TWO NONPARAMETRIC APPROACHES TO MEAN ABSOLUTE DEVIATION PORTFOLIO SELECTION MODEL

ZHIFENG DAI AND HUAN ZHU

College of Mathematics and Statistics, Changsha University of Science and Technology
Changsha 410114, China
Hunan Provincial Key Laboratory of Mathematical Modeling and Analysis in Engineering,
Changsha University of Science and Technology, Changsha 410114, China

FENGHUA WEN*

Department of finance, College of business, Central South University
Hunan, 410083, China

(Communicated by Stefan Wolfgang Pickl)

Abstract. In this paper, we apply two nonparametric approaches to mean absolute deviation (MAD) portfolio selection model. The first one is to use the nonparametric kernel mean estimation to replace the returns of assets with five different kernel functions. Then, we construct the nonparametric kernel mean estimation-based MAD portfolio model. The second one is to utilize the nonparametric kernel median estimation to replace the returns of assets with five different kernel functions. Then, we construct the nonparametric kernel median estimation-based MAD portfolio model. We also extend the two kinds of nonparametric approach to mean-Conditional Value-at-Risk portfolio model. Finally, we give the in-sample and out-of-sample analysis of the proposed strategies and compare the performance of the proposed models by using actual stock returns in Shanghai stock exchange of China. The experimental results show the nonparametric estimation-based portfolio models are more efficient than the original portfolio model.

1. Introduction. The pioneer work of Markowitz [27] on portfolio selection problem is the milestone of modern finance theory for optimal portfolio construction, investment diversification, and asset allocation. The MV model is formulated as a quadratic programming problem which is a trade off between the return (the portfolio mean) and the risk (the portfolio variance). If we have \( n \) assets to manage, to build the MV model, we have to estimate the covariance matrix with \( n(n+1)/2 \) constants through historical data. As the number of asset in the portfolio selection problem increases, the size of the covariance matrix becomes large, and it is difficult to estimate it. This computational difficulty in the estimation of covariance matrix can be substantially alleviated through the use of the factor models (Sharpe [35]) or sparse matrix techniques (Pang [30]).

2020 Mathematics Subject Classification. Primary: 90C30, 90C90; Secondary: 65K05.

Key words and phrases. Portfolio selection, nonparametric estimation method, mean absolute deviation model.

* Corresponding author: wfh@amss.ac.cn.
Different from Markowitz’s MV model, Konno and Yamazaki [24] took absolute deviation of portfolio as a risk measure and proposed the mean absolute deviation (MAD) model to select a portfolio. Their model can overcome the difficulties associated with the classical MV model while maintaining its advantages over equilibrium models. In particular, their model can be formulated as a linear program, so that large-scale portfolio selection problems may be easily solved. The similar ideas on the risk measure for a linear programming formulation extended the model such as Chiodi et al. [10]; Kellerer et al. [22]; Mansini et al. [26]; and Papahristodoulo and Dotzauer [29]; Perold [31]; Speranza ([41], [42]).

However, conventional portfolio selection models assume that the future condition of a stock market can be accurately predicted by historical data or the risk factors (random variables) with a known probability distribution. No matter how accurate the past data is, this premise will not exist in financial markets due to the high volatility of market environments. It is well known that the nonparametric estimation method is appealing in order to reduce these impacts (see Li and Racine [25]). Moreover, nonparametric estimation method enjoys some nice properties. One of these is that little or no restrictive prior information on functionals is needed. Another advantage is that it allows a wide range of data dependence, which makes it adaptable in the context of financial losses. (see Yao et al. [48]). Therefore, in recent years, many scholars paid attention to the computational formulas for risks measure by using the nonparametric or semiparametric estimation method. For example, applying the nonparametric estimations method to Value-at-Risk (VaR; see Morgan [21]), from different perspective (see, eg., Alemany et al. [1]; Bingham et al. [6]; Chen and Tang [8]; He et al.[18]; Jeong and Kang [20]; Schaumburg [38]); studying the non-parametric Conditional Value-at-Risk (CVaR; see Rockfellar and Uryasev [32]) estimator from different perspective (see, eg., Scaillet [34]; Cai and Wang [7]; Chen [9]; Xiao et al.[47]; Yu et al. [50]). However, these papers largely focus on applying nonparametric estimations method to measure risk, they do not consider the portfolio optimization problem.

Recently, Yao et al. [48] obtained the CVaR estimation formula by using the nonparametric estimator of the loss function density, and two mean-CVaR models based on nonparametric estimation are formed by selecting different bandwidths. And they proved that when the CVaR value is solved, the corresponding VaR value can also be solved. Finally, through the empirical analysis, it is shown that the mean-CVaR portfolio model based on nonparametric estimation is superior to the mean-CVaR portfolio model. Yao et al. [49] adopted a smooth non-parametric estimation to explore the safety-first portfolio optimization problem. The portfolio effective frontier of safety-first portfolio model based on the nonparametric estimation is smoother than the safety-first portfolio model based on the empirical distribution estimation. Therefore, the safety-first portfolio model with nonparametric estimation is superior to the safety-first portfolio model.

The MAD portfolio model also depends on the sample data to estimate the parameters. In addition, solutions to optimization problems can show remarkable sensitivity to parameters’ estimation error or uncertainty. If the sample data is limited, the small change of the expected return value given by the investor will cause a large change in the optimal solution of the portfolio. More importantly, only a finite number of observations are available, so we need to use optimization and statistical methods to deal with this drawback. At present, the nonparametric estimation method has not been applied to the MAD portfolio model. Therefore,
in this paper, we will use the absolute deviation value to measure the risk, and apply the nonparametric estimation method in the MAD model to find the optimal investment strategy. In particular, we first apply the nonparametric kernel mean estimation and the nonparametric kernel median estimation to replace the returns of assets with five different kernel functions, respectively. Then, we construct the nonparametric kernel mean estimation-based MAD portfolio model and the nonparametric kernel median estimation-based MAD portfolio model. Finally, we will apply the nonparametric kernel mean estimation and the nonparametric kernel median estimation to mean-CVaR portfolio model.

The rest of this paper is organized as follows. In Section 2, we introduce MAD portfolio model. In Section 3, we introduce kernel mean return nonparametric estimation and apply it to MAD portfolio optimization. In Section 4, we introduce kernel median return nonparametric estimation and apply it to MAD portfolio optimization. In Section 5, we extend the two kinds of nonparametric approach to mean-CVaR portfolio model. In Section 6, we report some empirical studies to test the proposed models.

Notations: We use boldface letter such as \( \mathbf{x} \) for vector to distinguish it from scalar \( x \).

2. MAD portfolio model. Firstly, we give the notations for the MAD portfolio optimization model in Konno and Yamazaki [24], which are as followings:

- “\( N \)” stands for the total number of stocks.
- “\( x_j \)” stands for units of asset \( j \) to be included in the portfolio, \( j = 1, 2, \ldots, N \).
- “\( T \)” stands for the length of the time horizon.
- “\( t \)” stands for each period over the time horizon, \( t = 1, 2, \ldots, T \).
- “\( \eta \)” stands for minimum rate of return required by an investor.
- “\( R_j \)” stands for a random variable representing the rate of return of asset \( j \).
- “\( \mu_j \)” stands for the expected return \( E[R_j] \), of asset \( j \).
- “\( r_{jt} \)” stands for the observed return of asset \( j \) during the period \( t, t = 1, 2, \ldots, T \).
- “\( u_j \)” stands for the maximum units of asset \( j \).
- “\( d_t \)” stands for deviation below the average rate of return at time period \( t, t = 1, 2, \ldots, T \).
- “\( y_t \)” stands for the return of a portfolio \( \mathbf{x} \) in scenario \( t \).

We now suppose that \( T \) different scenarios have been identified as possible at the target time length \( T \). The probability that the scenarios \( t, t = 1, 2, \ldots, T \), will happen is indicated by \( p_t \), with \( \sum_{t=1}^{T} p_t = 1 \). In general, we assume \( p_t = \frac{1}{T} \). We further assume that for each random variable \( R_j, j = 1, 2, \ldots, N \), its realization \( r_{jt} \) under scenario \( t \) is known. The set of the returns of all the assets \( \mathbf{r}_t = \{ r_{jt}, j = 1, 2, \ldots, N \} \) define the scenario \( t, t = 1, 2, \ldots, T \). The expected return of asset \( j, j = 1, 2, \ldots, N \), is calculated as

\[
\mu_j = \sum_{t=1}^{T} p_t r_{jt} = \frac{1}{T} \sum_{t=1}^{T} r_{jt} \quad t = 1, 2, \ldots, T. \tag{1}
\]

For each portfolio \( \mathbf{x} = (x_1, x_2, \ldots, x_N) \), we define a corresponding random variable

\[
R(\mathbf{x}) = \sum_{j=1}^{N} R_j x_j,
\]

which represents the portfolio return. The return \( y_t \) of a portfolio \( \mathbf{x} \) in scenario \( t \) can be computed as
\[ y_t = \sum_{j=1}^{N} r_{jt} x_j. \]  

Then we have the expected return of the portfolio \( x \) can be computed as a linear function of \( x \)

\[ \mu(x) = E[R(x)] = \sum_{t=1}^{T} p_t y_t = \sum_{t=1}^{T} px \left( \sum_{j=1}^{N} r_{jt} x_j \right) = \sum_{j=1}^{N} x_j \left( \sum_{t=1}^{T} p_t r_{jt} \right) = \sum_{j=1}^{N} \mu_j x_j \]  

The variance is the classical statistical quantity used to measure the dispersion of a random variable around its mean with \( L_2 \) norm. There are, however, other ways to measure the dispersion of a random variable, for example, the MAD is a \( L_1 \) norm.

The MAD is a dispersion measure that is defined as

\[ \text{MAD}(x) = E[|R(x) - E[R(x)]|] = E[\left| \sum_{j=1}^{N} R_j x_j - E[\sum_{j=1}^{N} R_j x_j] \right|]. \]  

The MAD measures the average of the absolute value of the difference between the random variable and its expected value. With respect to the variance, the MAD considers absolute values instead of squared values.

Recalling that the expected return of the portfolio can be calculated as (4), the MAD can be written as

\[ \text{MAD}(x) = \sum_{t=1}^{T} p_t \left| \sum_{j=1}^{N} r_{jt} x_j - \sum_{j=1}^{N} \mu_j x_j \right| = \frac{1}{T} \sum_{t=1}^{T} |y_t - \sum_{j=1}^{N} \mu_j x_j|. \]  

Then, the MAD portfolio optimization problem with the decision variables \( x = (x_1, x_2, \ldots, x_N) \) is as following

\[ \min_{x \in \mathbb{R}^N} \text{MAD}(x), \]

\[ \text{s.t.} \quad \sum_{j=1}^{N} x_j \mu_j \geq \eta, \]

\[ \sum_{j=1}^{N} x_j = 1, \]

\[ 0 \leq x_j \leq u_j, j = 1, \ldots, N. \]  

The objective function of the model (6) is non-smooth i.e. it is not differentiable. We now define the deviation in scenario \( t \) as \( d_t \), that is \( d_t = |y_t - \sum_{j=1}^{N} \mu_j x_j| = |y_t - \mu(x)| \). As \( d_t = |y_t - \mu(x)| = \max\{ (y_t - \mu(x)); -(y_t - \mu(x)) \} \), then the MAD portfolio optimization problem with the decision variables \( x = (x_1, x_2, \ldots, x_N) \) and \( d = (d_1, d_2, \ldots, d_T) \) can be written in the following equivalent linear form

\[ \min_{(x,d) \in \mathbb{R}^N \times \mathbb{R}^T} \text{MAD}(x) = \frac{1}{T} \sum_{t=1}^{T} d_t, \]

\[ \text{s.t.} \quad d_t \geq y_t - \mu(x), \quad t = 1, 2, \ldots, T, \]

\[ d_t \geq -(y_t - \mu(x)), \quad t = 1, 2, \ldots, T, \]

\[ \sum_{j=1}^{N} x_j \mu_j \geq \eta. \]
\[
\sum_{j=1}^{N} x_j = 1, \\
y_t = \sum_{j=1}^{N} r_{jt} x_j, \\
d_t \geq 0, \quad t = 1, 2, \ldots, T, \\
0 \leq x_j \leq u_j, \quad j = 1, \ldots, N.
\]

where \(\mu_j\) and \(\mu(\mathbf{x})\) are defined by (1) and (3), respectively.

In the case the rates of return are a multivariate normally distributed random variable, the rate of return of the portfolio is normally distributed. Then, the proportionality relation between the mean absolute deviation and the standard deviation occurs \(MAD(\mathbf{x}) = \sqrt{\frac{2}{\pi}} \sigma(\mathbf{x})\). As a consequence, minimizing the MAD is equivalent to minimizing the standard deviation, which means, in this specific case, the equivalence of the associated optimization problems. However, the MAD model does not require any specification of the return distribution. In addition, the objective function and the constraints of the above model are linear, the MAD portfolio optimization model is a linear programming problem (LP), which can be solved efficiently using the state-of-the-art solvers.

3. Kernel mean return nonparametric estimation with application to MAD portfolio optimization. However, the classic MAD portfolio optimization model assumes that risk factors (random variables) are with a known probability distribution or have a particular parametric form of distribution. But in most cases, we know little about the density or probability functional form of risk factors. As shown in Li and Racine [25], nonparametric estimation does not rely on data belonging to any particular distribution. In particular, it can be applied in situations where little or no restrictive prior information on functionals is known. Therefore, in recent years, many scholars study the computational formulas for risks measured by VaR or CVaR by using the nonparametric estimation method. To the best of our knowledge, no scholar has applied the nonparametric estimation method to the MAD portfolio optimization model. Moreover, because only a finite number of observations are available, it is difficult for the investor to obtain the optimal solution of the portfolio. Therefore, we will use the nonparametric estimation methods to deal with this drawback. In what following, we introduce the kernel estimation method.

Following in Li and Racine [25], we introduce the kernel function \(K(\cdot)\) which is a probability density function. For theoretical perspective, the following properties are indispensable for the kernel function \(K(\cdot)\):

\[
\int_{-\infty}^{\infty} K(z) dz = 1, \quad \int_{-\infty}^{\infty} z K(z) dz = 0, \quad \int_{-\infty}^{\infty} z^2 K(z) dz < +\infty. \quad (8)
\]

Some classic kernel functions with domain of definition and the set in which the function takes values are as followings:

(1) Rectangular:

\[
K(z) = \begin{cases} 
\frac{1}{2}, & \text{if } |z| < 1, \\
0, & \text{if } |z| \geq 1.
\end{cases} \quad (9)
\]
(2) Triangular:

\[ K(z) = \begin{cases} 
1 - |z|, & \text{if } |z| < 1, \\
0, & \text{if } |z| \geq 1.
\end{cases} \tag{10} \]

(3) Epanechnikov:

\[ K(z) = \begin{cases} 
\frac{3}{4}(1 - z^2), & \text{if } |z| < 1, \\
0, & \text{if } |z| \geq 1.
\end{cases} \tag{11} \]

(4) Biweight:

\[ K(z) = \begin{cases} 
\frac{15}{16}(1 - z^2)^2, & \text{if } |z| < 1, \\
0, & \text{if } |z| \geq 1.
\end{cases} \tag{12} \]

(5) Gaussian:

\[ K(z) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}z^2\right), z \in \mathbb{R}. \tag{13} \]

3.1. Kernel mean return nonparametric estimation. Now we introduce some preliminary knowledge of kernel mean return nonparametric estimation with the kernel function \( K(\cdot) \) which is a probability density function (see Li and Racine [25] for more details). A kernel mean estimation of one point can be seen as a weighted average of these observations where the weight is given to each observation decreases with its distance from the estimated point.

If we have \( T \) return observations made at moments \( t = 1, 2, \ldots, T \) denote by \( r_{jt} \) the rate of return of asset \( j \) at moment \( t \).

Recall that for every \( j \in \{1, 2, \ldots, N\} \), \( R_j \) is the random variable that represents the rate of return of asset \( j \). The kernel estimator for the probability density function of the random variable \( R_j \) is the following

\[ \hat{f}_j(x) = \frac{1}{Th_j} \sum_{t=1}^{T} K\left(\frac{r_{jt} - x}{h_j}\right), \tag{14} \]

where the kernel function \( K(\cdot) \) is selected from (9)-(13), and \( h_j = h_j(T) \) is bandwidth. The kernel function \( K \) determines the shape of the weighting function. The bandwidth \( h_j = h_j(T) \) is a smoothing parameter which depends on the sample size \( T \). Based on the same way, we can obtain the kernel estimation of the probability density function \( f(r_{jt}), j = 1, 2, \ldots, N \) for all assets.

For the choice of bandwidth \( h_j = h_j(T) \), theoretical and practical researches had been carried out on the question that how one should select \( K \) and \( h \) in order to optimize the properties of the estimator. There are some ways to choose \( K \) and \( h \): minimization of asymptotic errors, rule of thumb, cross validation, plug-in methods. More details can be found in Pagan and Ullah [28], Silverman [37], Subramanian [39].

In this paper, we use the rule of thumb in Hansen [17] to select the bandwidth \( h \). That is, \( h_j = h_j(T) = 1.06\sigma_j T^{-0.2} \), where \( \sigma_j \) is the standard deviation of \( T \) return observations \( r_{jt}, t = 1, \ldots, T \), and the estimated value of \( \sigma_j \) is

\[ \sigma_j = \sqrt{\frac{1}{T - 2} \sum_{t=1}^{T} (r_{jt} - \overline{r}_j)^2}, \]

where \( \overline{r}_j = \frac{1}{T} \sum_{t=1}^{T} r_{jt} \). It is obviously that the bandwidth \( h_j = h_j(T) \) is a sequence of positive real numbers which decreases to 0 as \( T \) tends to infinity.
3.2. MAD model based on nonparametric kernel mean estimation. Given the length of the time horizon \( T \) with return observations \( \{r_{jt}, j = 1, \ldots, N, t = 1, \ldots, T\} \), from a given portfolio, to build a more sophisticated estimation of mean-absolute deviation model, as in Athayde([2], [3]), we will replace all the observations \( \{r_{jt}, j = 1, \ldots, N, t = 1, \ldots, T\} \) by their kernel mean estimation \( \hat{r}_{jt} \). The kernel mean estimations \( \{\hat{r}_{jt}, j = 1, \ldots, N, t = 1, 2, \ldots, T\} \) are a weighted average of the observations, in which the weight given to each observation decreases with its distance from the point. In particular, following Athayde([2], [3]), the kernel mean estimator \( \hat{r}_{jt} \) of some return \( r_{jt} \) for a given asset \( j \) is then given by

\[
\hat{r}_{jt} = \frac{\sum_{l=1}^{T} r_{jl} K \left( \frac{r_{jt} - r_{jl}}{h} \right)}{\sum_{l=1}^{T} K \left( \frac{r_{jt} - r_{jl}}{h} \right)}.
\]

(15)

The kernel mean estimator scenario return \( \hat{y}_t \) of a portfolio \( \hat{x} \) in scenario \( t \) has the following form

\[
\hat{y}_t = \sum_{j=1}^{n} \hat{r}_{jt} x_j.
\]

(17)

Then the kernel mean estimator expected return of the portfolio \( \hat{x} \) can be computed as a linear function of \( \hat{x} \)

\[
\hat{\mu} = E[R(\hat{x})] = \sum_{t=1}^{T} p_t \hat{y}_t = \sum_{t=1}^{T} p_t (\sum_{j=1}^{N} \hat{r}_{jt} x_j) = \sum_{j=1}^{N} x_j (\sum_{t=1}^{T} p_t \hat{r}_{jt}) = \sum_{j=1}^{N} \hat{\mu}_j x_j
\]

(18)

Therefore, the MAD portfolio optimization problem based on nonparametric kernel mean estimation can be written in the following equivalent linear form

\[
\min_{\hat{x} \in \mathbb{R}^n} \quad \text{MAD}(\hat{x}) = \frac{1}{T} \sum_{t=1}^{T} \hat{d}_t,
\]

s.t. \( \hat{d}_t \geq \hat{y}_t - \hat{\mu} \), \( t = 1, 2, \ldots, T \),

\( \hat{d}_t \geq -(\hat{y}_t - \hat{\mu}) \), \( t = 1, 2, \ldots, T \),

\[
\sum_{j=1}^{N} x_j \hat{\mu}_j \geq \eta,
\]

(19)

\[
\sum_{j=1}^{N} x_j = 1,
\]

\( \hat{d}_t \geq 0 \), \( t = 1, 2, \ldots, T \),

\( 0 \leq x_j \leq u_j \), \( j = 1, \ldots, N \).

where \( \hat{d}_t = |\hat{y}_t - \hat{\mu}| = \max\{ (\hat{y}_t - \hat{\mu}); -(\hat{y}_t - \hat{\mu}) \} \), \( \hat{\mu}_j \), \( \hat{y}_t \) and \( \hat{\mu} \) are defined by (16), (17) and (18) respectively. It is obvious that the MAD portfolio optimization model
based on nonparametric kernel mean return estimation is still a LP problem, which can be solved efficiently using the state-of-the-art solvers.

4. **Kernel median return estimation with application to MAD portfolio optimization.** The disadvantage of the above kernel mean estimation is that it is sensitive to outliers and may be inappropriate in some cases, as when the distribution is highly asymmetric (see, Conine and Tamarkin [11]). This problem can be solved by using another useful descriptive statistic which is robust to heavy-tailed error distributions and outliers: the kernel median return nonparametric estimation.

It is a more complete picture of the distribution than the one given by its mean. For some relevant theoretical properties and applications about kernel median nonparametric estimation, the readers can refer to Berlinet et al. ([4]; [5]); Gannoun et al.[15]; Gooijer and Gannoun[16]; Salah et al.[33]; Shen[36]; Subramanian ([39], [40]); Zhao and Ma[52], Zhao and Chen[53], Zhao and Cui[54], among others.

4.1. **Kernel median return estimation.** As we know, the median is more robust and most adapted to model data than the mean when the conditional distribution is asymmetric or multi-modal in regression and prediction context.

If \((X, Z)\) stands for a \(R^t\)-valued random vector and \(F(\cdot | X = x)\) is the conditional distribution of \(Z\) for given \(X = x\), then the conditional median value \(v(x)\) of \(Z\) for given \(X = x\) is the solution of the following minimization problem:

\[
v(x) = \arg \min_{y \in \mathbb{R}} \int |z - y| F(dz | X = x). \tag{20}
\]

In order to get an estimator \(\hat{v}(x)\) of \(v(x)\), we can replace \(F(dz | X = x)\) in (20) by an appropriate estimator \(\hat{F}(\cdot | X = x)\). Then we have

\[
\hat{v}(x) = \arg \min_{y \in \mathbb{R}} \int |z - y| \hat{F}(dz | X = x). \tag{21}
\]

More details on applied and theoretical properties of median (or on quantiles) can be found in Gannoun et al.[15].

Given the length of the time horizon \(T\) with \(r_t = \{r_{jt}, j = 1, 2, \ldots, N\}\) defined by the scenarios \(t, t = 1, 2, \ldots, T\) from a given portfolio, as in Gannoun et al.[15], the kernel estimator \(\hat{F}(z | r_{jt})\) of the conditional distribution function \(F(z | r_t)\) can be defined by

\[
\hat{F}(z | r_{jt}) = \frac{\sum_{l=1}^T 1 \{z \leq z_l\} K\left(\frac{z_l - r_{jt}}{h}\right)}{\sum_{l=1}^T K\left(\frac{r_{jt} - r_{lt}}{h}\right)}, \tag{22}
\]

From Koenker [23] and (22), we can obtain a kernel median return estimation for a given asset \(j\) directly by:

\[
\hat{r}_{jt} = \arg \min_{z \in \mathbb{R}} \frac{\sum_{l=1}^T (r_{jt} - z) K\left(\frac{r_{jt} - r_{lt}}{h}\right)}{\sum_{l=1}^T K\left(\frac{r_{jt} - r_{lt}}{h}\right)}, \tag{23}
\]

where \(K(\cdot)\) is a kernel function, \(h\) is called bandwidth.

Following in Gannoun et al.[15], we can also use another way to obtain a kernel median return estimation for \(\hat{r}_{jt}\) by solving the following equation

\[
\hat{F}(z | r_{jt}) = \frac{\sum_{l=1}^T 1 \{z \leq z_l\} K\left(\frac{z_l - r_{jt}}{h}\right)}{\sum_{l=1}^T K\left(\frac{r_{jt} - r_{lt}}{h}\right)} = 0.5 \tag{24}
\]

It is easy to see that the conditional distribution function can be seen as a conditional expectation i.e. \(E(1_{\{Z \leq z\}} | X = x) = F(z | X = x)\). In addition, if
$F(\cdot|X = x)$ is continuous and strictly increasing, the conditional median of $Z$ for given $X = x$ can be obtained by the following

$$v(x) = F^{-1}(0.5|X = x).$$

4.2. MAD model based on kernel median return estimation. Given the length of the time horizon $T$ with return observations $\{r_{jt}, j = 1, \ldots, N, t = 1, \ldots, T\}$, from a given portfolio, to build a more sophisticated estimation of mean-absolute deviation model, we can replace all the returns observations $\{r_{jt}, j = 1, \ldots, N, t = 1, \ldots, T\}$ by their kernel median estimators (median regression) $\{\hat{r}_{jt}, j = 1, \ldots, N, t = 1, \ldots, T\}$. From the definition of kernel median estimators (23), we can see that the kernel median estimators of the returns observations $\{r_{jt}, j = 1, \ldots, N, t = 1, \ldots, T\}$ are a weighted median of the observations. Hence, following Athayde ([2], [3]), the expected return of asset $j, j = 1, 2, \ldots, N$, is calculated as

$$\hat{\mu}_j = \frac{T}{T} \sum_{t=1}^{T} \hat{y}_t x_j = \frac{N}{N} \sum_{j=1}^{N} \hat{\mu}_j x_j$$

(27)

And the kernel median estimator scenario return $\hat{y}_t$ of a portfolio $x$ in scenario $t$ can be computed as

$$\hat{y}_t = \sum_{j=1}^{n} \hat{r}_{jt} x_j.$$ 

(26)

Then we have the kernel median estimator expected return of the portfolio $x$ with the following linear function of $x$

$$m(x) = E[R(x)] = \sum_{t=1}^{T} p_t \hat{y}_t = \sum_{t=1}^{T} \sum_{j=1}^{N} p_t \hat{r}_{jt} x_j = \sum_{j=1}^{N} x_j \sum_{t=1}^{T} p_t \hat{r}_{jt} = \sum_{j=1}^{N} \hat{\mu}_j x_j$$

(27)

Therefore, the MAD portfolio optimization problem based on nonparametric kernel median return estimation can be written in the following equivalent linear form

$$\min_{x \in \mathbb{R}^n}, \quad \text{MAD}(x) = \frac{1}{T} \sum_{t=1}^{T} \hat{d}_t,$$

s.t. $\hat{d}_t \geq \hat{y}_t - m(x), \quad t = 1, 2, \ldots, T,$

$$\hat{d}_t \geq -(\hat{y}_t - m(x)), \quad t = 1, 2, \ldots, T,$$

$$\sum_{j=1}^{N} x_j \hat{\mu}_j \geq \eta,$$

$$\sum_{j=1}^{N} x_j = 1,$$

$$\hat{d}_t \geq 0, \quad t = 1, 2, \ldots, T,$$

$$0 \leq x_j \leq u_j, \quad j = 1, \ldots, N.$$

(28)

where $\hat{d}_t = |\hat{y}_t - \hat{\mu}| = \max\{ (\hat{y}_t - \hat{\mu}); -(\hat{y}_t - \hat{\mu}) \}, \hat{\mu}_j, \hat{y}_t$ and $m(x)$ are defined by (25), (26) and (27) respectively. It is obvious that the MAD portfolio optimization model
based on nonparametric kernel median return estimation is still a LP problem, which can be solved efficiently using the state-of-the-art solvers.

5. **Extend to mean-CVaR portfolio model.** As we know, the CVaR also depends the historical sample data. In this section, we will extend the nonparametric kernel mean return estimation and nonparametric kernel median return estimation to mean-CVaR portfolio model. Following the definition of CVaR in Rockafellar and Uryasev [32], we briefly give the mean-CVaR portfolio model. The readers can refer to Rockafellar and Uryasev [32] for more details.

Suppose we have \( T \) return observations \( r_t = \{r_{jt}, j = 1, 2, \ldots, N\} \) defined by the scenarios \( t, t = 1, 2, \ldots, T \) from a given portfolio. We also assume that the probability that the scenarios \( t, t = 1, 2, \ldots, T \), will happen is indicated by \( p_t \). With \( \sum_{t=1}^{T} p_t = 1 \), we have \( p_t = \frac{1}{T} \). Further, the loss function is defined as

\[
 f(x, r_t) = -x \cdot r_t,
\]

for the scenarios \( t, t = 1, 2, \ldots, T \). And the expected return of the portfolio \( x \) can be computed as in (3).

For a given confidence level \( \beta \), the mean-CVaR portfolio model becomes the following LP problem with variables \((x, d, \alpha, \theta) \in \mathbb{R}^N \times \mathbb{R}^T \times \mathbb{R} \times \mathbb{R} \).

\[
\begin{align*}
\min \quad & \theta \\
\text{s.t.} \quad & \alpha + \frac{1}{T(1 - \beta)} \sum_{t=1}^{T} d_t \leq \theta, \\
& d_t \geq -x \cdot r_t - \alpha, \quad t = 1, \ldots, T, \\
& \sum_{j=1}^{N} \mu_j x_j \geq \eta, \\
& \sum_{j=1}^{N} x_j = 1, \\
& d \geq 0, \quad x \geq 0.
\end{align*}
\]

5.1. **Mean-CVaR model based on kernel mean return estimation.** As in 3.2, we can also replace all the returns observations \( \{r_{jt}, j = 1, \ldots, N, t = 1, \ldots, T\} \) by their kernel mean estimators \( \{\hat{r}_{jt}, j = 1, \ldots, N, t = 1, \ldots, T\} \) to obtain the mean-CVaR model based on kernel mean return estimation.

Using the the kernel mean estimator scenario return \( \{\hat{r}_{jt}, j = 1, 2, \ldots, N, t = 1, \ldots, T\} \) in (15) and the kernel mean estimator expected return of the portfolio \( x \) in (18), the mean-CVaR model based on kernel mean return estimation portfolio model becomes the following LP problem with variables \((x, d, \alpha, \theta) \in \mathbb{R}^N \times \mathbb{R}^T \times \mathbb{R} \times \mathbb{R} \).

\[
\begin{align*}
\min \quad & \theta \\
\text{s.t.} \quad & \alpha + \frac{1}{T(1 - \beta)} \sum_{t=1}^{T} d_t \leq \theta, \\
& d_t \geq -x \cdot \hat{r}_t - \alpha, \quad t = 1, \ldots, T, \\
& \sum_{j=1}^{N} \hat{\mu}_j x_j \geq \eta,
\end{align*}
\]
\[ \sum_{j=1}^{N} x_j = 1, \]
\[ d \geq 0, \quad x \geq 0. \]

5.2. Mean-CVaR model based on kernel median return estimation. Similarly, we can replace all the returns observations \( \{r_{jt}, j = 1, \ldots, N, t = 1, \ldots, T\} \) by their kernel median estimators \( \{\hat{r}_{jt}, j = 1, \ldots, N, t = 1, \ldots, T\} \) in (23). Using the kernel median estimator expected return of the portfolio \( x \) in (27), the mean-CVaR portfolio model based on kernel median return estimation becomes the following LP problem with variables \((x, d, \alpha, \theta) \in \mathbb{R}^N \times \mathbb{R}^T \times \mathbb{R} \times \mathbb{R}\).

\[
\begin{align*}
\min & \quad \theta \\
\text{s.t.} & \quad \alpha + \frac{1}{T(1-\beta)} \sum_{t=1}^{T} d_t \leq \theta, \\
& \quad d_t \geq -x \cdot \hat{r}_t - \alpha, \quad t = 1, \ldots, T, \\
& \quad \sum_{j=1}^{N} \hat{\mu}_j x_j \geq \eta, \\
& \quad \sum_{j=1}^{N} x_j = 1, \\
& \quad d \geq 0, \quad x \geq 0.
\end{align*}
\]

6. Empirical studies. In this section, we apply the models described above to construct optimal portfolios and evaluate their in-sample and out-of-sample performance. In order to investigate the performance of the proposed methods, we employ the real market data of A shares listed on the Shanghai stock exchange. In the empirical analysis, we do not take transaction costs and taxes into account.

6.1. Data and models. In our empirical studies, the portfolio models have the following meanings:

- “MAD-1” stands for MAD portfolio model (7).
- “MAD-2” stands for MAD portfolio model based on nonparametric kernel mean estimation model (19).
- “MAD-3” stands for MAD portfolio model based on nonparametric kernel median return estimation (28).
- “Mean-CVaR-1” stands for mean-CVaR portfolio model (30) with confidence level \( \beta = 0.95 \).
- “Mean-CVaR-2” stands for mean-CVaR portfolio model based on nonparametric kernel mean estimation model (31) with confidence level \( \beta = 0.95 \).
- “Mean-CVaR-3” stands for mean-CVaR portfolio model based on nonparametric kernel median return estimation (32) with confidence level \( \beta = 0.95 \).

In our empirical studies, all off the tested portfolio models are LP problem, which can be solved efficiently using the state-of-the-art solvers, for example, optimization toolbox (linprog) in matlab. Hence, all empirical studies are computed by using Matlab 2016 on personal computer.

In our empirical studies, the data comes from great wisdom 365 which was set up in Shanghai, China, in 2000. At every trading day, it can provide high-speed
In particular, we choose 20 stocks of A shares listed on the Shanghai stock exchange, and use the closing price of the 20 stocks from November 1, 2013 to October 31, 2016 (a total of 732 data) in empirical analysis. The total data is divided into two parts. The first part is from November 1, 2013 to October 31, 2015 (a total of 488 data) which is used to calculate the optimal solution of the MAD model. The second part is from November 1, 2015 to October 31, 2016 (a total of 244 data) which is used to measure the out-of-sample performance of the portfolio. In order to test the efficiency of the method, we also use the Shanghai Composite Index as the benchmark. Therefore, we also select the daily closing index data of the Shanghai Composite Index from November 1, 2015 to October 31, 2016 (a total of 244 data).

Table 1. Descriptive Statistics of daily return of 20 stocks from Shanghai A shares

| Stock name | Number | Max  | Min  | Mean  | Sd    | Skewness | Kurtosis |
|------------|--------|------|------|-------|-------|----------|----------|
| P.R.E.     | 600048 | 0.10038 | -0.35544 | 0.00049 | 0.03308 | -1.47313 | 22.2281  |
| T.J.H.     | 600717 | 0.10025 | -0.10029 | 0.00074 | 0.03040 | -0.08191 | 5.4232   |
| H.E.       | 600060 | 0.10035 | -0.11024 | 0.00109 | 0.03355 | -0.00870 | 5.15926  |
| H.N.I.     | 600011 | 0.10050 | -0.10036 | 0.00064 | 0.02654 | -0.28255 | 7.6319   |
| N.J.H.T.   | 600064 | 0.10028 | -0.32864 | 0.00112 | 0.03146 | -1.69324 | 20.8171  |
| S.H.I.     | 600031 | 0.10066 | -0.10042 | 0.00015 | 0.02843 | -0.14598 | 6.8630   |
| S.H.A.     | 600009 | 0.09996 | -0.10007 | 0.00112 | 0.02464 | -0.09756 | 7.3309   |
| T.R.T.     | 600885 | 0.10022 | -0.10016 | 0.00109 | 0.02834 | -0.04908 | 6.7348   |
| Ch.M.B.    | 600036 | 0.09571 | -0.09914 | 0.00087 | 0.01999 | 0.44853  | 8.3975   |
| Ch.S.      | 600150 | 0.10016 | -0.10011 | 0.00101 | 0.03593 | -0.04488 | 4.6552   |
| Ch.U.      | 600050 | 0.10103 | -0.10056 | 0.00096 | 0.02846 | 0.21470  | 6.7539   |
| Sino.      | 600028 | 0.10035 | -0.10040 | 0.00037 | 0.02138 | -0.15015 | 8.9252   |
| Ch.S.      | 600118 | 0.10027 | -0.10009 | 0.00170 | 0.03683 | -0.03555 | 4.7026   |
| C.S.       | 600030 | 0.10043 | -0.10012 | 0.00091 | 0.03038 | 0.23151  | 5.9877   |
| C.S.M.     | 601098 | 0.10027 | -0.10017 | 0.00120 | 0.02910 | -0.14500 | 5.4879   |
| Ch.L.I.    | 601628 | 0.10036 | -0.10007 | 0.00095 | 0.02689 | 0.62490  | 6.7596   |
| O.F.X.     | 600612 | 0.10010 | -0.10006 | 0.00109 | 0.02811 | 0.29691  | 5.8163   |
| C.Q.B.     | 600132 | 0.10027 | -0.10699 | 0.00042 | 0.02841 | -0.37571 | 6.9771   |
| J.J.I.     | 600650 | 0.10037 | -0.10018 | 0.00202 | 0.03929 | 0.20423  | 4.3077   |
| Q.J.B.     | 600706 | 0.10028 | -0.10024 | 0.00125 | 0.03459 | -0.26875 | 4.7732   |

Table 1 lists the maximum and minimum, standard deviation, skewness, kurtosis of the daily return of 20 stocks selected from the Shanghai stock exchange from November 1, 2013 to October 31, 2016. For example, the daily return of Sinopec’s maximum return is 0.10035, the minimum return is -0.10040, the return mean is 0.00037, the standard deviation is 0.02138, the skewness is -0.15015, the kurtosis is 8.92527.

Therefore, the daily stock return of Shanghai stock exchange market is not subject to normal distribution, and there are high peak and heavier-tail. The non-parametric estimation methods mentioned in this paper does not need to know the specific form of the distribution of returns, and can overcome the impact of high peak and heavier-tail phenomenon.
6.2. **Comparison of efficient frontier.** In this subsection, we use the selected 20 stocks of A shares listed on the Shanghai stock exchange from November 1, 2013 to October 31, 2015 (total of 488 data) to calculate the optimal solution of the MAD-1, 2 and 3 portfolio models, and the Mean-CVaR-1, 2 and 3 portfolio models. That is, we can obtain the risk measure and optimal investment weight \( x_i, i = 1, 2, \ldots, N \). Then we give efficient frontier of the MAD-1, 2, 3 portfolio models, and the Mean-CVaR-1, 2, 3 portfolio models as the following figures. The kernel function used in portfolio models MAD-2 and 3, and the portfolio models Mean-CVaR-2 and 3 are the Gauss kernel function (13).

![Figure 1. Comparision of efficient frontier of MAD portfolio models](image1)

![Figure 2. Efficient frontier of the MAD model based on the kernel median estimation](image2)

![Figure 3. Efficient frontier of the mean-CVaR portfolio models](image3)
Fig. 1 shows the comparison of efficient frontier of the MAD-1, 2 and 3 portfolio models, and Fig. 2 is the efficient frontier of the MAD model based on the kernel median estimation. Fig. 3 shows the comparison of efficient frontier of the Mean-CVaR-1, 2 and 3 portfolio models.

It can be observed from Fig. 1 that the efficient frontiers of the MAD-2 and 3 portfolio models are in the upper left of the MAD-1 portfolio model when the data selected from Shanghai stock exchange market. For the same expected return of a portfolio, the minimum absolute deviation of the MAD-2 and 3 portfolio models are smaller than that of the MAD-1 portfolio model, where the minimum absolute deviation of the MAD-3 portfolio model is the smallest. Therefore, the risk generated by using MAD model for nonparametric methods is smaller than the MAD model without using nonparametric method, and the risk produced by using the MAD-3 portfolio model (the MAD model of kernel median estimation) is smaller than that by using the MAD-2 portfolio model. The efficient frontiers of the MAD-3 portfolio model in Fig. 1 is a straight line perpendicular to the abscissa, which is due to in the range of the expected yield level \( \eta < 0.017 \), the MAD-3 portfolio model is insensitive to \( \eta \) the value of this range, that is, when \( \eta < 0.017 \), the minimum absolute deviation of the MAD-3 portfolio model does not change with the change of \( \eta \). Fig. 2 shows the range of the MAD-3 portfolio model which is sensitive to the expected yield level \( \eta \). It can be observed from Fig. 2 that the minimum absolute deviation of the MAD-3 portfolio model will increase with the increase of when \( \eta > 0.017 \).

It can also be seen from Fig. 3 that the efficient frontiers of the Mean-CVaR-2 and 3 portfolio models are in the upper left of the Mean-CVaR-1 portfolio model. For the same expected return of a portfolio, the CVaR of the Mean-CVaR-2 and 3 portfolio models are smaller than that of the Mean-CVaR-1 portfolio model, where the CVaR of the Mean-CVaR-3 portfolio model is the smallest. Therefore, the risk generated by using CVaR model for nonparametric methods is smaller than the CVaR model without using nonparametric method, and the risk produced by using the Mean-CVaR-3 portfolio model (the mean-CVaR model of kernel median estimation) is smaller than that by using the Mean-CVaR-2 portfolio model. From the above figures, we can see that both the nonparametric estimation-based MAD portfolio models and the nonparametric estimation-based mean-CVaR portfolio models enjoy better in-sample performance.

6.3. Comparison of efficient frontier based on different kernel functions.
In the above subsection, we use the Gaussian kernel function to compare the efficient frontier of portfolio models, while in this subsection we try to analyse whether the kernel function has an influence on the optimal solution of the MAD portfolio model and the mean-CVaR portfolio model.

It can be seen from Fig. 4 and Fig. 5 that the efficient frontiers of the MAD portfolio models and the mean-CVaR portfolio models of kernel mean estimation obtained from the four kernel function: rectangular, triangular, biweight, epanechnikov are similar, and the Gaussian is the best. So it is reasonable to consider the Gaussian as a kernel function in the above subsection.

It can be seen from Fig. 6 that the efficient frontiers of the MAD model of kernel median estimation obtained from the five kernel function are roughly same. If \( \eta < 0.024 \), the effect of the Gaussian is better than that of the rectangular, the triangular, the biweight and the epanechnikov. However, if \( \eta > 0.024 \), the effect of the rectangular, the triangular, the biweight and the epanechnikov are better than that of the Gaussian. In general, there is no much difference. The efficient frontier
of CVaR models of kernel median estimation in Fig.7 shows that the different kernel functions have small effect. Of course, the effect of the Gaussian is better than that of the rectangular, the triangular, the biweight and the epanechnikov which is not related to the value of $\eta$.

6.4. The validity of the portfolio models based on nonparametric methods. In this subsection, we choose the daily return of the Shanghai Composite Index from November 1, 2015 to October 31, 2016 as a benchmark, then compare the forecast portfolio return of the MAD-1, 2, 3 portfolio models and the Mean-CVaR-1, 2 and 3 portfolio models with the benchmark respectively. Descriptive
Figure 7. Efficient frontier of mean-CVaR models of kernel median estimation under different kernel functions.

statistics of daily return of Shanghai Composite Index from November 1, 2015 to October 31, 2016 are listed in Table 2.

Table 2. Descriptive statistics of daily return of Shanghai Composite Index

|          | Max    | Min    | average | sd      | Skewness | Kurtosis |
|----------|--------|--------|---------|---------|----------|----------|
| Shanghai S-I-R | 0.04310| -0.07045| -0.00016| 0.01566 | -1.37612| 8.38382 |

In this case, we use the rolling sample prediction method to obtain the portfolio return. Each time we roll five data (ie, the daily return data for one week), if it is the MAD model of kernel mean estimation, the specific steps are as follows (the steps of the other portfolio models are similar): First, give a desired return, there is a total of 487 daily yield data from November 1, 2013 to October 31,2015, which can be used to calculate the optimal investment weight of model 2, then using the optimal investment weight to forecast the first five days’ (2015/11/2-2015/11/6) portfolio return from November 1, 2013 to October 31,2015. And the first five daily return data can be removed, then adding the top five data from November 1, 2015 to October 31,2016 in so that to get new 487 daily return data. The optimal investment weight of model 2 can be calculated with the new 487 daily return data, then using this optimal investment weight to forecast the 6th portfolio yield to 10th portfolio return from November 1, 2015 to October 31,2016.

And so on, we forecast 5-day’s portfolio return each time, then excluding the first five daily return and add the new 5-day return in the tested dataset. So the tested dataset is always 487, the number of scrolling are 49 times, and eventually generating 243 predicted portfolio return values.

The figures 8-10 show the comparison out-of-sample portfolio return of the MAD-1, 2, 3 portfolio models with the daily return of Shanghai Composite Index where the expected return \( \eta = 0.002 \). The figures 11-13 give the comparison out-of-sample portfolio return of the Mean-CVaR-1, 2, 3 portfolio models with the daily return of Shanghai Composite Index where the expected return \( \eta = 0.002 \).

Among them, 48% of forecast portfolio return of the MAD-1 portfolio model is higher than the daily return of Shanghai Composite Index, 52% of the MAD-2 portfolio model’s forecast portfolio return is higher than the daily return of Shanghai Composite Index, 58% of the MAD-3 portfolio model’s forecast return is higher than the daily return of Shanghai Composite Index.
We can also see that, 47% of forecast portfolio return of the Mean-CVaR-1 portfolio model is higher than the daily return of Shanghai Composite Index, 54% of the Mean-CVar-2 portfolio model’s forecast portfolio return is higher than the daily return of Shanghai Composite Index, 57% of the Mean-CVaR-3 portfolio model’s forecast return is higher than the daily return of Shanghai Composite Index.
Therefore, the MAD model based on nonparametric estimation and the mean-CVaR model based on nonparametric estimation are more efficient than that without nonparametric estimation. And the MAD model of kernel median estimation is the most efficient among the MAD-1, 2, 3 portfolio models. Similarly, the mean-CVaR model of kernel median estimation is also the most efficient among the Mean-CVaR-1, 2, 3 portfolio models. The reason is that the kernel median return nonparametric estimation is better suited for skewed distributions to derive at central tendency since it is much more robust and sensible. As statistics test shows
the daily stock return of Shanghai stock exchange market is not subject to normal distribution, and there are high peak and heavier-tail.

7. Conclusion. In this paper, we have studied the MAD portfolio selection problem by using the nonparametric estimation method. We firstly use the nonparametric kernel mean estimation and the nonparametric kernel median estimation to replace the returns of assets, respectively, with five different kernel functions. Then, the nonparametric kernel mean estimation-based MAD portfolio model and the nonparametric kernel median estimation-based MAD portfolio model are proposed. Further, we also apply the nonparametric kernel mean estimation and the nonparametric kernel median estimation to the mean-CVaR portfolio model. The proposed nonparametric estimation-based portfolio models all are LP problem, which can be solved efficiently using the state-of-the-art solvers.

The in-sample and out-of-sample analysis of the proposed strategies show that the nonparametric estimation-based portfolio models are more efficient than the classic portfolio model. In addition, the portfolio model based on nonparametric kernel median estimation is more efficient than the portfolio model based on nonparametric kernel mean estimation.

Although we don’t apply the nonparametric kernel mean estimation and the nonparametric kernel median estimation to mean-VaR portfolio model, we can except that the nonparametric estimation-based mean-VaR portfolio models also have similar performance as the the nonparametric estimation-based MAD portfolio models and the nonparametric estimation-based mean-CVaR portfolio models.

The nonparametric estimation methods introduced in this paper have generality and adaptability. For example, it can be used to investigate the multi-stage mean-risk portfolio optimization problem, where the risk is measured by absolute deviation, VaR, CVaR, and so on. On the other hand, we will try to apply nonparametric estimation methods to the spars and stable portfolio (see Dai and Wen [12], [14], and [51]) and empirical finance (see Wen et al.[43]-[46]). These will be our future research topics.

Acknowledgments. We are grateful to the anonymous referees and editor for their useful comments, which have made the paper clearer and more comprehensive than the earlier version. This work was supported by the NSF of China granted (71771030, 11301041, 71873146, 71671018), Scientific Research Fund of Hunan Provincial Education Department(16B005).

REFERENCES

[1] R. Alemany, C. Bolancé and M. Guillén, A nonparametric approach to calculating value-at-risk, *Insurance Math. Econ.*, 52 (2013), 255–262.
[2] G. Athayde, *Building a Mean-Downside Risk Portfolio Frontier*, Developments in Forecast Combination and Portfolio Choice, John Wiley and Sons, 2001.
[3] G. Athayde, *The Mean-downside Risk Portfolio Frontier: A Non-parametric Approach*, Published in Advances in portfolio construction and implementation, 2003.
[4] A. Berlinet, B. Cadre and A. Gannoun, On the conditional $L_1$-median and its estimation, *J. Nonparametr. Stat.*, 13 (2001a), 631–645.
[5] A. Berlinet, A. Gannoun and E. Matzner, Asymptotic normality of convergent estimates of conditional quantiles, *Stats.*, 35 (2001b), 139–169.
[6] N. Bingham, R. Kiesel and R. Schmidt, A semi-parametric approach to risk management, *Quant. Financ.*, 3 (2003), 426–441.
[7] Z. Cai and X. Wang, Nonparametric estimation of conditional VaR and expected shortfall, *J. Econom.*, 147 (2008), 120–130.
[8] S. Chen and C. Y. Tang, Nonparametric inference of value-at-risk for dependent financial returns, *J. Financ. Econ.*, 3 (2005), 227–255.

[9] S. Chen, Nonparametric estimation of expected shortfall, *J. Financ. Econ.*, 6 (2008), 87–107.

[10] L. Chiodi, R. Mansini and M. G. Speranza, Semi-absolute deviation rule for mutual funds portfolio selection, *Ann. Oper. Res.*, 124 (2003), 245–265.

[11] T. E. Conine, and M. J. Tamarkin, On diversification given asymmetry in returns, *J. Financ.*, 36 (1981), 1143–1155.

[12] Z. Dai and F. Wen, Some improved sparse and stable portfolio optimization problems, *Finan. Res. Lett.*, 27 (2018), 46–52.

[13] Z. Dai and F. Wen, A generalized approach to sparse and stable portfolio optimization problem, *J. Ind. Manag. Optim.*, 14 (2018), 1651–1666.

[14] Z. Dai and F. Wang, Sparse and robust mean-variance portfolio optimization problems. *Physica A*, 2019, accepted.

[15] A. Gannoun, J. Saracco and K. Yu, Nonparametric time series prediction by conditional median and quantiles, *J. Stat. Plan. Inferenc.*, 117 (2003), 207–223.

[16] J. G. Gooijer and A. Gannoun, Tr multivariate conditional median estimation, *Commun. Stat. Simul. C.*, 36 (2007), 165–176.

[17] B. E. Hansen, Bandwidth selection for nonparametric distribution estimation. Working Paper, 2004.

[18] Z. He, L. He and F. Wen, Risk compensation and market returns: The role of investor sentiment in the stock market, *Emerg. Mark. Financ. Tr.*, 55 (2019), 704–718.

[19] C. Huang, Z. Yang, T. Yi and X. Zou, On the basins of attraction for a class of delay differential equations with non-monotone bistable nonlinearities, *J. Differential Equations*, 256 (2014), 2101–2114.

[20] S. O. Jeong and K. H. Kang, Nonparametric estimation of value-at-risk, *J. Appl. Stat.*, 36 (2009), 1225–1238.

[21] P. Jorion, *Value at Risk: The New Benchmark for controlling Derivatives Risk*, New York: McGraw-Hill. 1997.

[22] H. Kellerer, R. Mansini and M. G. Speranza, Selecting portfolios with fixed costs and minimum transaction lots, *Ann. Oper. Res.*, 99 (2000), 287–304.

[23] R. Koenker, *Quantile Regression*, Cambridge Books, Cambridge University Press, Issue. 38, 2005.

[24] H. Konno and H. Yamazaki, Mean-absolute deviation portfolio optimization model and its applications to Tokyo stock market, *Manage. Sci.*, 37 (1991), 501–623.

[25] Q. Li and J. S. Racine, *Nonparametric Econometrics: Theory and Practice*, Princeton University Press, 2007.

[26] R. Mansini, W. Ogryczak and M. G. Speranza, LP solvable models for portfolio optimization: A classification and computational comparison, *IMA J. Manag. Math.*, 14 (2003), 187–220.

[27] H. Markowitz, Portfolio Selection, *J. Financ.*, 7 (1952), 77–91.

[28] A. R. Pagan and A. Ullah, *Nonparametric Econometrics*, Cambridge University Press. 1999.

[29] C. Papahristodoulou and E. Dotzauer, Optimal portfolios using linear programming models, *J. Oper. Res. Soc.*, 55 (2004), 1169–1177.

[30] J. S. Pang, A new efficient algorithm for a class of portfolio selection problems, *Oper. Res.*, 28 (1980), 754–767.

[31] A. Perold, Large scale portfolio selections, *Manage. Sci.*, 30 (1984), 1143–1160.

[32] T. Rockfeller and S. Uryasev, Optimization of conditional value-at-risk, *J. Risk.*, 2 (2000), 21–34.

[33] H. B. Salah, M. Chaouch, A. Gannoun, C. D. Peretti and A. Trabelsi, Mean and median-based nonparametric estimation of returns in mean-downside risk portfolio frontier, *Ann. Oper. Res.*, 262 (2018), 653–681.

[34] O. Scaillet, Nonparametric estimation and sensitivity analysis of expected shortfall, *Math. Financ.*, 14 (2004), 115–129.

[35] W. F. Sharpe, Capital asset prices: A theory of market equilibrium under conditions of risk, *J. Financ.*, 19 (1964), 425–442.

[36] P. Shen, Median regression model with left truncated and right censored data, *J. Stat. Plan. Infer.*, 142 (2012), 1757–1766.

[37] B. W. Silverman, *Density Estimation for Statistics and Data Analysis*, Chapman and Hall, New York. 1986.
[38] J. Schaumburg, Predicting extreme value at risk: Nonparametric quantile regression with refinements from extreme value theory, *Comput. Stat. Data Anal.*, 56 (2012), 4081–4096.

[39] S. Subramanian, Median regression using nonparametric kernel estimation, *J. Nonparametr. Stat.*, 14 (2002), 583–605.

[40] S. Subramanian, Median regression analysis from data with left and right censored observations, *Stat. Methodol.*, 4 (2007), 121–131.

[41] M. G. Speranza, Linear programming models for portfolio optimization, *Finance.*, 14 (1993), 107–123.

[42] M. G. Speranza, A heuristic algorithm for a portfolio optimization model applied to the Milan stock market, *Comput. Oper. Res.*, 23 (1996), 433–441.

[43] F. Wen, J. Xiao, C. Huang and X. Xia, Interaction between oil and US dollar exchange rate: Nonlinear causality, time-varying influence and structural breaks in volatility, *Appl. Econ.*, 50 (2018), 319–334.

[44] F. Wen, F. Min and Y. J. Zhang et al., Crude oil price shocks, monetary policy, and China’s economy, *Int. J. Financ. Econ.*, (2018) online.

[45] F. Wen, X. Yang and W. Zhou, Tail dependence networks of global stock markets, *Int. J. Financ. Econ.*, 24 (2019), 558–567.

[46] F. Wen, J. Xiao and X. Xia et al., Oil prices and chinese stock market: Nonlinear causality and volatility persistence, *Emerg. Mark. Financ. Tr.*, 55 (2019), 1247–1263.

[47] J. Xiao, M. Zhou and F. Wen et al., Asymmetric impacts of oil price uncertainty on Chinese stock returns under different market conditions: Evidence from oil volatility index, *Emerg. Econ.*, 74 (2018), 777–786.

[48] H. Yao, Z. Li and Y. Lai, Mean-CVaR portfolio selection: A nonparametric estimation framework, *Comput. Oper. Res.*, 40 (2013), 1014–1022.

[49] H. Yao, Y. Li and K. Benson, A smooth non-parametric estimation framework for safety-first portfolio optimization, *Quant. Financ.*, 15 (2015), 1865–1884.

[50] K. Yu, A. Allay, S. Yang and D. J. Hand, Kernel quantile based estimation of expected shortfall, *J. Risk.*, 12 (2010), 15-32.

[51] G. Yuan, Z. H. Meng and Y. Li, A modified Hestenes and Stiefel conjugate gradient algorithm for large-scale nonsmooth minimizations and nonlinear equations, *J. Optim. Theory. Appl.*, 168 (2016), 129–152.

[52] G. Zhao and Y. Y. Ma, Robust nonparametric kernel regression estimator, *Stat. Probabil. Lett.*, 116 (2016), 72–79.

[53] Y. Zhao and F. Chen, Empirical likelihood inference for censored median regression model via nonparametric kernel estimation, *J. Multivariate Anal.*, 99 (2008), 215–231.

[54] Y. Zhao and H. Cui, Empirical likelihood for median regression model with designed censoring variables, *J. Multivariate Anal.*, 101 (2010), 240–251.

Received June 2018; 1st revision October 2018; final revision December 2018.

E-mail address: zhifengdai823@163.com
E-mail address: sxjmgroup@163.com
E-mail address: wfh@amss.ac.cn