estimation. Angelidis and Benos

[3] calculated VaR for Greek Stocks by employing nonparametric methods, such as historical and filtered historical simulation. Recently, the nonparametric quantile regression, along with the extreme value theory, is applied by Schaumburg [25] to predict VaR. All together Using VaR as a risk controlling parameter is the same as variance; a similar framework is applied: variance is replaced by VaR and then it is decreased in a mean-VaR space. In this study value at risk is decreased in a mean-value at risk framework with negative data. Note that value at risk can be negative, so it is unlikely that variance to get non-negative values.

Conventional DEA models, as used by Morey and Morey [23], assume non-negative values for inputs and outputs. These models cannot be used for the case in which DMUs include both negative and positive inputs and/or outputs. Portela et al. [24] consider a DEA model which can be applied in cases where input/output data take positive and negative values. There are also other models can be used for negative data such as Modified slacks-based measure model (MSBM), Sharp et. al. [26], semi-oriented radial measure (SORM), Emrouznejad [11]. In 2015 Lio et al. [18] demonstrated that the linearized diversification models can provide an effective way to approximate portfolio efficiency (PE or Markowitz frontier) provided that the frontier is concave. In this paper, we should prove that with increasing sample size, the frontiers of the linear models with both variance and value at risk, as risk measure, gradually approximate the frontiers of the mean-variance and mean-value at risk models and non-linear model with negative data.

The rest of the paper is organized as follows: Section 2 briefly reviews the portfolio performance literature and have quick look at value at risk as a substitute of variance as a risk parameter. Section 3 goes through the convergence property of the RDM models under the mean-variance framework, which indicates that suitable RDM models with sufficient data can be used to effectively approximate the Portfolio Efficiency (PE). Section 4 presents computational results using Iranian stock companies data and finally conclusions are given in section 5.

## 2 Background

Portfolio theory to investing is published by Markowitz [20]. This approach starts by assuming that an investor has a given sum of money to invest at the present time. This money will be invested for a time as the investor's holding period. The end of the holding period, the investor will sell all of the assets that were bought at the beginning of the period and then either consume or reinvest. Since portfolio is a collection of assets, it is better that to select an optimal portfolio from a set of possible portfolios. Hence the investor should recognize the returns (and portfolio returns), expected (mean) return and standard deviation of return. This means that the investor wants to both maximize expected return and minimize uncertainty (risk). Rate of return (or simply the return) of the investor's wealth from the beginning to the end of the period is calculated as follows:

$$\text{Return} = \frac{(\text{end-of-period wealth}) - (\text{beginning-of-period wealth})}{\text{beginning-of-period wealth}} \quad (2.1)$$

Since Portfolio is a collection of assets, its return  $r_p$  can be calculated in a similar manner. Thus according to Markowitz, the investor should view the rate of return associated to any one of these portfolios as what is called in statistics a random variable. These variables can be described expected the return (mean or  $\bar{r}_p$ ) and standard deviation of return. Expected return and deviation standard of return are calculated as follows:

$$\bar{r}_p = \sum_{j=1}^n \lambda_j \bar{r}_j, \sigma_p = \left[ \sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j \Omega_{ij} \right]^{\frac{1}{2}} \quad (2.2)$$

Where:

$n$  = The number of assets in the portfolio

$\bar{r}_p$  = The expected return of the portfolio

$\lambda_j$  = The proportion of the portfolio's initial value invested in asset  $i$

$\bar{r}_j$  = The expected return of asset  $i$

$\sigma_p$  = The deviation standard of the portfolio

$\Omega_{ij}$  = The covariance of the returns between asset

$i$  and asset  $j$

$$\min \sigma_p = \left[ \sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j \Omega_{ij} \right]^{\frac{1}{2}} \tag{2.3}$$

$$\sum_{j=1}^n \lambda_j \bar{r}_j \geq \alpha$$

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

$$\lambda_j \geq 0 \quad j = 1, \dots, n$$

In the above, optimal portfolio from the set of portfolios will be chosen that maximum expected return for varying levels of risk and minimum risk for varying levels of expected return [27]. Data Envelopment Analysis is a nonparametric method for evaluating the efficiency of systems with multiple inputs and multiple outputs. In this section we present some basic definitions, models and concepts that will be used in other sections in DEA. They will not be discussed in details. Consider  $j$ , ( $j = 1, \dots, n$ ) where each consumes  $m$  inputs to produce  $s$  outputs. Suppose that the observed input and output vectors of  $j$  are  $X_j = (x_{1j}, \dots, x_{mj})$  and  $Y_j = (y_{1j}, \dots, y_{sj})$  respectively, and let  $X_j \geq 0$  and  $X_j \neq 0$ ,  $Y_j$  and  $Y_j \neq 0$ . A basic DEA formulation in input orientation is as follows:

$$\min \theta - \varepsilon \left( \sum_{r=1}^s s_r^+ + \sum_{i=1}^m s_i^- \right)$$

$$\text{s.t.} \quad \sum_{j=1}^n \lambda_j x_{ij} + s_i^- = \theta x_{io} \quad i = 1, \dots, m,$$

$$\sum_{j=1}^n \lambda_j y_{rj} + s_r^+ = y_{ro} \quad r = 1, \dots, s,$$

$$\tag{2.4}$$

Where  $\lambda$  is a  $n$ -vector of  $\lambda$  variables,  $s^+$  as-vector of output slacks,  $s^-$  an  $m$ -vector of input slacks and set  $\Lambda$  is defined as follows:

$$\Lambda = \begin{cases} \{\lambda \in \mathbb{R}_+^n & \text{with constant RTS,} \\ \{\lambda \in \mathbb{R}_+^n, 1 \lambda \leq 1\} & \text{with non-increasing RTS,} \\ \{\lambda \in \mathbb{R}_+^n, 1 \lambda = 1\} & \text{with variable RTS} \end{cases} \tag{2.5}$$

Note that subscript 'o' refers to the unit under the evaluation. A is efficient if  $\theta = 1$  and all slack variables  $s^-, s^+$  equal zero; otherwise it is inefficient. In the DEA formulation above, the left-hand sides in the constraints define an efficient portfolio.  $\theta$  is a multiplier defines the distance from the efficient frontier. The slack variables

are used to ensure that the efficient point is fully efficient. This model is used for asset selection. The portfolio performance evaluation literature is vast. In recent years these models have been used to evaluate the portfolio efficiency. Also in the Markowitz theory, it is required to characterize the whole efficient frontier but the proposed models by Joro & Na do not need to characterize the whole efficient frontier but only the projection points. The distance between the asset and its projection which means the ratio between the variance of the projection point and the variance of the asset is considered as an efficiency measure ( $\theta$ ). In this framework, there is  $n$  assets,  $\lambda_j$  is the weight of asset  $j$  in the projection point,  $r_j$  is the expected return of asset  $j$ ,  $\mu_o$  and  $\delta_o^2$  are the expected return and variance of the asset under evaluation respectively. Efficiency measure  $\theta$  can be solved via following model:

$$\min \theta - \varepsilon(s_1 + s_2)$$

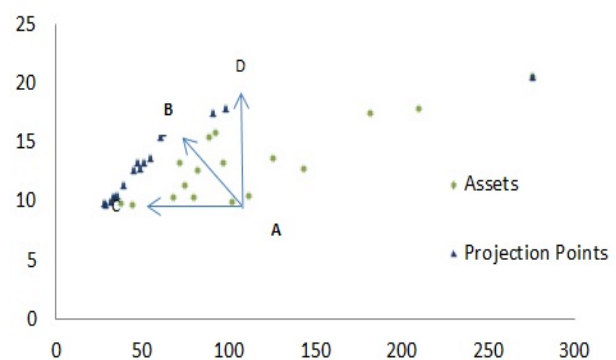
$$\text{s.t.} \quad E \left[ \sum_{j=1}^n \lambda_j r_j \right] - s_1 = \mu_o,$$

$$E \left[ \left( \sum_{j=1}^n \lambda_j (r_j - \mu_j) \right)^2 \right] + s_2 = \theta \delta_o^2$$

$$\sum_{j=1}^n \lambda_j \leq 1 \quad \forall \lambda \geq 0$$

$$\tag{2.6}$$

Model (2.6) is revealed by the non-parametric efficiency analysis Data Envelopment Analysis (DEA). Fig 1 illustrates different projection that



**Figure 1:** Different projections (input oriented, output oriented, combination oriented).

consist of input oriented, output oriented and combination oriented in models of data envelopment analysis. C is the projection point obtained

via fixing expected return and minimizing variance,  $B$  via maximizing return and minimizing variance simultaneously, and  $D$  via fixing variance and maximizing return.

In the conventional DEA models, each  $j$  ( $j = 1, \dots, n$ ) is specified by a pair of non-negative input and output vectors  $(x_j, y_j) \in \mathbb{R}_+^{m+s}$ , in which inputs  $x_{ij}$  ( $i = 1, \dots, m$ ) are utilized to produce outputs,  $y_{rj}$  ( $r = 1, \dots, s$ ). These models can not be used for the case in which DMUs include both negative and positive inputs and/or outputs. Poltera et al. (2004) [24] consider a DEA model which can be applied in the cases where input/ output data take positive and negative values. Rang Directional Measure (RDM) model proposed by Poltera et al. goes as follows:

$$\begin{aligned} \max \quad & \beta \\ \text{s.t.} \quad & \sum_{j=1}^n \lambda_j x_{ij} \leq x_{io} - \beta R_{io} \quad i = 1, \dots, m, \\ & \sum_{j=1}^n \lambda_j y_{rj} \geq y_{ro} + \beta R_{ro} \quad r = 1, \dots, s, \\ & \sum_{j=1}^n \lambda_j = 1, \end{aligned} \tag{2.7}$$

Ideal point ( $I$ ) within the presence of negative data, is  $I = (\max_j \{y_{rj} : r = 1, \dots, s\}, \min_j \{x_{ij} : i = 1, \dots, m\})$  where

$$\begin{aligned} R_{io} &= x_{io} - \min_j \{x_{ij} : j = 1, \dots, n\}, i = 1, \dots, m, \\ R_{ro} &= \max_j \{y_{rj} : j = 1, \dots, n\} - y_{ro}, r = 1, \dots, s. \end{aligned} \tag{2.8}$$

Here, according to used inputs and outputs, variance (risk parameter) is used as input and mean of returns is used as output in RDM model.

The other models solve negative data such as Modified slacks-based measure model (MSBM), Emrouznejad [11], semi-oriented radial measure (SORM), Sharp et al. [26] and etc. Extremely, we present following non-linear mean-variance RDM

model on the basis of negative data:

$$\begin{aligned} \max \quad & \beta \\ \text{s.t.} \quad & E \left[ \sum_{j=1}^n \lambda_j r_j \right] \geq \mu_o + \beta R_{\mu_o} \\ & E \left[ \left( \sum_{j=1}^n \lambda_j (r_j - \mu_j) \right)^2 \right] \leq \sigma_o^2 - \beta R_{\sigma_o^2} \\ & \sum_{j=1}^n \lambda_j = 1 \quad \lambda \geq 0 \end{aligned} \tag{2.9}$$

Ideal point ( $I$ ) within the presence of negative data, is  $I = (\min_j \{\sigma_j^2\}, \max_j \{\mu_j\})$  where

$$\begin{aligned} R_{\mu_o} &= \max_j \{\mu_j : j = 1, \dots, n\} - \mu_o \\ R_{\sigma_o^2} &= \sigma_o^2 - \min_j \{\sigma_j^2 : j = 1, \dots, n\} \end{aligned} \tag{2.10}$$

The above model can be expressed as following:

$$\begin{aligned} \max \quad & \beta \\ \text{s.t.} \quad & E[r(\lambda)] \geq \mu_o + \beta R_{\mu_o} \\ & \text{Var}[r(\lambda)] \leq \sigma_o^2 - \beta R_{\sigma_o^2} \\ & \sum_{j=1}^n \lambda_j = 1 \quad \lambda \geq 0 \end{aligned} \tag{2.11}$$

However, it can be shown that as  $n$  increases, model (2.7) Converges to models (2.3) & (2.9).

Beside variance as a risk parameter which has its positive and negative sides, value at risk (VaR) is another risk parameter with different characteristics. To calculate VaR generally there is no need that return's distributions come from a normal basis, although, the way which is used to obtain VaR, is important.

Value at Risk (VaR) is defined as maximum amount of invest that one may loss in a specified time period. Statistically VaR is defined as the percentile of a distribution.

$$p(\Delta P_k > \text{VaR}) = 1 - \alpha$$

Calculation of VaR can be done through different methods. Historical and Monte Carlo simulations and variance-covariance are three mostly used methods. This paper uses Historical simulation and variance-covariance method to calculate VaR and tries to compare portfolios efficiencies by using mean-VaR models. Mean-VaR models basically are same as mean-variance models. However, method used to calculate VaR determines that the model is linear or non-linear.

In historical simulation VaR is calculated based on what happened before. In this method there is no need to be aware of returns distributions over time. Simply returns are order in an ascending way and preferred percentile is value at risk. In fact this method is completely based on what happened before and this is its downside point. When no historical data is available or data have trend, using this method is either impossible or leads to inaccurate results. In contrast with historical simulation, variance-covariance method is based on returns distributions. In this method, in the first step, a normality check should be used to get sure, returns come from normal distribution. On the next step, VaR or appropriate percentile is calculated through formulas obtained based on normal distribution. Value at risk which is calculated from variance covariance method has its own negative aspects. Wrong distribution assumption, non-stationary variables causes by changes happen over time are two of this methods negative points. To calculate VaR from normally distributed returns, consider VaR formulas.

$$p(\Delta P_k > -\text{VaR}) = 1 - \alpha \tag{2.12}$$

$$\Delta_h P_t = P_{t+h} - P_t$$

By normalizing equation (2.12) we have:

$$P\left(\frac{\Delta_h P_t - \mu_t}{\sigma_t} < \frac{-\text{VaR} - \mu_t}{\sigma_t}\right) = \alpha \tag{2.13}$$

where  $Z_\alpha = \frac{\Delta_h P_t - \mu_t}{\sigma_t}$  and  $Z_\alpha \equiv \Phi^{-1}(\alpha)$ ,  $\frac{1}{2} < \alpha < 1$ , therefore,

$$-Z_\alpha = \frac{-\text{VaR}_\alpha - \mu_t}{\sigma_t},$$

and so,

$$\text{VaR}_\alpha = \sigma_t Z_\alpha - \mu_t \tag{2.14}$$

In later sections mean-Var models are introduced and are used to evaluate portfolios efficiencies.

### 3 Theoretical foundation of RDM approach: Convergence property

Linear DEA models can't obtain true solution for estimating portfolio efficient and non-linear RDM frontier. As know, non-linear RDM and Markowitz frontier are concave; therefore we should prove that with increasing sample size,

the frontiers of the RDM linear models gradually approximate the frontiers of the mean-variance model and RDM non-linear model.

Assumption: Suppose there exists a probability density function  $p(x)$  of  $x \in \Omega$  satisfying  $\forall x^0 \in \Omega$ , there exists a set  $S(x^0) = U(x^0, \xi) \cap \Omega$  such that  $\int_{S(x^0)} p(x) dx > 0$ , where  $U(x^0, \xi)$  is a neighborhood of  $x^0$ .

Let  $\Psi = \left\{ (r, \sigma) \mid \sum_{j=1}^n \lambda_j r_j \geq r, \sum_{j=1}^n \lambda_j \sigma_j \leq \sigma, \sum_{j=1}^n \lambda_j = 1, \lambda_j \geq 0, j = 1, \dots, n \right\}$ , then the

RDM frontier is formed by the outer envelope (upper left boundary) of  $\Psi$  as show in Figure 2.

**Theorem 3.1** Let  $\bar{r}_p = h(\sigma_p)$  be the portfolio frontier without risk-free assets and  $\bar{r}_p^* + \beta R_{\bar{r}_p^*} = h_n^*(\sigma_p - \beta R_{\sigma_p})$  be the RDM frontier with  $n$  portfolio samples. Then  $h_n^*(\sigma_p - \beta R_{\sigma_p})$  converges to  $h(\sigma_p)$  in probability when  $n \rightarrow +\infty$ .

**Proof.** For any  $A = (\sigma_p^a, h(\sigma_p^a))$  on the efficient portfolio frontier, there exists  $x^a \in \Omega$ , such that  $(\sigma_p(x^a), \bar{r}_p(x^a)) = (\sigma_p^a, h(\sigma_p^a))$ .

Since  $k(x) \in (\sigma_p(x), \bar{r}_p(x))$  is continuous on  $x$ , there exists  $\varepsilon > 0$ , such that  $k^{-1}(U(A, \varepsilon))$  is an open set, where  $U(A, \varepsilon)$  is a neighborhood of  $A$  and

$$x^a \in k^{-1}(U(A, \varepsilon))$$

Thus

$$\exists \xi > 0, \text{ s.t. } S(x^a) = U(x^a, \xi) \cap \Omega \subseteq k^{-1}(U(A, \varepsilon)).$$

Due to the assumption on the probability density function  $p(x)$ , we have

$$q(U(A, \varepsilon)) = \int_{k^{-1}(U(A, \varepsilon))} p(x) dx > 0$$

$$\geq \int_{S(x^a)} p(x) dx > 0$$

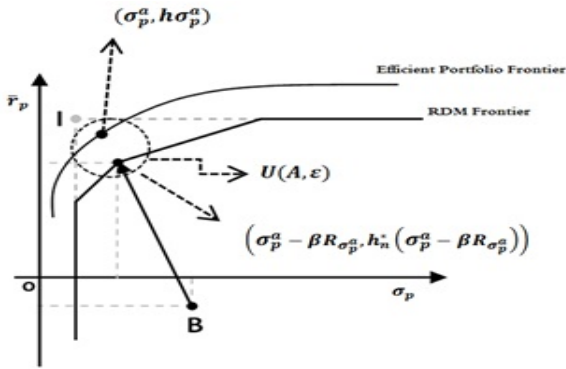
Let  $T$  represents the event that the expected returns and standard derivations of all the  $n$  portfolio samples that are not in  $U(A, \varepsilon)$ . Therefore, the probability of  $T$  can be expressed as

$$Pr(T) = (1 - q(U(A, \varepsilon)))^n$$

It follows from the definition of the concavity of efficient portfolio frontier and RDM frontier that  $Pr\{ |h(\sigma_p^a) - h_n^*(\sigma_p^a - \beta R_{\sigma_p^a})| > \varepsilon \} =$

$$Pr\{h(\sigma_p^a), h_n^*(\sigma_p^a - \beta R_{\sigma_p^a}) > \varepsilon\} \leq Pr(T) = (1 - q(U(A, \varepsilon)))^n \rightarrow 0, \text{ when } n \rightarrow \infty.$$

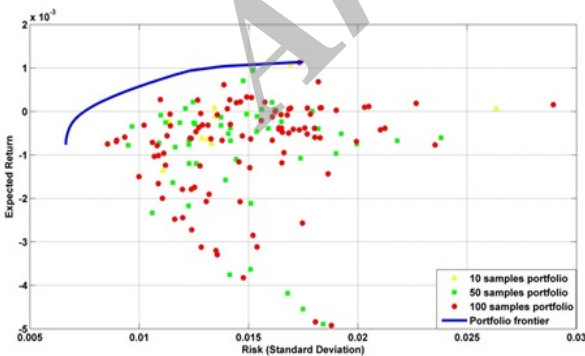
Because  $\sigma_p^a$  is arbitrary, we obtain the conclusion that  $h_n^*(\sigma_p - \beta R_{\sigma_p})$  converges to  $h = (\sigma_p)$  in probability, as shown in Figure 2.



**Figure 2:** Convergence explanation. It can be seen as  $n$  increases RDM frontier converges to Markowitz frontier.

### 4 Application in Iranian Stock Companies

In this section, we verify the validity of the above-discussed results using illustrative examples. 15 stocks from the Iranian stock companies are selected, which are monthly data from 21 April 2014 to 21 June 2014. Their statistical properties are shown in Table ???. We then randomly generated  $n = 10, 50, 100$  weights using MATLAB to construct portfolio samples. Efficiencies of sample portfolios are evaluated with model (2.7). In



**Figure 3:** Portfolios with different sample sizes.

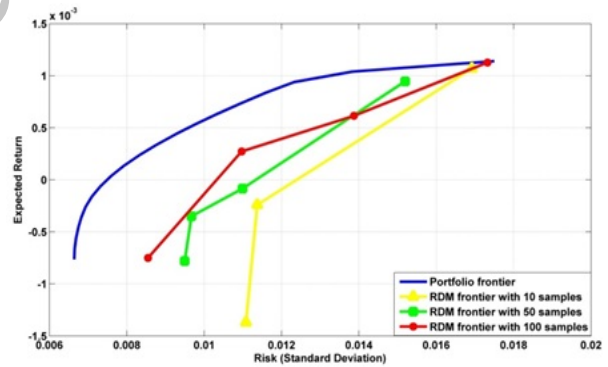
Figure 3, portfolios with sample sizes 10, 50 and 100 constructed by random weights are shown. In this figure blue curve shows efficient frontier obtained by Markowitz model (non-linear model).

By evaluating portfolios' efficiencies using linear models, it can be seen linear efficient frontier converges to non-linear efficient frontier as  $n$  increases (Figure 3). In Table 1 statistics of 10

**Table 1:** Basic statistics of 10 random portfolios made by 15 under evaluation assets.

Portfolio number	Mean	variance	Efficiency Mean-var model ( $\beta$ )
1	-0.0013	0.00012	0.00
2	-0.0001	0.00018	0.18
3	-0.0006	0.00016	0.33
4	-5.5E-05	0.00022	0.30
5	-0.0002	0.00012	0.00
6	-0.0007	0.00017	0.40
7	0.0010	0.00028	0.00
8	6.08E-05	0.00069	0.76
9	-0.0006	0.00017	0.36
10	8E-05	0.00018	0.07

sample portfolios are provided. Statistics for 50 and 100 samples can be calculated in a same way. As we can see in Table 1 portfolios 1, 5 and 7 are efficient ones. In fact these portfolios are yellow dots in Figure 3, where the efficient frontier breaks. In Figure 4 efficient RDM frontier for 50



**Figure 4:** This figure shows as number of samples increases, linear efficient frontier gets closer to non-linear efficient frontier (Portfolio Frontier).

and 100 samples of portfolios are shown. It is obvious by increasing number of samples, linear RDM frontier converges to non-linear frontier.

The three polylines in Figure 4 and 5 are the envelope frontiers constructed by RDM linear models with 10, 50 and 100 samples. The top curve is the efficient frontier calculated by the Markowitz mean-variance model.

Figure 5 also shows projection of under eval-

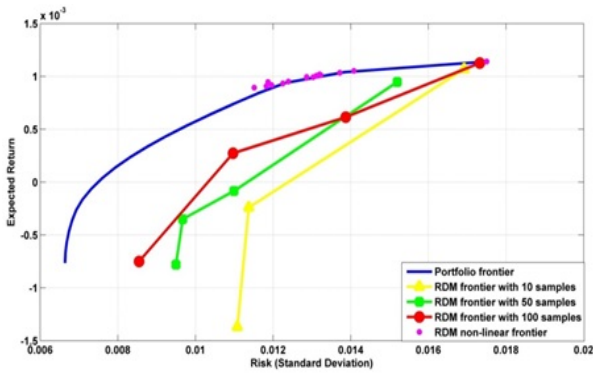


Figure 5: Purple dots represent projection of under evaluation assets on the efficient frontier.

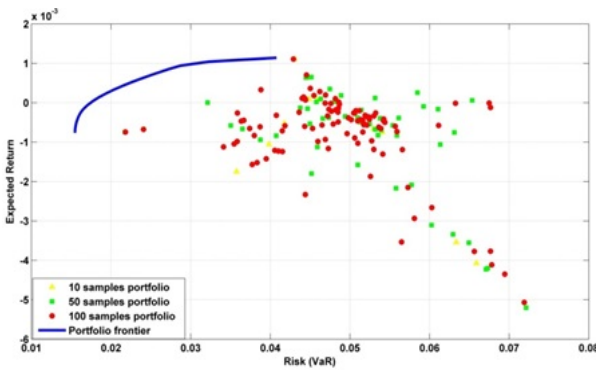


Figure 6: Mean-VaR region and portfolios. In this plot value at risk is calculated in confidence level of 99%.

uation assets on the efficient frontier which are obtained by solving of non-linear RDM model.

As mentioned in section 2, one may uses value at risk as a risk parameter due to its positive aspect. In continue, value at risk through two different methods for 15 under evaluation assets are calculated. First of all, VaR is calculated by using historical simulation. In this method returns are sorted in an ascending way and appropriate percentile is calculated. Statistics are provided in Table 2. Means are calculated through equation (2.2).

In Table 2, it can be found as the level of value at risk confidence level increases, amount of value at risk gets larger. It illustrates by increasing confidence level investor gets more sure how much money may lose in a specified period of investment. Same results for 10 sample portfolios made by 15 assets are provided below. (Table 3)

Figures 6-8 show portfolios position in a mean-value at risk region. In figures 9-11, we can see as the number of samples increases same as mean-

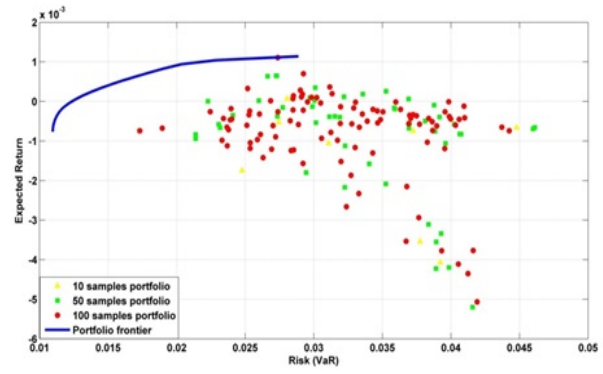


Figure 7: Mean-VaR region and portfolios. In this plot value at risk is calculated in confidence level of 95%.

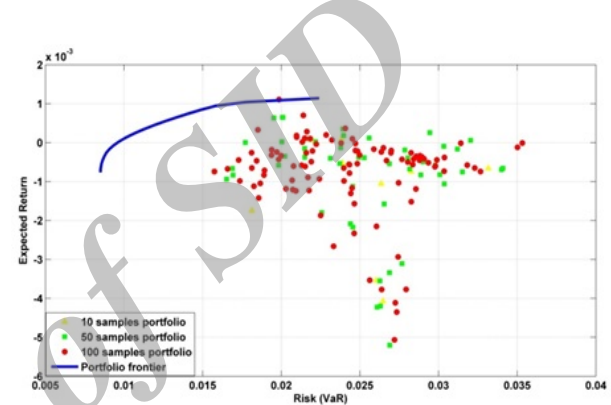


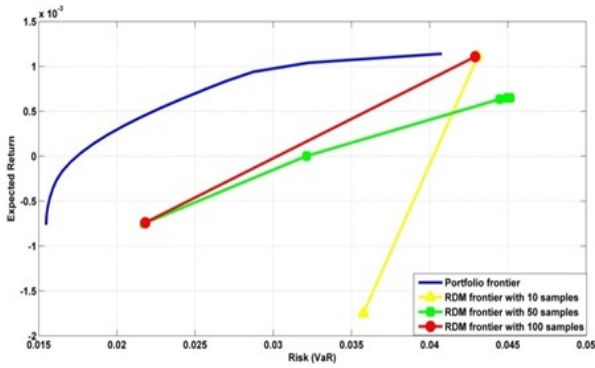
Figure 8: Mean-VaR region and portfolios. In this plot value at risk is calculated in confidence level of 90%.

Table 2: Mean and value at risk on under evaluation asstes.

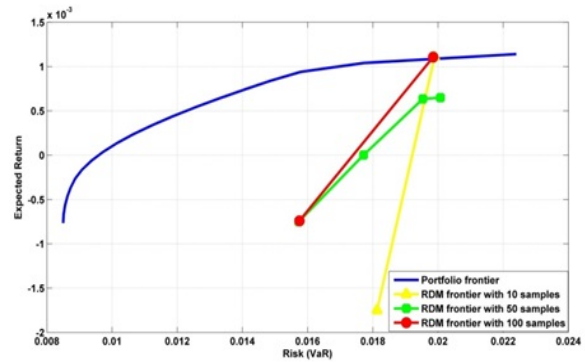
Asset number	Mean	Value at Risk		
		90%	95%	99%
1	-0.0007	0.0158	0.0173	0.0217
2	-0.0003	0.0286	0.0370	0.0524
3	-0.0007	0.0347	0.0470	0.0542
4	0.0007	0.0197	0.0269	0.0451
5	0.0001	0.0243	0.0296	0.0439
6	0.0001	0.0373	0.0420	0.0731
7	-0.0053	0.0271	0.0420	0.0729
8	-0.0006	0.0299	0.0405	0.0559
9	-0.0004	0.0191	0.0222	0.0503
10	-0.0011	0.0139	0.0194	0.0298
11	0.0001	0.0210	0.0294	0.0454
12	0.0011	0.0198	0.0273	0.0431
13	-0.0025	0.0233	0.0315	0.0432
14	-0.0003	0.0291	0.0411	0.0545
15	-0.00144	0.0288	0.0321	0.0377

variance framework linear efficient frontier converges to non-linear frontier.

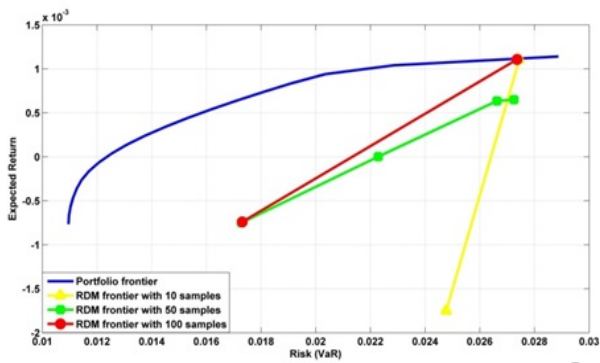




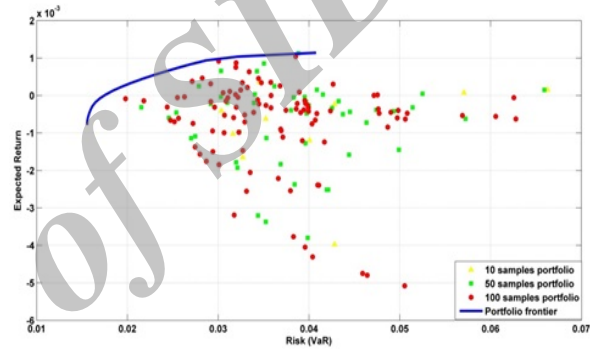
**Figure 9:** In this figure it can be seen that same as mean-variance models as the number of samples increases linear efficient converges to non-linear frontier. In this figure value at risk is calculated on 99% confidence level.



**Figure 11:** In this figure it can be seen that same as mean-variance models as the number of samples increases linear efficient converges to non-linear frontier. Value confidence level in this figure is 90%.



**Figure 10:** In this figure it can be seen that same as mean-variance models as the number of samples increases linear efficient converges to non-linear frontier. Value confidence level in this figure is 95%.



**Figure 12:** Sample portfolios in mean-VaR region. Value at risk in calculated based on a 99% confidence level.

**Table 3:** Mean and value at risk of 10 portfolios made by under evaluation assets are calculated. Three last columns are their efficiencies and mean-VaR models outputs. Based on this model assets 5 and 8 are efficient.

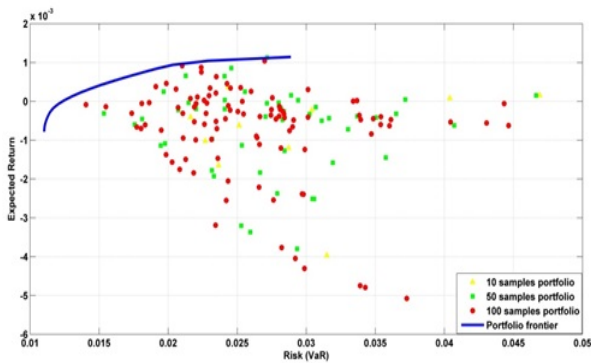
Portfolio number	Mean	Value at Risk			Efficiency Mean-VaR model ( $\beta$ )		
		90%	95%	99%	90%	95%	99%
1	-0.0007	0.0282	0.0372	0.0540	0.84	0.81	0.68
2	-0.0005	0.0238	0.0274	0.0417	0.74	0.36	0.28
3	-0.0010	0.0263	0.0310	0.0398	0.82	0.68	0.24
4	-0.0035	0.0260	0.0377	0.0633	0.84	0.85	0.82
5	-0.0017	0.0181	0.0247	0.0357	0.00	0.00	0.00
6	6.84E-05	0.0219	0.0280	0.0458	0.61	0.37	0.43
7	-0.0040	0.0264	0.0392	0.0658	0.85	0.86	0.83
8	0.0011	0.0198	0.0274	0.0430	0.00	0.00	0.00
9	9.43E-05	0.0238	0.0300	0.0448	0.72	0.57	0.38
10	-0.0006	0.0331	0.0447	0.0529	0.89	0.88	0.66

In all mentioned figure, value at risk is calculated through historical method, and it was clear that linear frontier is convergence to non-linear frontier. However in some cases frontiers cross

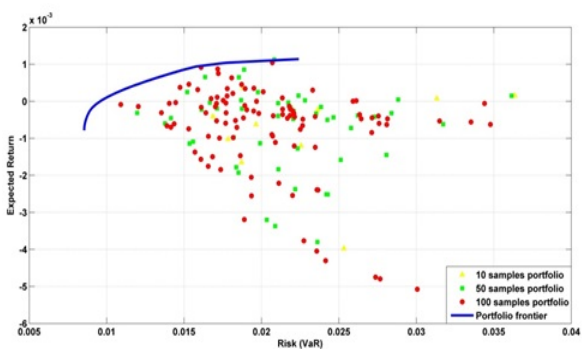
each other, mainly they ordered in the way we expect.

Same results are obtained if values at risks are calculated via variance-covariance method. In this method, returns have to come from normal distribution. First of all by using Anderson-Darling normality test [1], distributions of returns of under evaluation assets are checked. Returns of 13 assets were normally distributed. For normally distributed assets, expected returns and their value at risks are calculated by using formulas (2.2). Results are provided in Table 4. Mean and values at risks of normally distributed assets on all level of risk confidence.. Same as historical method, as the risk confidence level increases value at risk of an assets gets larger. Therefore, investor gets surer the amount of risk that may face.

As said before three series of portfolios are made to show by increasing number of samples,



**Figure 13:** Sample portfolios in mean-VaR region. Value at risk in calculated based on a 95% confidence level.



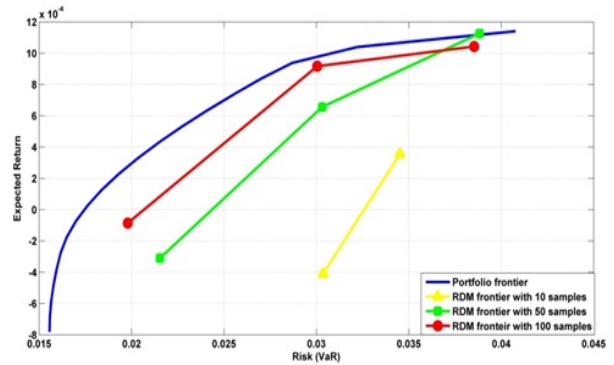
**Figure 14:** Sample portfolios in mean-VaR region. Value at risk in calculated based on a 90% confidence level.

**Table 4:** Mean and values at risks of normally distributed assets on all level of risk confidence.

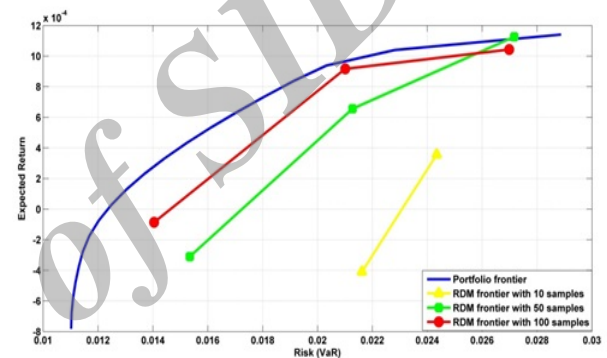
Asset number	Mean	Value at Risk		
		90%	95%	99%
1	-0.0007	0.0154	0.0197	0.0275
2	-0.0003	0.0309	0.0397	0.0560
3	-0.0007	0.0319	0.0410	0.0576
4	0.0007	0.0234	0.0303	0.0431
5	0.0001	0.0272	0.0351	0.0497
6	0.0001	0.0374	0.0482	0.0682
7	-0.0053	0.0311	0.0386	0.0523
8	-0.0006	0.0350	0.0450	0.0633
9	-0.0004	0.0228	0.0293	0.0412
10	0.0011	0.0212	0.0277	0.0396
11	-0.0025	0.0250	0.0315	0.0434
12	-0.0003	0.0278	0.0358	0.0504
13	-0.0014	0.0254	0.0323	0.0451

linear frontier converges to non-linear one. In Table 5 efficiencies of 10 random portfolios made by normal assets are provided.

In Figures 15-17, linear efficient frontier of each



**Figure 15:** Linear and non-linear frontiers with different sample size of portfolios. Linear frontier converges to non-linear frontier and n increases. Value at risk in this figure is calculated in a 99% confidence level.



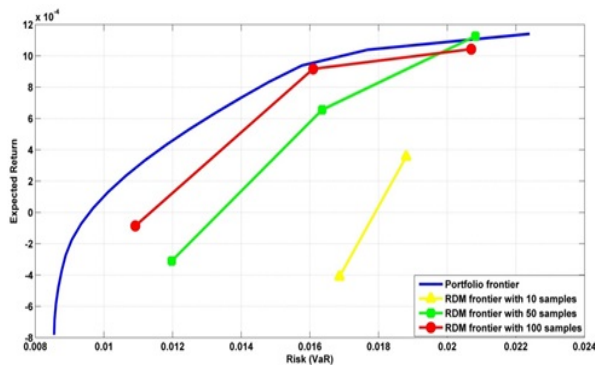
**Figure 16:** Linear and non-linear frontiers with different sample size of portfolios. Linear frontier converges to non-linear frontier and n increases. Value at risk in this figure is calculated in a 95% confidence level.

series of sample portfolios are shown.

**Table 5:** Mean, value at risk and efficiencies of 10 sample portfolios.

Portfolio number	Mean	Value at Risk			Efficiency Mean-VaR model ( $\beta$ )		
		90%	95%	99%	90%	95%	99%
1	-0.0039	0.0253	0.0314	0.0428	0.90	0.89	0.88
2	-0.0012	0.0225	0.0287	0.0400	0.80	0.78	0.77
3	-0.0006	0.0196	0.0251	0.0352	0.63	0.61	0.59
4	-0.0004	0.0168	0.0216	0.0303	0.00	0.00	0.00
5	-0.0016	0.0187	0.0236	0.0327	0.72	0.70	0.69
6	-0.0002	0.0236	0.0303	0.0428	0.76	0.75	0.73
7	7.6E-05	0.0313	0.0403	0.0570	0.87	0.86	0.85
8	-0.0010	0.0178	0.0226	0.03162	0.57	0.54	0.53
9	0.0003	0.0188	0.0243	0.0345	0.00	0.00	0.00
10	0.0001	0.0363	0.0469	0.0662	0.90	0.90	0.89

In Table 5 Mean, value at risk and efficiencies of 10 sample portfolios. it can be seen that portfolios number 4 and 9 are efficient on all levels of risk confidence. In figures 12-14 all sample portfolios and non-linear efficient frontier are shown.



**Figure 17:** Linear and non-linear frontiers with different sample size of portfolios. Linear frontier converges to non-linear frontier and  $n$  increases. Value at risk in this figure is calculated in a 90% confidence level.

## 5 Conclusion

In this paper, under section 2, mean-variance, linear and non-linear RDM models are discussed. In later parts value at risk as a new risk parameter was discussed. We had also a quick review over methods of VaR calculation and talked about positive and negative aspects of each method. In section 3 a theorem discussed and proved that by increasing number of samples linear RDM frontier convergence to Markowitz frontier. So RDM linear models can be used to estimate portfolios efficiencies and actual efficient frontier.

In the last section, all discussed topics, with a random sample of stocks data from Tehran stock, was tested. 15 stocks from Tehran stock were randomly gathered and their prices over 60 days were gathered. It was shown, as the number of samples increase, whether consider variance or value at risk as a risk parameter, RDM linear frontier converges to Markowitz frontier.

## References

- [1] T. W. Anderson, D. A. Darling, *A Test of Goodness-of-Fit*, Journal of the American Statistical Association 49 (1954) 765-769.
- [2] D. Ariditti, *Skewness and investors decisions: A reply*, Journal of Financial and Quantitative Analysis 10 (1975) 173-176.
- [3] T. Angelidis, A. Benos, *Value-at-Risk for Greek stocks* Multinational Finance Journal 12 (2008) 67-104.
- [4] W. J. Baumol, *An Expected Gain-Confidence Limit Criterion for Portfolio Selection*, Management Science 10 (1963) 174-182.
- [5] N.H. Bingham, R. Kiesel, R. Schmidt, *A Semi-Parametric Approach To Risk Management*, Quantitative Finance 6 (2003) 426-441.
- [6] W. Bric, K. Kerstens, J. B. Lesourd, *Single Period Markowitz Portfolio selection, Performance Gauging and Duality: A Variation on The Luenberger shortage Function*, Journal of Optimization Theory and Applications 120 (2004) 1-27.
- [7] W. Bric, K. Kerstens, O. Jokung, *Mean-Variance-Skewness Portfolio Performance Gauging: A General Shortage Function and Dual Approach*, Management Science 53 (2007) 135-149.
- [8] A. Charnes, W. W. Cooper, A. Y. Seiford, *Data Envelopment Analysis: Theory, Methodology and Applications*, Kluwer Academic Publishers, Boston 1994.
- [9] A. Charnes, W. W. Cooper, E. Rhodes, *Measuring Efficiency of Decision Making Units*, European Journal of Operational Research 2 (1978) 429-444.
- [10] S. X. Chen, C. Y. Tang, *Nonparametric Inference of Value at Risk for Dependent Financial Returns*, Journal of financial econometrics 12 (2005) 227-55.
- [11] A. Emrouznejad, *A Semi-Oriented Radial Measure for Measuring The Efficiency Of Decision Making Units With Negative Data, Using DEA*, European journal of Operational Research 200 (2010) 297-304.
- [12] Y. K. Ho, Y. L. Cheung, *Behavior of Intra-Daily Stock Return on an Asian Emerging Market*, Applied Economics 23 (1991) 957-966.
- [13] S.O. Jeong, K. H. Kang, *Nonparametric Estimation of Value-At-Risk*, Journal of Applied Statistics 10 (2009) 1225-38.
- [14] T. Joro, P. Na, *Portfolio Performance Evaluation in a Mean-Variance-Skewness Framework*, European Journal of Operational Research 175 (2005) 446-461.

- [15] A. Kane, *Skewness Preference and Portfolio Choice*, Journal of Financial and Quantitative Analysis 17 (1982) 15-25.
- [16] K. Kerstens, A. Mounir, I. Woestyne, *Geometric Representation of The Mean-Variance-Skewness Portfolio Frontier Based Upon The Shortage Function*, European Journal of Operational Research 10 (2011) 1-33.
- [17] A. Kraus, R. H. Litzenberger, *Skewness Preference and the Valuation of Risk Assets*, Journal of Finance 31 (1976) 1085-1100.
- [18] W. Liu, Z. Zhou, D. Liu, H. Xiao, *Estimation of Portfolio Efficiency via DEA*, Omega 52 (2015) 107-118.
- [19] B. Mandelbrot, *The Variation of Certain Speculative Prices*, Journal of Business 36 (1963) 394-419.
- [20] H. M. Markowitz, *Portfolio Selection*, Journal of Finance 7 (1952) 77-91.
- [21] M. Mhiri, J. Prigent, *International Portfolio Optimization with Higher Moments*, International Journal of Economics and Finance 5 (2010) 157-169.
- [22] T. Mitton, K. Vorknik, *Equilibrium under Diversification and the Preference Of Skewness*, Review of Financial Studies 20 (2007) 1255-1288.
- [23] M. R. Morey, R. C. Morey, *Mutual Fund Performance Appraisals: A Multi-Horizon Perspective With Endogenous Benchmarking*, Omega 27 (1999) 241-258.
- [24] M. C. Portela, E. Thanassoulis, G. Simpson, *A directional distance approach to deal with negative data in DEA: An application to bank branches*, Journal of Operational Research Society 55 (2004) 1111-1121.
- [25] J. Schaumburg, *Predicting Extreme Value at Risk: Nonparametric Quantile Regression with Refinements from Extreme Value Theory*, Computational Statistics and Data Analysis 56 (2012) 4081-4096.
- [26] J. A. Sharp, W. Meng, W. Liu, *A Modified Slacks-Based Measure Model for Data Envelopment Analysis with Natural Negative Outputs and Inputs*, Journal of the Operational Research Society 57 (2006) 1-6.
- [27] W. F. Sharpe, *Investment*, Third Edition, Prentice-Hall (1985).
- [28] P. Silvapulle, CW. Granger, *Large Returns, Conditional Correlation and Portfolio Diversification: A Value-At-Risk Approach*, Quantitative Finance 10 (2001) 542-51.



Masoud Sanei is Associate professor of Applied Mathematics in operational Research at Islamic Azad University Central Tehran Branches. His research interest are in the areas of Applied Mathematics, Data Envelopment Analysis, Supply Chains, Finance. He has published research articles in international journals of Mathematics. He is referee and editor of mathematical journals.



Shokoofeh Banihashemi is Assistant professor of Applied Mathematics in operational research at Allameh Tabatabaai University. Her research interests are in the areas of Applied Mathematics, Data Envelopment Analysis, Supply Chains, Finance. She has published research articles in international journals of Mathematics. She is referee of mathematical and economics journals.



Marzieh Kaveh: she has received her master in Applied Mathematics in Operational Research at Islamic Azad University Central Tehran Branches.

## Estimation of portfolio efficient frontier by different measures of risk via DEA

M. Sanei, S. Banihashemi, M. Kaveh

تخمین مرز کارای سبددارایی با سنج‌های متفاوت ریسک از طریق تحلیل پوششی داده‌ها

### چکیده:

در این مقاله، مدل‌های خطی تحلیل پوششی داده‌ها برای تخمین مرز کارای مارکوویتز مورد استفاده قرار می‌گیرند. مدل‌های سنتی تحلیل پوششی داده‌ها مقادیر نامنفی برای ورودی‌ها و خروجی‌ها در نظر می‌گیرند. اگرچه واریانس به عنوان سنج ریسک، تنها متغیری در این مدل‌ها است که مقادیر نامنفی می‌پذیرد. بنابراین مدل‌های داده منفی که ریسک‌دارایی‌ها به عنوان ورودی و میانگین بازده به عنوان خروجی در نظر گرفته می‌شوند. در ابتدا واریانس به عنوان سنج ریسک در نظر گرفته می‌شد. اگرچه هم در تئوری و هم در عمل اشاره دارد که واریانس سنج خوبی برای ریسک نمی‌باشد. پس ارزش در معرض خطر به عنوان سنج ریسک جدید معرفی شده است. در این مقاله ما ثابت می‌کنیم که با افزایش سایز نمونه مرزهای مدل‌های خطی با سنج‌های ریسک واریانس و ارزش در معرض خطر تدریجاً به مرزهای مدل‌های میانگین-واریانس و میانگین-ارزش در معرض خطر و مدل غیر خطی داده منفی میل می‌کنند. نهایتاً مثال عددی با سنج‌های ریسک واریانس و ارزش در معرض خطر که از روش‌های شبه‌سازی تاریخی و واریانس-کوواریانس بدست آمده است برای اثبات مفید و موثر بودن ادعایمان استفاده شده است.