Research Article

Statistically Efficient Construction of α-Risk-Minimizing Portfolio

Hiroyuki Taniai[1] and Takayuki Shiohama[2]

[1] *School of International Liberal Studies, Waseda University, 1-6-1 Nishi-Waseda, Shinjuku, Tokyo 169-8050, Japan*

[2] *Department of Management Science, Faculty of Engineering, Tokyo University of Science, 1-3 Kagurazaka, Shinjuku, Tokyo 162-8601, Japan*

Correspondence should be addressed to Hiroyuki Taniai, htaniai@aoni.waseda.jp

We propose a semiparametrically efficient estimator for α-risk-minimizing portfolio weights. Based on the work of Bassett et al. (2004), an α-risk-minimizing portfolio optimization is formulated as a linear quantile regression problem. The quantile regression method uses a pseudolikelihood based on an asymmetric Laplace reference density, and asymptotic properties such as consistency and asymptotic normality are obtained. We apply the results of Hallin et al. (2008) to the problem of constructing α-risk-minimizing portfolios using residual signs and ranks and a general reference density. Monte Carlo simulations assess the performance of the proposed method. Empirical applications are also investigated.

1. Introduction

Since the first formation of Markowitz's mean-variance model, portfolio optimization and construction have been a critical part of asset and fund management. At the same time, portfolio risk assessment has become an essential tool in risk management. Yet there are well-known shortcomings of variance as a risk measure for the purposes of portfolio optimization; namely, variance is a good risk measure only for elliptical and symmetric return distributions.

The proper mathematical characterization of risk is of central importance in finance. The choice of an adequate risk measure is a complex task that, in principle, involves deep consideration of the attitudes of market players and the structure of markets. Recently, value at risk (VaR) has gained widespread use, in practice as well as in regulation. VaR has been criticized, however, because as a quantile is no reason to be convex, and indeed, it is easy to construct portfolios for which VaR seriously violates convexity. The shortcomings of VaR led to the introduction of coherent risk measures. Artzner et al. [1] and Föllmer and Schied [2]

question whether VaR qualifies as such a measure, and both find that VaR is not an adequate measure of risk. Unlike VaR, expected shortfall (or tail VaR), which is defined as the expected portfolio tail return, has been shown to have all necessary characteristics of a coherent risk measure. In this paper, we use α-risk as a risk measure that satisfies the conditions of coherent risk measure (see [3]). Variants of the α-risk measure include expected shortfall and tail VaR. The α-risk-minimizing portfolio, introduced as a *pessimistic portfolio* in Bassett et al. [3], can be formulated as a problem of linear quantile regression.

Since the seminal work by Koenker and Bassett [4], quantile regression (QR) has become more widely used to describe the conditional distribution of a random variable given a set of covariates. One common finding in the extant literature is that the quantile regression estimator has good asymptotic properties under various data dependence structures, and for a wide variety of conditional quantile models and data structures. A comprehensive guide to quantile regression is provided by Koenker [5].

Quantile regression methods use a pseudolikelihood based on an asymmetric Laplace reference density (see [6]). Komunjer [7] introduced a class of "tick-exponential" distribution, which includes an asymmetric Laplace density as a particular case, and showed that the tick-exponential QMLE reduces to the standard quantile regression estimator of Koenker and Bassett [4].

In quantile regression, one must know the conditional error density at zero, and incorrect specification of the conditional error density leads to inefficient estimators. Yet correct specification is difficult, because reliable shape information may be scarce. Zhao [8], Whang [9], and Komunjer and Vuong [10] propose efficiency corrections for the univariate quantile regression model.

This paper describes a semiparametrically efficient estimation of an α-risk-minimizing portfolio in place of an asymmetric Laplace reference density (a standard quantile regression estimator), by using any other α-quantile zero reference density f, based on residual ranks and signs. A $\sqrt{n}$-consistent and asymptotically normal one-step estimator is proposed. Like all semiparametric estimators in the literature, our method relies on the availability of a $\sqrt{n}$-consistent first-round estimator, a natural choice being the standard quantile regression estimator. Under correct specifications, they attain the semiparametric efficiency bound associated with f.

The remainder of this paper is organized as follows. In Section 2, we introduce the setup and definition of an α-risk-minimizing portfolio and present its equivalent formation under quantile regression settings. Section 3 contains theoretical results for our one-step estimator, and Section 4 describes its computation and performance. Section 5 gives empirical applications, and Section 6 our conclusions.

2. α-Risk-Minimizing Portfolio Formulation

"α-risk" can be considered a coherent measure of risk as discussed in Artzner et al. [1]. The α-risk of X, say $\rho_{v_\alpha}(X)$, is defined as

$$\rho_{v_\alpha}(X) := -\int_0^1 F_X^\leftarrow(t)dv_\alpha(t) = -\frac{1}{\alpha}\int_0^\alpha F_X^\leftarrow(t)dt, \quad \alpha \in (0,1), \tag{2.1}$$

where $v_\alpha(t) := \min\{t/\alpha, 1\}$ and $F_X^\leftarrow(\alpha) := \inf\{x : F_X(x) \geq \alpha\}$ denote the quantile function of a random variable X with distribution function F_X. Here, we recall the definition of expected

shortfall and the relationship among the tail risk measures in finance. The α-expected shortfall defined for $\alpha \in (0, 1)$ as

$$\mathrm{ES}_{(\alpha)}(X) = -\frac{1}{\alpha} \int_0^\alpha F_X^\leftarrow(p)\, dp \tag{2.2}$$

can be shown to be a risk measure that satisfies the axioms of a coherent measure of risk. It is worth mentioning that the expected shortfall is closely related but not coincident to the notion of conditional value at risk $\mathrm{CVaR}_{(\alpha)}$ defined in Uryasev [11] and Pflug [12]. We note that expected shortfall and conditional VaR or tail conditional expectations are identical "extreme" risk measures only for continuous distributions, that is,

$$\mathrm{CVaR}_{(\alpha)}(X) = \mathrm{TCE}_{(\alpha)}(X) = -\mathrm{E}[X \mid X > F_X^\leftarrow(\alpha)]. \tag{2.3}$$

To avoid confusion, in this paper, we use the term "α-risk measure" instead of terms like expected shortfall, CVaR, or tail conditional expectation.

Bassett et al. [3] showed that a portfolio with minimized α-risk can be constructed via the quantile regression (QR) methods of Koenker and Bassett [4]. QR is based on the fact that a quantile can be characterized as the minimizer of some expected asymmetric absolute loss function, namely,

$$\begin{aligned} F_X^\leftarrow(\alpha) &= \underset{\theta}{\operatorname{argmin}}\, \mathrm{E}\big[(\alpha \mathbb{1}\{X - \theta \geq 0\} + (1 - \alpha)\mathbb{1}\{X - \theta < 0\})|X - \theta|\big] \\ &=: \underset{\theta}{\operatorname{argmin}}\, \mathrm{E}\big[\rho_\alpha(X - \theta)\big], \end{aligned} \tag{2.4}$$

where $\rho_\alpha(u) := u(\alpha - \mathbb{1}\{u < 0\})$, $u \in \mathbb{R}$ is called the *check function* (see [5]), and $\mathbb{1}A$ is the indicator function defined by $\mathbb{1}A = \mathbb{1}_A(\omega) := 1$ if $\omega \in A$, $:= 0$ if $\omega \notin A$. To construct the optimal (i.e., α-risk minimized) portfolio, the following lemma is needed.

Lemma 2.1 (Theorem 2 of [3]). *Let X be a real-valued random variable with $\mathrm{E}X = \mu < \infty$, then*

$$\min_{\theta \in \mathbb{R}} \mathrm{E}\big[\rho_\alpha(X - \theta)\big] = \alpha\big(\mu + \rho_{\nu_\alpha}(X)\big). \tag{2.5}$$

Then, $Y = Y(\boldsymbol{\pi}) = \mathbf{X}'\boldsymbol{\pi}$ denotes a portfolio consisting of d different assets $\mathbf{X} := (X_1, \ldots, X_d)'$ with allocation weights $\boldsymbol{\pi} := (\pi_1, \ldots, \pi_d)'$ (subject to $\sum_{j=1}^d \pi_j = 1$), and the optimization problem under study is, for some prespecified expected return μ_0,

$$\min_{\boldsymbol{\pi}} \rho_{\nu_\alpha}(\mathbf{X}'\boldsymbol{\pi}) \quad \text{subject to } \mathrm{E}(\mathbf{X}'\boldsymbol{\pi}) = \mu_0, \qquad \mathbf{1}_d'\boldsymbol{\pi} = 1. \tag{2.6}$$

The sample or empirical analogue of this problem can be expressed as

$$\min_{\mathbf{b} \in \mathbb{R}^d} \sum_{i=1}^{n+1} \rho_\alpha(\mathbf{Z}_i - (\mathbf{A_w b})_i), \tag{2.7}$$

where X_{ij} denotes the jth sample value of asset i, $\overline{X}_i := n^{-1} \sum_{j=1}^{n} X_{ij}$,

$$\mathbf{Z} = (Z_1, \ldots, Z_n, Z_{n+1})' := \left(X_{11}, \ldots, X_{n1}, \kappa \left(\overline{X}_1 - \mu_0 \right) \right)',$$

$$\mathbf{A_w} := \begin{bmatrix} 1 & X_{11} - X_{21} & \cdots & X_{11} - X_{d1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & X_{1n} - X_{2n} & \cdots & X_{1n} - X_{dn} \\ \hline 0 & \kappa\left(\overline{X}_1 - \overline{X}_2\right) & \cdots & \kappa\left(\overline{X}_1 - \overline{X}_d\right) \end{bmatrix} =: \begin{bmatrix} \mathbf{W}'_1 \\ \vdots \\ \mathbf{W}'_n \\ \mathbf{W}'_{n+1} \end{bmatrix}, \tag{2.8}$$

with some κ sufficiently large. The minimizer of (2.7), namely,

$$\widehat{\boldsymbol{\beta}}^{(n)}(\alpha) := \underset{\mathbf{b} \in \mathbb{R}^d}{\operatorname{argmin}} \sum_{i=1}^{n+1} \rho_\alpha (Z_i - (\mathbf{A_w b})_i) = \underset{\mathbf{b} \in \mathbb{R}^d}{\operatorname{argmin}} \sum_{i=1}^{n+1} \rho_\alpha (Z_i - \mathbf{W}'_i \mathbf{b}) \tag{2.9}$$

and $\widehat{\pi}_1^{(n)}(\alpha) := 1 - \sum_{i=2}^{d} \widehat{\beta}_i^{(n)}(\alpha)$, provides the optimal weights yielding the minimal α-risk. The large sample properties of $\widehat{\boldsymbol{\beta}}^{(n)}(\alpha)$, especially its $\sqrt{n}$-consistency, can be implied from the standard arguments and assumptions in the QR context (see [5]).

Let $\overline{\mathbf{W}}^{(n)}$ and $\widehat{\boldsymbol{\Sigma}}_{\mathbf{W}}^{(n)}$ be the mean vector and the covariance matrix of $\mathbf{W}_i$ which are given by

$$\overline{\mathbf{W}}^{(n)} := \frac{1}{n+1} \sum_{i=1}^{n+1} \mathbf{W}_i, \qquad \widehat{\boldsymbol{\Sigma}}_{\mathbf{W}}^{(n)} := \frac{1}{n+1} \sum_{i=1}^{n+1} \left(\mathbf{W}_i - \overline{\mathbf{W}}^{(n)} \right) \left(\mathbf{W}_i - \overline{\mathbf{W}}^{(n)} \right)', \tag{2.10}$$

respectively. Here, the (p, q)th element of $\widehat{\boldsymbol{\Sigma}}_{\mathbf{W}}^{(n)}$ is

$$\widehat{\sigma}_{W,pq}^{(n)} = \begin{cases} \dfrac{n}{(n+1)^2} & \text{if } p = q = 1, \\[2mm] \dfrac{n(1-\kappa)}{(n+1)^2} \left(\overline{X}_1 - \overline{X}_p \right) & \text{if } q = 1, \ p = 2, \ldots, d, \\[2mm] \dfrac{n(1-\kappa)}{(n+1)^2} \left(\overline{X}_1 - \overline{X}_q \right) & \text{if } p = 1, \ q = 2, \ldots, d, \\[2mm] \widehat{\sigma}_{pq} + \dfrac{n(\kappa-1)^2}{(n+1)^2} \left(\overline{X}_1 - \overline{X}_p \right) \left(\overline{X}_1 - \overline{X}_q \right) & \text{if } p, q = 2, \ldots, d, \end{cases} \tag{2.11}$$

where

$$\widehat{\sigma}_{pq} := \frac{1}{n+1} \sum_{i=1}^{n} (X_{1i} - X_{pi})(X_{1i} - X_{qi}) - \frac{n}{n+1} \left(\overline{X}_1 - \overline{X}_p \right) \left(\overline{X}_1 - \overline{X}_q \right). \tag{2.12}$$

Let $\widehat{\mathbf{D}}_{\Sigma_W} := \mathrm{diag}\{\widehat{\sigma}_{W,11}^{(n)}, \ldots, \widehat{\sigma}_{W,dd}^{(n)}\}$. Then the correlation matrix of $\{\mathbf{W}_i\}_{i=1,\ldots,n+1}$ becomes $\mathbf{R} := \widehat{\mathbf{D}}_{\Sigma_W}^{-1/2}\widehat{\mathbf{\Sigma}}_{\mathbf{W}}^{(n)}\widehat{\mathbf{D}}_{\Sigma_W}^{-1/2}$, and the (p,q)th elemant of $\mathbf{R}$ is given by

$$
\begin{aligned}
&r_{pq} \\
&= \frac{1 + \left((n+1)^2/n(\kappa-1)^2\right)\cdot\left(\widehat{\sigma}_{pq}/\left(\overline{X}_1 - \overline{X}_p\right)\left(\overline{X}_1 - \overline{X}_q\right)\right)}{\left(1 + \left((n+1)^2/n(\kappa-1)^2\right)\cdot\left(\widehat{\sigma}_{pp}/\left(\overline{X}_1 - \overline{X}_p\right)^2\right)\right)^{1/2}\left(1 + \left((n+1)^2/n(\kappa-1)^2\right)\cdot\left(\widehat{\sigma}_{qq}/\left(\overline{X}_1 - \overline{X}_q\right)^2\right)\right)^{1/2}},
\end{aligned}
$$

$$(2.13)$$

for $p, q = 2, \ldots, d$. The above correlation coefficient can take values close to 1 when n/κ^2 is close to 0 with $(\overline{X}_1 - \overline{X}_p) \neq 0$ and $(\overline{X}_1 - \overline{X}_q) \neq 0$. Hence, the correlation of the estimated portfolio weights is possibly highly correlated among assets whose sample means differ from $\overline{X}_1$, while these problems are ignorable in an asymptotic inference problem if we take $\kappa = O(n^{1/2})$.

Thus far, we have seen that the α-risk-minimizing portfolio can be obtained by (2.9), which was the result of Bassett et al. [3]. In what follows, we show that semiparametrically efficient inference of the optimal weights $\widehat{\boldsymbol{\beta}}^{(n)}(\alpha)$ is feasible. The quantity estimated by (2.9) can be regarded as a QR coefficient $\boldsymbol{\beta}(\alpha)$, defined by

$$
F_{Z_i}^{\leftarrow}(\alpha \mid \mathbf{w}_i) := F_{Z_i|\mathbf{W}_i=\mathbf{w}_i}^{\leftarrow}(\alpha) =: \mathbf{w}_i'\boldsymbol{\beta}(\alpha), \tag{2.14}
$$

where $F_{X|S}^{\leftarrow}(\cdot)$ denotes a conditional quantile function, that is, $F_{X|S}^{\leftarrow}(\alpha) := \inf\{x : P(X \leq x|S) \geq \alpha\}$. Note that here the QR model (2.14) has a random coefficient regression (RCR) interpretation of the form $Z_i = \mathbf{W}_i'\boldsymbol{\beta}(U_i)$ with componentwise monotone increasing function $\boldsymbol{\beta}$ and random variables U_i that are uniformly distributed over $[0,1]$, that is, $U_i \sim \mathrm{Uniform}\,[0,1]$ (see [5]). Here, a choice such that $\boldsymbol{\beta}(u) = [\beta_1(u), \beta_2(u), \ldots, \beta_d(u)]' := [b_1 + F_\xi^{\leftarrow}(u), b_2, \ldots, b_d]'$ with F_ξ the distribution function of some independent and identically distributed (i.i.d.) n-tuple $(\xi_1, \ldots, \xi_n)$ yields

$$
Z_i = \mathbf{W}_i'\boldsymbol{\beta}(U_i) = \mathbf{W}_i'\boldsymbol{\beta}(F_\xi(\xi_i)) = \mathbf{W}_i'\begin{bmatrix} b_1 + \xi_i \\ b_2 \\ \vdots \\ b_d \end{bmatrix}. \tag{2.15}
$$

Hence, recalling that the first component of $\mathbf{W}_i$ is 1, it follows that, for any fixed $\alpha \in [0,1]$, the QR coefficient $\boldsymbol{\beta}(\alpha)$ can be characterized as the parameter $\mathbf{b} \in \mathbb{R}^d$ of a model such as

$$
Z_i = \mathbf{W}_i'\mathbf{b} + \xi_i, \quad \xi_i \overset{\mathrm{iid}}{\sim} G, \tag{2.16}
$$

where the density g of G is subject to

$$
g \in \mathcal{F}^\alpha := \left\{ f : \int_{-\infty}^0 f(x)dx = \alpha = 1 - \int_0^\infty f(x)dx \right\}, \tag{2.17}
$$

that is, $G^{\leftarrow}(\alpha) = 0$. Let us describe this model as $(\mathcal{Z}^{(n)}, \mathcal{A}^{(n)}, \mathcal{P}_Q^{(n)})$, with $\mathcal{P}_Q^{(n)} := \{P_{\mathbf{b},g}^{(n)} \mid \mathbf{b} \in \mathbb{R}^d, \ g \in \mathcal{F}^\alpha\}$, where $P_{\mathbf{b},g}^{(n)}$ denotes the distribution of an observation $\{Z_i\}_{i=1}^n$. This model (2.16) is a fixed-α submodel of (2.14) and is the parametric submodel through which we will achieve semiparametric efficiency.

The model (2.16) is a quantile-restricted linear regression model. But here we have no knowledge about the true density g, other than that it belongs to $\mathcal{F}^\alpha$, which allows us to identify $\mathbf{b}$. So, we arbitrarily choose f from $\mathcal{F}^\alpha$ and call it the "reference density" and correspondingly define a "reference model"

$$Z_i = \mathbf{W}_i'\mathbf{b} + e_i, \qquad e_i \overset{\text{iid}}{\sim} F, \tag{2.18}$$

where the density f of F is subject to $f \in \mathcal{F}^\alpha$. The goal of the next section is to construct an asymptotically efficient version of $\widehat{\boldsymbol{\beta}}_n(\alpha)$ based on some feasible $f \in \mathcal{F}^\alpha$, that is, attaining the semiparametric lower bound at correctly specified density $f = g$ that nevertheless remains $\sqrt{n}$-consistent under a misspecified density ($f \neq g$).

3. Semiparametrically Efficient Estimation

The procedure that we will apply here to achieve semiparametric efficiency is based on the invariance principle, as introduced by Hallin and Werker [13]. To this end, first we should have locally asymptotic normality (LAN; see, e.g., van der Vaart [14]) for a parametric submodel $P_{\mathbf{b},g}^{(n)}$, namely,

$$\log \frac{P_{\mathbf{b}+\mathbf{h}/\sqrt{n},g}^{(n)}}{P_{\mathbf{b},g}^{(n)}} = \mathbf{h}'\boldsymbol{\Delta}_{\mathbf{b};g}^{(n)} + \frac{1}{2}\mathbf{h}'\mathbf{I}_{\mathbf{b};g}\mathbf{h} + o_P(1), \qquad \mathbf{h} \in \mathbb{R}^d,$$

$$\boldsymbol{\Delta}_{\mathbf{b};g}^{(n)} \overset{d}{\longrightarrow} \mathcal{N}(\mathbf{0}, \mathbf{I}_{\mathbf{b};g}), \tag{3.1}$$

where all the stochastic convergences are taken under $P_{\mathbf{b},g} := P_{\mathbf{b},g}^{(\infty)}$. Here, the random vector $\boldsymbol{\Delta}_{\mathbf{b};g}^{(n)}$ is called the *central sequence*, and the positive definite matrix $\mathbf{I}_{\mathbf{b};g}$ is the *information matrix*. To ensure the LAN condition for model (2.18), the following assumption is required.

Assumption 3.1. The reference density f has finite Fisher information for location:

$$0 < \mathcal{I}_f := \int_{-\infty}^{\infty} \varphi_f(x)^2 f(x)dx = \int_0^1 \varphi_f(F^{\leftarrow}(u))^2 du < \infty, \qquad \text{where } \varphi_f(x) := \frac{-f'(x)}{f(x)}. \tag{3.2}$$

Assumption 3.2. The regression vectors $\mathbf{W}_i$ satisfy, under $P_{\mathbf{b},g}$,

$$\overline{\mathbf{W}}^{(n)} \overset{P}{\longrightarrow} \boldsymbol{\mu}_{\mathbf{W}}, \qquad \widehat{\boldsymbol{\Sigma}}_{\mathbf{W}}^{(n)} \overset{P}{\longrightarrow} \Sigma_{\mathbf{W}}, \tag{3.3}$$

for some vector $\boldsymbol{\mu}_{\mathbf{W}}$ and positive definite $\Sigma_{\mathbf{W}}$, where $\overline{\mathbf{W}}^{(n)}$ and $\widehat{\boldsymbol{\Sigma}}_{\mathbf{W}}^{(n)}$ are defined by (2.10).

Then, by Theorem 2.1 and Example 4.1 of Drost et al. [15], model (2.18) satisfies the uniform LAN condition for any $\mathbf{b}_n$ of the form $\mathbf{b} + O(n^{-1/2})$, with central sequence and information matrix

$$\boldsymbol{\Delta}_{\mathbf{b}_n;f}^{(n)} = \frac{1}{\sqrt{n+1}} \sum_{i=1}^{n+1} \varphi_f(e_{\mathbf{b}_n,i}) \mathbf{W}_i, \qquad \mathbf{I}_{*;f} = \mathfrak{I}_f(\Sigma_\mathbf{W} + \boldsymbol{\mu}_\mathbf{W}\boldsymbol{\mu}_\mathbf{W}'), \tag{3.4}$$

where $e_{\mathbf{b}_n,i}$ denotes the residual (i.e., $e_{\mathbf{b}_n,i} := Z_i - \mathbf{W}_i'\mathbf{b}_n$). Consequently, we have the contiguity $P_{\mathbf{b}_n+\mathbf{h}/\sqrt{n},f}^{(n)} \vartriangleright P_{\mathbf{b}_n,f}^{(n)}$, and of course $P_{\mathbf{b}_n,f}^{(n)} \vartriangleright P_{\mathbf{b},f}^{(n)}$ as well. Recall that the contiguity $P^{(n)} \vartriangleright Q^{(n)}$ means that for any sequence $S^{(n)}$, if $P^{(n)}(S^{(n)}) \to 0$, then $Q^{(n)}(S^{(n)}) \to 0$ also. The reason why we have specified uniform LAN, rather than LAN at single $\mathbf{b}$, is the one-step improvement, which will be discussed later.

By following Hallin and Werker [13], a semiparametrically efficient procedure can be obtained by projecting $\boldsymbol{\Delta}_{\mathbf{b}_n,f}^{(n)}$ on some σ-field to which the generating group for $\{P_{\mathbf{b}_n,f}^{(n)} \mid f \in \mathcal{F}^\alpha\}$ becomes *maximal invariant* (see, e.g., Schmetterer [16]). For the quantile-restricted regression model (2.16), such a σ-field is studied by Hallin et al. [6] and found to be generated by signs and ranks of the residuals. Here, let us denote the sign of a residual as $S_{\mathbf{b}_n,i}$, the rank of a residual as $R_{\mathbf{b}_n,i}^{(n)}$, and the σ-field generated by them as

$$\mathcal{B}_{\mathbf{b}_n}^{(n)} := \sigma\left(S_{\mathbf{b}_n,1}, \ldots, S_{\mathbf{b}_n,n}; R_{\mathbf{b}_n,1}^{(n)}, \ldots, R_{\mathbf{b}_n,n}^{(n)} \right). \tag{3.5}$$

Then, "good" inference should be based on

$$\begin{aligned}
\underset{\sim}{\boldsymbol{\Delta}}_{\mathbf{b}_n,f}^{(n)} &:= \mathrm{E}_{\mathbf{b}_n,f}^{(n)}\left[\boldsymbol{\Delta}_{\mathbf{b}_n,f}^{(n)} \mid \mathcal{B}_{\mathbf{b}_n}^{(n)} \right] = \frac{1}{\sqrt{n+1}} \sum_{i=1}^{n+1} \mathrm{E}_{\mathbf{b}_n,f}^{(n)}\left[\varphi_f(e_{\mathbf{b}_n,i}) \mid \mathcal{B}_{\mathbf{b}_n}^{(n)} \right] \mathbf{W}_i \\
&= \frac{1}{\sqrt{n+1}} \sum_{i=1}^{n+1} \mathrm{E}_{\mathbf{b}_n,f}^{(n)}\left[\varphi_f[F^\leftarrow(U_{\mathbf{b}_n,i})] \mid \mathcal{B}_{\mathbf{b}_n}^{(n)} \right] \mathbf{W}_i \\
&= \frac{1}{\sqrt{n+1}} \sum_{i=1}^{n+1} \varphi_f[F^\leftarrow(V_{\mathbf{b}_n,i})] \mathbf{W}_i + o_P(1),
\end{aligned} \tag{3.6}$$

where $U_{\mathbf{b}_n,i} := F(e_{\mathbf{b}_n,i})$ is i.i.d. uniform on $[0,1]$ under $P_{\mathbf{b}_n,f}^{(n)}$ and hence approximated by

$$V_{\mathbf{b}_n,i}^{(n)} := \begin{cases} \alpha \cdot \dfrac{R_{\mathbf{b}_n,i}^{(n)}}{N_{\mathbf{b}_n,L}^{(n)} + 1} & \text{if } R_{\mathbf{b}_n,i}^{(n)} \leq N_{\mathbf{b}_n,L}^{(n)}, \\[2ex] \alpha + (1-\alpha) \cdot \dfrac{R_{\mathbf{b}_n,i}^{(n)} - N_{\mathbf{b}_n,L}^{(n)}}{n - N_{\mathbf{b}_n,L}^{(n)} + 1} & \text{otherwise,} \end{cases} \tag{3.7}$$

with $N_{\mathbf{b}_n,L}^{(n)} := \#\{i \in \{1,\ldots,n\} \mid S_{\mathbf{b}_n,i} \leq 0\}$. In short, we are first rewriting the residual $e_{\mathbf{b}_n,i}$ as $F^\leftarrow(U_{\mathbf{b}_n,i})$ with realizations $U_{\mathbf{b}_n,i}$ of a $[0,1]$-uniform random variable, and then approximating those $U_{\mathbf{b}_n,i}$ as $V_{\mathbf{b}_n,i}^{(n)}$ given $\{N_{\mathbf{b}_n,L}^{(n)}; R_{\mathbf{b}_n,1}^{(n)}, \ldots, R_{\mathbf{b}_n,n}^{(n)}\}$.

Using this rank-based central sequence, we can construct the *one-step estimator* (see, e.g., Bickel [17]; Bickel et al. [18]) as follows.

Definition 3.3. For any sequence of estimators $\widehat{\theta}_n$, the discretized estimator $\overline{\theta}_n$ is defined to be the nearest vertex of $\{\theta : \theta = (1/\sqrt{n})(i_1, i_2, \ldots, i_d)', i_j : \text{integers}\}$.

Definition 3.4. Let $\overline{\boldsymbol{\beta}}^{(n)}(\alpha)$ be the discretized version of $\widehat{\boldsymbol{\beta}}^{(n)}(\alpha)$ defined at (2.9). We define the (rank-based) one-step estimator of **b** based on reference density $f \in \mathcal{F}^\alpha$ as

$$\widehat{\mathbf{b}}_f^{(n)} := \overline{\boldsymbol{\beta}}^{(n)}(\alpha) + \widehat{\boldsymbol{\Sigma}}_{fg}^{-1} \frac{\boldsymbol{\Delta}_{\sim \overline{\boldsymbol{\beta}}^{(n)}(\alpha),f}^{(n)}}{\sqrt{n}}, \quad \text{with } \widehat{\boldsymbol{\Sigma}}_{fg} := \widehat{\mathcal{D}}_{fg}\Sigma_{\mathbf{W}} - \frac{f(0)}{1-\alpha}\widehat{\mu}_{\varphi_g}^- \boldsymbol{\mu}_{\mathbf{W}}\boldsymbol{\mu}_{\mathbf{W}}', \tag{3.8}$$

where $\widehat{\mathcal{D}}_{fg}$ and $\widehat{\mu}_{\varphi_g}^-$ are consistent estimates of

$$\mathcal{D}_{fg} := \int_0^1 \varphi_f[F^{\leftarrow}(u)]\varphi_g[G^{\leftarrow}(u)]du, \tag{3.9}$$

$$\mu_{\varphi_g}^- := \mathrm{E}\left[\varphi_g[g^{\leftarrow}(U)] \mid U \le \alpha\right] \quad \left(= \frac{-g(0)}{\alpha}\right), \tag{3.10}$$

respectively.

Consistent estimates $\widehat{\mathcal{D}}_{fg}$ and $\widehat{\mu}_{\varphi_g}^-$ can be obtained in the manner of Hallin et al. [19], which is done without the kernel estimation of g, though here we omit the details.

Lemma 3.5 (Section 4.1 of [6]). *Under $P_{\mathbf{b},g}$ with $g \in \mathcal{F}^\alpha$,*

$$\sqrt{n}\left(\widehat{\mathbf{b}}_f^{(n)} - \mathbf{b}\right) \xrightarrow{d} \mathcal{N}\left(\mathbf{0}, \ \Sigma_{fg}^{-1}\Sigma_{ff}\Sigma_{fg}^{-1}\right). \tag{3.11}$$

*Therefore, the one-step estimator $\widehat{\mathbf{b}}_f^{(n)}$ defined by (3.8) for **b** is semiparametrically efficient at $f = g$.*

In our original notation, the above statement can be rewritten as, for some $\alpha \in (0,1)$ fixed, $\sqrt{n}(\widehat{\mathbf{b}}_f^{(n)} - \boldsymbol{\beta}(\alpha)) \xrightarrow{d} \mathcal{N}(\mathbf{0}, \ \Sigma_{fg}^{-1}\Sigma_{ff}\Sigma_{fg}^{-1})$.

Recall that the standard QR estimator, defined at (2.9), is asymptotically normal (see Koenker [5]):

$$\sqrt{n}\left(\widehat{\boldsymbol{\beta}}^{(n)}(\alpha) - \boldsymbol{\beta}(\alpha)\right) \xrightarrow{d} \mathcal{N}\left(\mathbf{0}, \ \mathbf{D}^{-1}\right), \tag{3.12}$$

where

$$\mathbf{D} := \frac{g^2(0)}{\alpha(1-\alpha)}\mathbf{D}_0, \quad \text{with } \mathbf{D}_0 = \lim_{n\to\infty}\frac{1}{n+1}\sum_{i=1}^{n+1}\mathbf{W}_i\mathbf{W}_i'. \tag{3.13}$$

Denote the true portfolio weight with respect to risk probability α by $\boldsymbol{\pi} = (1 - \mathbf{1}'_{d-1}\boldsymbol{\pi}_2(\alpha),$ $\boldsymbol{\pi}_2(\alpha)')'$, where $\boldsymbol{\pi}_2(\alpha) = (\pi_2(\alpha), \dots, \pi_d(\alpha))'$, and its standard quantile regression and our one-step estimators by $\widehat{\boldsymbol{\pi}}^{(\mathrm{QR})} := (1 - \mathbf{1}'_{d-1}\widehat{\boldsymbol{\beta}}_2^{(n)}(\alpha), \widehat{\boldsymbol{\beta}}_2^{(n)}(\alpha)')'$ and $\widehat{\boldsymbol{\pi}}_f^{(\mathrm{OS})} := (1 - \mathbf{1}'_{d-1}\widehat{\mathbf{b}}_{f,2}^{(n)}, \widehat{\mathbf{b}}_{f,2}^{(n)'})'$, respectively. Denote the block matrix of the covariance matrix of standard quantile and one-step estimators by

$$\mathbf{D}^{-1} = \begin{pmatrix} D_{11} & \mathbf{D}_{12} \\ \mathbf{D}'_{12} & \mathbf{D}_{22} \end{pmatrix}, \qquad \boldsymbol{\Sigma}_{fg}^{-1}\boldsymbol{\Sigma}_{ff}\boldsymbol{\Sigma}_{fg}^{-1} = \begin{pmatrix} \Sigma_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}'_{12} & \boldsymbol{\Sigma}_{22} \end{pmatrix}, \tag{3.14}$$

where submatrices $\mathbf{D}_{22}$ and $\boldsymbol{\Sigma}_{22}$ are $(d-1) \times (d-1)$ symmetric matrices for the covariance of portfolio weights $\boldsymbol{\pi}_2$. Then we obtain the variances of the α-risk-minimizing portfolio constructed by the standard quantile, and the one-step estimators are stated in the following proposition. Since direct evaluation gives the following statement, we skip its proof.

Proposition 3.6. *The asymptotic conditional variances of an α-risk-minimizing portfolio using the standard quantile regression and one-step estimators given at $\mathbf{X} = \mathbf{x}$ are, respectively,*

$$\mathrm{Var}\left(\mathbf{X}'\widehat{\boldsymbol{\pi}}^{(\mathrm{QR})} \mid \mathbf{X} = \mathbf{x} = [x_1, \mathbf{x}'_2]'\right) = x_1^2 \cdot \mathrm{tr}\left[\mathbf{1}_{d-1}\mathbf{1}'_{d-1}\mathbf{D}_{22}\right] + 2x_1 \cdot \mathrm{tr}\left[\mathbf{1}_{d-1}\mathbf{x}'_2\mathbf{D}_{22}\right] + \mathrm{tr}\left[\mathbf{x}_2\mathbf{x}'_2\mathbf{D}_{22}\right],$$

$$\mathrm{Var}\left(\mathbf{X}'\widehat{\boldsymbol{\pi}}_f^{(\mathrm{OS})} \mid \mathbf{X} = \mathbf{x} = [x_1, \mathbf{x}'_2]'\right) = x_1^2 \cdot \mathrm{tr}\left[\mathbf{1}_{d-1}\mathbf{1}'_{d-1}\boldsymbol{\Sigma}_{22}\right] + 2x_1 \cdot \mathrm{tr}\left[\mathbf{1}_{d-1}\mathbf{x}'_2\boldsymbol{\Sigma}_{22}\right] + \mathrm{tr}\left[\mathbf{x}_2\mathbf{x}'_2\boldsymbol{\Sigma}_{22}\right],$$

$$\tag{3.15}$$

where $\mathbf{x}_2 = (x_2, \dots, x_d)'$.

For any positive definite matrices A and B, we say $A \leq B$ if $B - A$ is nonnegative definite. To compare the efficiency of the standard quantile regression estimator and the one-step estimator, we need to show that $\boldsymbol{\Sigma}_{fg}^{-1}\boldsymbol{\Sigma}_{ff}\boldsymbol{\Sigma}_{fg}^{-1} \leq \mathbf{D}^{-1}$. To see this, as in Section 3 of Koenker and Zhao [20], let us consider

$$\boldsymbol{\Sigma} := \begin{pmatrix} \boldsymbol{\Sigma}_{fg}\boldsymbol{\Sigma}_{ff}^{-1}\boldsymbol{\Sigma}_{fg} & \boldsymbol{\Sigma}_{fg} \\ \boldsymbol{\Sigma}_{fg} & \boldsymbol{\Sigma}_{fg}\mathbf{D}^{-1}\boldsymbol{\Sigma}_{fg} \end{pmatrix}. \tag{3.16}$$

Note that $\boldsymbol{\Sigma}$ is a nonnegative definite matrix. If $\boldsymbol{\Sigma}_{fg}^{-1}\mathbf{D}\boldsymbol{\Sigma}_{fg}^{-1}$ is a positive definite, then there exists orthogonal matrix $\mathbf{P}$, such that

$$\mathbf{P}'\boldsymbol{\Sigma}\mathbf{P} = \begin{pmatrix} \boldsymbol{\Sigma}_{fg}\boldsymbol{\Sigma}_{ff}^{-1}\boldsymbol{\Sigma}_{fg} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Sigma}_{fg}\mathbf{D}^{-1}\boldsymbol{\Sigma}_{fg} - \boldsymbol{\Sigma}_{ff} \end{pmatrix}, \tag{3.17}$$

so $\Sigma_{fg}\mathbf{D}^{-1}\Sigma_{fg} - \Sigma_{ff}$ is nonnegative definite. Hence, $\mathbf{D}^{-1} - \Sigma_{fg}^{-1}\Sigma_{ff}\Sigma_{fg}^{-1}$ is nonnegative definite if Σ_{fg} is nonsingular. This result assures that the one-step estimator is asymptotically more efficient than the standard quantile regression estimator. From this result, it is easy to see that

$$\mathrm{Var}\left(\mathbf{X}'\widehat{\boldsymbol{\pi}}_f^{(\mathrm{OS})} \mid \mathbf{X} = \mathbf{x}\right) \leq \mathrm{Var}\left(\mathbf{X}'\widehat{\boldsymbol{\pi}}^{(\mathrm{QR})} \mid \mathbf{X} = \mathbf{x}\right). \tag{3.18}$$

Also, by taking expectation on both sides, the same inequality holds for unconditional variances.

4. Numerical Studies

In this section, we examine the finite sample properties of the proposed one-step estimator described in Section 3 for the cases where $\alpha = 0.1$ and 0.5. Our simulations are performed with two data generating processes to focus on the underlying true density g and how the choice of the reference density f might affect the finite sample performances.

The first data-generating process (DGP1) is the same as that investigated by Bassett et al. [3]. For DGP1, we consider the construction of an α-minimizing portfolio from four independently distributed assets, that is, asset 1 is normally distributed with mean 0.05 and standard deviation 0.02. Asset 2 is a reversed χ_3^2 density with location and scale chosen so that its mean and variance are identical to those of asset 1. Asset 3 is normally distributed with mean 0.09 and standard deviation 0.05. Finally, asset 4 has a χ_3^2 density with identical mean and standard deviation to asset 3. DGP2 is a four-dimensional normal distribution with mean vectors the same as those of DGP1, and covariance matrix $\Sigma = [\sigma_{ij}]_{i,j=1,\ldots,4}$ with $\mathrm{diag}\,\Sigma = (0.02, 0.02, 0.05, 0.05)$ and $\sigma_{i,j}$ for $i \neq j$ is $\sigma_{ii}\sigma_{jj}\rho$. Here, we set $\rho = 0.5$, which indicates that the asset returns have correlation 0.5. Notice that both DGP1 and DGP2 have the same mean and variance structures. The underlying true conditional densities of $\mathbf{u} = \mathbf{Z} - \mathbf{A}_\mathbf{w}'\mathbf{b}$ for DGP1 and DGP2 are a mixture of the normal χ_3^2 and reversed χ_3^2 distribution and the normal distributions, respectively. A simulation of the estimator, for sample, size $n = 100, 500$, and 1000 consists of 1000 replications. We choose prespecified expected return μ_0 at 0.07.

For each scenario, we computed standard quantile regression estimates $\widehat{\boldsymbol{\beta}}^{(n)}(\alpha)$ with corresponding portfolio weights $\widehat{\boldsymbol{\pi}}^{(\mathrm{QR})} = (1 - \sum_{j=2}^d \widehat{\beta}_j^{(n)}(\alpha), \widehat{\beta}_2^{(n)}(\alpha), \ldots, \widehat{\beta}_d^{(n)}(\alpha))$, and our one-step estimates are defined by (3.8) for various choices of the reference density f and actual density g in the α-minimizing portfolio allocation problem.

To make the problem a pure location model, we set the variance of the estimated residual to have one, that is, $\widehat{\mathbf{u}} = \widetilde{\mathbf{u}}/\sqrt{\mathrm{Var}(\widetilde{\mathbf{u}})}$, where $\widetilde{\mathbf{u}} = [\widetilde{u}_i]_{i=1,\ldots,n+1} = Z_i - (\mathbf{A}_\mathbf{w}\overline{\mathbf{b}}^{(n)}(\alpha))_i$. The true density g can be estimated by the kernel estimator for DGP1,

$$\widehat{g}(u) = \frac{1}{(n+1)h} \sum_{i=1}^{n+1} K\left(\frac{u - \widehat{u}_i}{h}\right), \tag{4.1}$$

where K is a kernel function, and h is a bandwidth. The first derivative $g'(u)$ is estimated by

$$\widehat{g}'(u) = \frac{1}{(n+1)h^2} \sum_{i=1}^{n+1} K'\left(\frac{u - \widehat{u}}{h}\right). \tag{4.2}$$

As for the DGP2, the actual density g becomes normal because the portfolio is constructed by normally distributed returns. We use the normal distribution ($\mathcal{N}$), the asymmetric Laplace distribution (AL), the logistic distribution (LGT), and the asymmetric power distribution (APD) with $\lambda = 1.5$ for the reference density f.

The density function of the asymmetric power distribution introduced by Komunjer [7] is given by

$$
f(u) = \begin{cases} \dfrac{\delta_{\alpha,\lambda}^{1/\lambda}}{\Gamma(1 + 1/\lambda)} \exp\left[-\dfrac{\delta_{\alpha,\lambda}}{\alpha^\lambda} |u|^\lambda \right] & \text{if } u \leq 0, \\[3mm] \dfrac{\delta_{\alpha,\lambda}^{1/\lambda}}{\Gamma(1 + 1/\lambda)} \exp\left[-\dfrac{\delta_{\alpha,\lambda}}{(1 - \alpha)^\lambda} |u|^\lambda \right] & \text{if } u > 0, \end{cases} \tag{4.3}
$$

where $0 \leq \alpha < 1,\ \lambda > 0$, and

$$
\delta_{\alpha,\lambda} = \frac{2\alpha^\lambda (1 - \alpha)^\lambda}{\alpha^\lambda + (1 - \alpha)^\lambda}. \tag{4.4}
$$

When $\alpha = 0.5$, the APD pdf is symmetric around zero. In this case, the APD density reduces to the standard generalized power distribution (GPD) [21, pages 194-195]. Special cases of the GPD include uniform ($\lambda = \infty$), Gaussian ($\lambda = 2$), and Laplace ($\lambda = 1$) distributions. When $\alpha \neq 0.5$, the APD pdf is asymmetric. Special cases include asymmetric Laplace ($\lambda = 1$), the two-piece normal ($\lambda = 2$) distributions.

For a given sample size, we compute simulated mean and standard deviation of $\widehat{\boldsymbol{\pi}}^{(\mathrm{QR})}$ and $\widehat{\boldsymbol{\pi}}_f^{(\mathrm{OS})}$ and the relative efficiency $\mathrm{Var}(\widehat{\pi}_{f,i}^{(\mathrm{OS})}) / \mathrm{Var}(\widehat{\pi}_i^{(\mathrm{QR})})$ for $i = 1, \ldots, 4$.

Table 1 gives the results of the relative efficiencies for DGP 1. When $\alpha = 0.1$, we see that the efficiency gains of one-step estimators with asymmetric Laplace reference density are large compared with other reference densities with $n = 1000$, while these efficiency gains are less when sample size is $n = 100$. When $\alpha = 0.5$, relative efficiency of assets 3 and 4 with asymmetric Laplace reference density is minimum, while for assets 1 and 2, relative efficiency with normal reference density is minimum. This is because of the covariance structure of $\widehat{\boldsymbol{\Sigma}}_{\mathbf{W}}^{(n)}$ defined by (2.10). As can be seen in Section 2, if $\mu_1 \neq \mu_p$ and $\mu_1 \neq \mu_q$, the (p, q)th element of the correlation matrix defined by (2.13) has a value close to unity. In this case, the asymptotic variance of the usual quantile regression estimator becomes large, which leads to unsatisfactorily large variances in assets 3 and 4. However, the asymptotic variance of our one-step estimator does not have such problems.

Table 2 gives the results of the relative efficiencies for DGP2. In line with efficiency at a correctly specified reference density $f = \mathcal{N}$, we see that the relative efficiency is minimal for all assets and sample sizes with $\alpha = 0.1$ and 0.5. Even though we misspecify the reference density $f \neq \mathcal{N}$, there exists some sort of efficiency gain except for assets 1 and 2 of the asymmetric Laplace reference density with $n = 100$ and $\alpha = 0.1$. Efficiency gains for the normal reference density and logistic reference density are almost the same because the underlying true density is a symmetric normal distribution, and the asymmetric power reference density with $\lambda = 1.5$ outperforms the asymmetric Laplace reference density.

Figure 1 plots the kernel densities for the estimated portfolio weights for DGP2 with $\alpha = 0.5$ and $n = 1000$. We see that the standard quantile regression estimators have long

Table 1: $\text{Var}(\hat{\boldsymbol{\pi}}_f^{(OS)})/\text{Var}(\hat{\boldsymbol{\pi}}^{(QR)})$ for DGP 1 [3]: in this case, we estimate an unknown g (which must be $\mathcal{N}$-χ^2 mixed) using the kernel method.

f	π							
	π_1 (.1)	π_2 (.1)	π_3 (.1)	π_4 (.1)	π_1 (.5)	π_2 (.5)	π_3 (.5)	π_4 (.5)
$n = 100$								
AL	1.1054	1.1088	1.0036	1.0082	1.0211	1.0193	0.9925	0.9570
N	0.9773	0.9826	0.9093	0.9331	0.9412	0.9383	0.9227	0.9490
LGT	0.9731	0.9788	0.9017	0.9289	0.9458	0.9436	0.9228	0.9463
APD$_{1.5}$	0.9851	0.9838	0.9827	0.9893	0.9517	0.9492	0.9273	0.9471
$n = 500$								
AL	1.0266	1.0068	0.9549	0.9871	0.9596	0.9563	0.3626	0.5478
N	0.9762	0.9778	0.8966	0.9340	0.9186	0.9138	0.7522	0.8285
LGT	0.9739	0.9769	0.8893	0.9301	0.9255	0.9225	0.7277	0.8102
APD$_{1.5}$	0.9765	0.9712	0.9773	0.9896	0.9291	0.9262	0.6922	0.7842
$n = 1000$								
AL	0.8702	0.9084	0.5621	0.8193	0.9150	0.8985	0.2225	0.3614
N	0.9416	0.9485	0.8680	0.9112	0.7850	0.7714	0.4019	0.5214
LGT	0.9453	0.9532	0.8640	0.9085	0.8101	0.7981	0.4043	0.5199
APD$_{1.5}$	0.9290	0.9296	0.9242	0.9645	0.8206	0.8077	0.3572	0.4806

Table 2: $\text{Var}(\hat{\boldsymbol{\pi}}_f^{(OS)})/\text{Var}(\hat{\boldsymbol{\pi}}^{(QR)})$ for DGP2 (multinormal). In this case, residual density is a normal distribution. Hence, we adopt $g = \mathcal{N}$.

f	π							
	π_1 (.1)	π_2 (.1)	π_3 (.1)	π_4 (.1)	π_1 (.5)	π_2 (.5)	π_3 (.5)	π_4 (.5)
$n = 100$								
AL	1.0890	1.0996	0.7311	0.7221	0.7428	0.7584	0.6652	0.6370
N	0.5891	0.6013	0.4105	0.4199	0.5896	0.6010	0.5451	0.5347
LGT	0.6153	0.6280	0.4113	0.4229	0.6026	0.6154	0.5503	0.5356
APD$_{1.5}$	0.8512	0.8603	0.6120	0.6069	0.6076	0.6204	0.5580	0.5420
$n = 500$								
AL	0.8970	0.8814	0.7742	0.7402	0.6944	0.7030	0.6235	0.5965
N	0.5090	0.4993	0.4883	0.4648	0.5618	0.5666	0.5169	0.4949
LGT	0.5184	0.5096	0.4892	0.4678	0.5703	0.5761	0.5321	0.5086
APD$_{1.5}$	0.7271	0.7183	0.6746	0.6423	0.5723	0.5784	0.5317	0.5099
$n = 1000$								
AL	0.8836	0.8633	0.8015	0.8712	0.8778	0.8612	0.6726	0.7390
N	0.4800	0.4655	0.4861	0.5095	0.7071	0.6944	0.5746	0.6261
LGT	0.4897	0.4750	0.4939	0.5198	0.7238	0.7094	0.5831	0.6347
APD$_{1.5}$	0.7144	0.7041	0.6617	0.7088	0.7261	0.7126	0.5845	0.6395

tails on both sides for all assets, whereas one-step estimators have a narrower interval and higher peak at the true weight. This confirms that the one-step estimators are more semiparametrically efficient than the standard ones.
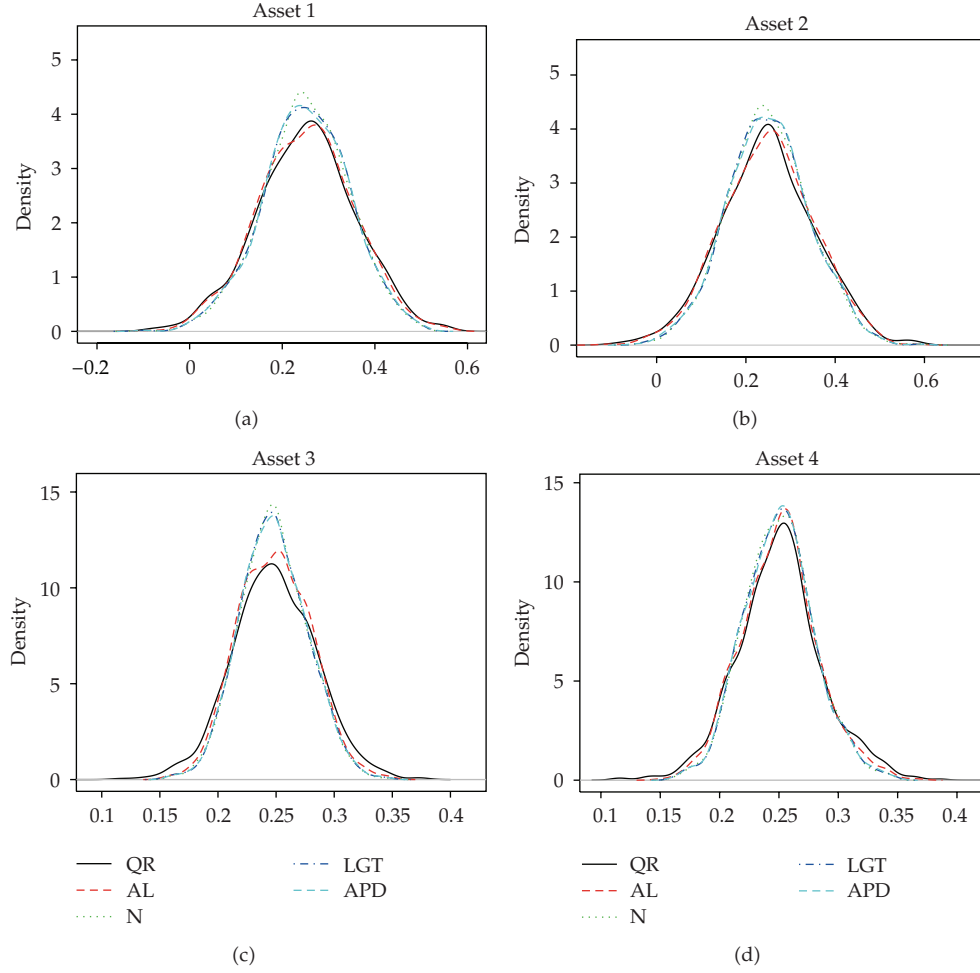
Figure 1: Kernel density plots for the portfolio weights. Panels (a) to (d) correspond to the kernel density for assets 1 to 4, respectively. The density shows the standard quantile estimator (QR; solid line); the estimator with an asymmetric Laplace distribution reference density (AL; dashed line); normal distribution (N; dotted line); logistic distribution (LGT; dotted-dashed line); asymmetric power distribution (APD; long dashed line).

5. Empirical Application

We apply our methodology to weekly log returns of the 96 stocks of the TOPIX large 100 index. The samples run from January 5, 2007, to December 2, 2011, for a total of 257 observations. The stock prices are adjusted to take into account events such as stock splits on individual securities. Preliminary tests reveal that most log return series have high values of kurtosis and negative values of skewness in general, which indicates that the log returns are non-Gaussian.

We computed the optimal portfolio allocations for $\alpha = 0.01, 0.05, 0.1, 0.2, 0.3, 0.4$, and 0.5. We set $\kappa = 1000$ and $\mu = -0.002$, which is the third quartile of the average log-return distribution. For the first-round estimates, we used the standard quantile regression estimator, and for the one-step estimates, we chose a normal distribution as a reference
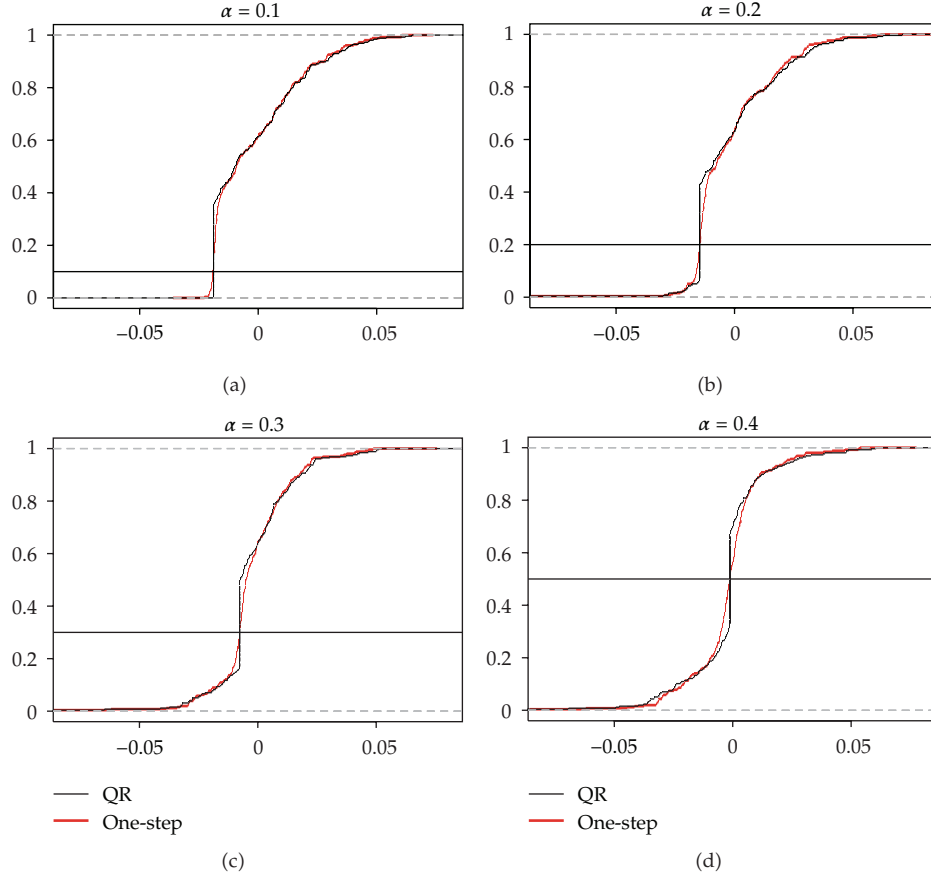
Figure 2: Empirical cumulative distribution function of the α-risk-minimizing portfolio based on the standard quantile regression estimator (thin line) and one-step estimator (thick line) for α = 0.1, 0.2, 0.3, and 0.5, which corresponds to (a) to (d), respectively.

density. Since we do not have enough information about the shape of the portfolio distributions for the various choices of α, the actual density g is estimated by the kernel method.

Figure 2 plots the cumulative distribution functions of the α-risk-minimizing portfolios obtained by the standard quantile regression estimates and one-step estimates for $\alpha = 0.1, 0.2, 0.3$, and 0.5. Summary statistics for the distributions of the different portfolios are reported in Table 3.

Figure 2 and Table 3 clearly show that the optimal α-risk-minimizing portfolio manages to reduce the occurrence of events in the left tail when α is small for both standard QR estimates and one-step estimates. The standard deviation of the one-step estimates of an α-minimizing portfolio is smaller than that of the standard QR estimates. We can also observe that the range of a constructed portfolio with one-step estimates is much smaller than that of standard QR estimates, due to the semiparametric efficiency properties of our one-step estimators. When α becomes large, the difference in the standard deviation of the constructed portfolio between standard QR estimates and one-step estimates tends to become large. Hence, efficiency gains are large for $\alpha = 0.5$, which is the mean absolute deviation portfolio.
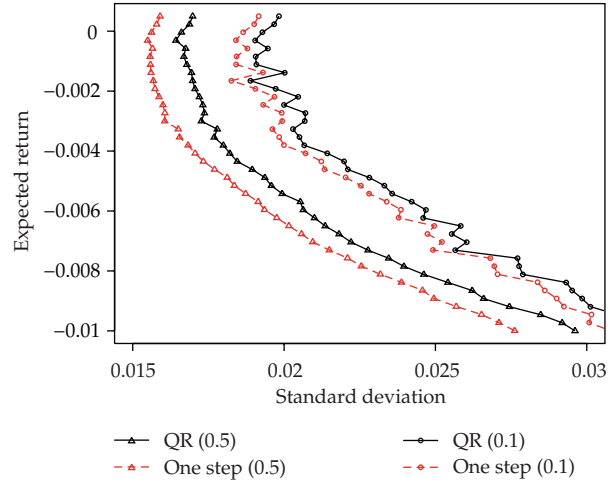
Figure 3: Efficient frontiers for an α-risk-minimizing portfolio based on the standard quantile regression estimator and the one-step one. The lines with triangles and circles represent the pair of obtained standard deviation and mean for the portfolio with $\alpha = 0.5$ and 0.1, respectively. The solid and dashed lines represent risks and returns for the standard quantile regression and one-step portfolios, respectively.

Table 3: Summary statistics for the α-minimizing-portfolios using quantile regression methods $\mathbf{X}'\widehat{\pi}^{(QR)}$ and one-step estimates $\mathbf{X}'\widehat{\pi}_f^{(OS)}$.

| α | $\mathbf{X}'\widehat{\pi}^{(QR)}$ | | | | | $\mathbf{X}'\widehat{\pi}_f^{(OS)}$ | | | | |
	Min	Max	Mean	Std. dev.	$Q(\alpha)$	Min	max	Mean	Std. dev.	$Q(\alpha)$
0.01	−0.0210	0.0557	−0.0020	0.0187	−0.0210	−0.0210	0.0555	−0.0020	0.0186	−0.0209
0.05	−0.0192	0.0756	−0.0020	0.0196	−0.0192	−0.0202	0.0733	−0.0020	0.0192	−0.0194
0.1	−0.0190	0.0649	−0.0020	0.0196	−0.0190	−0.0227	0.0606	−0.0020	0.0189	−0.0191
0.2	−0.0959	0.0708	−0.0020	0.0189	−0.0149	−0.0944	0.0624	−0.0020	0.0179	−0.0146
0.3	−0.1247	0.0512	−0.0020	0.0171	−0.0078	−0.1200	0.0487	−0.0020	0.0163	−0.0077
0.4	−0.1307	0.0545	−0.0020	0.0173	−0.0038	−0.1230	0.0482	−0.0020	0.0162	−0.0040
0.5	−0.1011	0.0605	−0.0020	0.0171	−0.0012	−0.0958	0.0537	−0.0020	0.0158	−0.0013

Note: the corresponding summary statistics of the TOPIX log returns for minimum, maximum, mean, and standard deviation are −0.2202, 0.0924, −0.0032, and 0.0328, respectively. Also, quantiles for the TOPIX log returns for $\alpha = 0.01, 0.05, 0.1, 0.2, 0.3, 0.4,$ and 0.5 are −0.0981, −0.0536, −0.0158, −0.0068, and 0.0006, respectively.

Another interesting finding is that the standard QR-constructed portfolios have high-density peaks at the required quantiles for all values of α, whereas the portfolio constructed by one-step estimates has a quite moderate density reduction at the required quantiles.

Consistent with economic intuition, higher risk aversion is associated with a shorter left tail. In the case where $\alpha \leq 0.1$, maximum loss is limited to less than −0.02. This result is particularly striking given that the sample includes the stock market crash of October 2008 due to the US subprime mortgage crisis and the bankruptcy of Lehman Brothers, which resulted in a weekly loss of more than −0.220 for TOPIX. The sample also includes the stock market crash of March 2011 due to the catastrophic earthquake and tsunami that hit Japan, which resulted in a weekly loss of −0.104.

Figure 3 presents empirical efficient frontiers corresponding to the standard quantile regression-based portfolios and one-step estimates of a portfolio with $\alpha = 0.1$ and 0.5.

Figure 3 clearly illustrates that the standard quantile regression-based portfolio is completely inefficient, far from the one-step frontier.

6. Summary and Conclusions

This paper considered a semiparametrically efficient estimation of an α-risk-minimizing portfolio. A one-step estimator based on residual signs and ranks was proposed, and simulations were performed to compare the finite sample relative efficiencies for the standard quantile regression estimators and the one-step one. These simulations confirmed our theoretical findings. An empirical application to construct a portfolio using 96 Japanese stocks was investigated and confirms that the one-step α-risk-minimizing portfolio has smaller variance that is obtained by the standard quantile regression estimator.

Further research topics include (1) construction of portfolios without short-sale constraints and (2) extending the results to the covariates of time series with heteroskedastic returns. For the former, we impose nonnegativity of the weights by using a penalty function containing a term that diverges to infinity as any of the weights becomes negative (see [22]). For the latter, we refer to Hallin et al. [6] and Taniai [23].

Acknowledgments

References

[1] P. Artzner, F. Delbaen, J.-M. Eber, and D. Heath, "Coherent measures of risk," *Mathematical Finance*, vol. 9, no. 3, pp. 203–228, 1999.

[2] H. Föllmer and A. Schied, "Convex measures of risk and trading constraints," *Finance and Stochastics*, vol. 6, no. 4, pp. 429–447, 2002.

[3] G. W. J. Bassett, R. Koenker, and G. Kordas, "Pessimistic portfolio allocation and choquet expected utility," *Journal of Financial Econometrics*, vol. 2, no. 4, pp. 477–492, 2004.

[4] R. Koenker and G. Bassett, Jr., "Regression quantiles," *Econometrica*, vol. 46, no. 1, pp. 33–50, 1978.

[5] R. Koenker, *Quantile Regression*, vol. 38 of *Econometric Society Monographs*, Cambridge University Press, Cambridge, UK, 2005.

[6] M. Hallin, C. Vermandele, and B. J. M. Werker, "Semiparametrically efficient inference based on signs and ranks for median-restricted models," *Journal of the Royal Statistical Society B*, vol. 70, no. 2, pp. 389–412, 2008.

[7] I. Komunjer, "Asymmetric power distribution: theory and applications to risk measurement," *Journal of Applied Econometrics*, vol. 22, no. 5, pp. 891–921, 2007.

[8] Q. Zhao, "Asymptotically efficient median regression in the presence of heteroskedasticity of unknown form," *Econometric Theory*, vol. 17, no. 4, pp. 765–784, 2001.

[9] Y.-J. Whang, "Smoothed empirical likelihood methods for quantile regression models," *Econometric Theory*, vol. 22, no. 2, pp. 173–205, 2006.

[10] I. Komunjer and Q. Vuong, "Semiparametric efficiency bound in time-series models for conditional quantiles," *Econometric Theory*, vol. 26, no. 2, pp. 383–405, 2010.

[11] S. Uryasev, "Conditional value-at-risk: optimization algorithms and applications," in *Proceedings of the IEEE/IAFE/INFORMS Conference on Computational Intelligence for Financial Engineering (CIFEr '00)*, pp. 49–57, 2000.

[12] G. C. Pflug, "Some remarks on the value-at-risk and the conditional value-at-risk," in *Probabilistic Constrained Optimization*, vol. 49 of *Nonconvex Optimization and Its Applications*, pp. 272–281, Kluwer Academic Publishers, Dodrecht, The Netherlands, 2000.

[13] M. Hallin and B. J. M. Werker, "Semi-parametric efficiency, distribution-freeness and invariance," *Bernoulli*, vol. 9, no. 1, pp. 137–165, 2003.

[14] A. W. van der Vaart, *Asymptotic Statistics*, vol. 3 of *Cambridge Series in Statistical and Probabilistic Mathematics*, Cambridge University Press, Cambridge, UK, 1998.

[15] F. C. Drost, C. A. J. Klaassen, and B. J. M. Werker, "Adaptive estimation in time-series models," *The Annals of Statistics*, vol. 25, no. 2, pp. 786–817, 1997.

[16] L. Schmetterer, *Introduction to Mathematical Statistics*, Springer, Berlin, Germany, 1974.

[17] P. J. Bickel, "On adaptive estimation," *The Annals of Statistics*, vol. 10, no. 3, pp. 647–671, 1982.

[18] P. J. Bickel, C. A. J. Klaassen, Y. Ritov, and J. A. Wellner, *Efficient and Adaptive Estimation for Semiparametric Models*, Johns Hopkins Series in the Mathematical Sciences, Johns Hopkins University Press, Baltimore, Md, USA, 1993.

[19] M. Hallin, H. Oja, and D. Paindaveine, "Semiparametrically efficient rank-based inference for shape. II. Optimal *R*-estimation of shape," *The Annals of Statistics*, vol. 34, no. 6, pp. 2757–2789, 2006.

[20] R. Koenker and Q. Zhao, "Conditional quantile estimation and inference for ARCH models," *Econometric Theory*, vol. 12, no. 5, pp. 793–813, 1996.

[21] N. L. Johnson, S. Kotz, and N. Balakrishnan, *Continuous Univariate Distributions. Vol. 1*, Wiley Series in Probability and Mathematical Statistics: Applied Probability and Statistics, John Wiley & Sons, New York, NY, USA, 2nd edition, 1994.

[22] S. Leorato, F. Peracchi, and A. V. Tanase, "Asymptotically efficient estimation of the conditional expected shortfall," *Computational Statistics & Data Analysis*, vol. 56, no. 4, pp. 768–784, 2012.

[23] H. Taniai, *Inference for the quantiles of ARCH processes*, Ph.D. thesis, Université libre de Bruxelles, Brussels, Belgium, 2009.