Proceedings of the International MultiConference of Engineers and Computer Scientists 2014 Vol II,
IMECS 2014, March 12 - 14, 2014, Hong Kong

Weighted Type of Quantile Regression and its Application

Xuejun Jiang, Tian Xia, and Dejun Xie

***Abstract*—In this paper we introduce a weighted composite quantile regression (CQR) estimation approach and study its application in nonlinear models such as exponential models and ARCH type of models. The weighted CQR is augmented by using a data-driven weighting scheme. With the error distribution unspecified, the proposed estimators share robustness from quantile regression and achieve nearly the same efficiency as the oracle maximum likelihood estimator(MLE) for a variety of error distributions including the normal, mixed-normal, Student's t, Cauchy distributions and etc,. We also suggest an algorithm for fast implementation of the proposed methodology. Simulations are conducted to compare the performance of different estimators, and the proposed approach is used to analyze the daily S&P 500 Composite index, which endorse our theoretical results.**

***Index Terms*—weighted CQR, Oracle MLE, Extended interior algorithm, Double threshold ARCH models.**

I. Introduction

Quantile regression (QR), introduced by Koenker and Bassett (1978), receives increasing attention in econometrics and statistics for its advantages over mean regression at least in two aspects: (i) the stochastic relationship between random variables can be portrayed much better and with much more accuracy using quantile regression than simple mean regression (see for example Chaudhuri, Doksum and Samarov, 1997); (ii) quantile regression provides more robust and consequently more efficient estimates than the mean regression when the error is non-normal (Koenker and Bassett, 1978; Koenker and Zhao, 1996). This motivates us to work along the line of QR for developing more efficient estimation methods.

The CQR in Zou and Yuan (2008) is robust compared to traditional QR. The CQR they used is a sum of different quantile regression (QR) [Koenker and Bassett (1978)] at predetermined quantiles, which uses equal weights for different QR (see Section 2 for details). Intuitively, equal weights are not optimal in general, and hence a more efficient CQR should exist. Therefore, in this article we introduce a "weighted CQR" estimation method and let the data decide the weights to improve efficiency while keeping robustness from the QR. The weighted CQR method is applicable to various models, but in this article we focus only on the nonlinear model

$$y_i = f(\mathbf{x}_i, \boldsymbol{\beta}) + \varepsilon_i, \quad i = 1, \ldots, n, \tag{I.1}$$

where ε_i's are independent random errors with unknown distribution function $G(\cdot)$ and density $g(\cdot)$, and the function $f(\cdot, \boldsymbol{\beta})$ is known up to a p-dimensional vector of parameters $\boldsymbol{\beta}$.

The work described in this paper was partially supported by NSF Grants 11101432 and 11361013.

Xuejun Jiang (jiang.xj at sustc.edu.cn) and Dejun Xie (xie.dj at sustc.edu.cn) are with South University of Science and Technology of China.

Tian Xia (xiatian718 at hotmail.com) is with University of Guizhou Finance and Economics, Guiyang, China.

We will address the issue of choosing weights when using the weighted CQR. Since the weights in the weighted CQR are allowed to be negative, the proposed weighted CQR is different from the common QR and the CQR (see also Section 2). When the weights are all equal and the model is linear with a fixed number of parameters, our method reduces to that of Zou and Yuan (2008). Since the proposed weighted CQR involves a vector of weights, we develop a data-driven weighting strategy which maximizes the efficiency of the weighted CQR estimators. The resulting estimation is adaptive in the sense that it performs asymptotically the same as if the theoretically optimal weights were used. The adaptive estimation is robust against outliers and heavy-tailed error distributions, like the Cauchy distribution, and efficient as nearly as the oracle MLE for a variety of error distributions(see Table I). This is a great advantage of the proposed estimation method, since the adaptive weighted CQR estimators do not require the form of error distribution and achieves nearly the Cramér-Rao lower bound.

The weighted CQR estimators admit no close form and involve minimizing complicate nonlinear functions, so it is challenging to derive asymptotic properties and to implement the methodology. Theoretically, we will establish asymptotic normality of the resulting estimators and show their optimality, no matter whether the error variance is finite or not. Practically, we will develop an algorithm for fast implementation of the proposed methodology. This algorithm solves a succession of linearized weighted CQR problems, each of whose dual problems is derived. We will use the "interior point algorithm" [see Jiang, X., et.cl.(2012), Vanderbei, et.cl.(1986), Koenker and Park (1996)] to solve these dual problems. The resulting algorithm is easy to implement. Simulations endorse our discovery.

The rest of the article is organized as follows. In Section II we introduce the weighted CQR for model (I.1). In section III we consider the computation method for the proposed methodology and conduct simulations to demonstrate the performance of it. In Section IV we apply the proposed methods to analyse a real dataset. Finally, Proofs for theorems are put in supplementary materials to save space.

II. Weighted composite quantile regression

Our idea can be well motivated from the linear model,

$$y_i = \mathbf{x}_i' \boldsymbol{\beta} + \varepsilon_i, \ \text{ for } i = 1, \ldots, n, \tag{II.2}$$

where $\{\varepsilon_i\}$ are i.i.d. noise with unknown distribution $G(\cdot)$ and density $g(\cdot)$.

By Koenker and Basset (1978), the τ-th QR estimate of $\boldsymbol{\beta}$ can be obtained via minimizing

$$\sum_{i=1}^{n} \rho_\tau(y_i - \mathbf{x}_i'\boldsymbol{\beta} - b_\tau)$$

over $\boldsymbol{\beta}$ and b_τ, where $\rho_\tau(u) = u(\tau - I(u < 0))$ is the check function with derivative $\psi_\tau(u) = \tau - I(u < 0)$ for $u \neq 0$. Noticing that the regression coefficients are the same across different QR models, Zou and Yuan (2008) proposed to estimate $\boldsymbol{\beta}$ by minimizing

$$\sum_{k=1}^{K}\sum_{i=1}^{n} \rho_{\tau_k}(y_i - \mathbf{x}_i'\boldsymbol{\beta} - b_{\tau_k}), \tag{II.3}$$

over $\boldsymbol{\beta}$ and b_{τ_k}, where $\{\tau_k\}_{k=1}^{K}$ are predetermined over $(0,1)$. This is the aforementioned CQR in the introduction section.

Note that the CQR method uses the same weight for different QR models. Intuitively, it is more effective if different weights are used. Applying the weighting scheme to model (I.1), one can estimate $\boldsymbol{\beta}$ by minimizing

$$L_n(\boldsymbol{\beta}, \mathbf{b}) \equiv \sum_{k=1}^{K}\omega_k \sum_{i=1}^{n} \rho_{\tau_k}(y_i - f(\mathbf{x}_i, \boldsymbol{\beta}) - b_{\tau_k}), \tag{II.4}$$

over $\boldsymbol{\beta}$ and $\mathbf{b} = (b_{\tau_1}, \ldots, b_{\tau_K})'$, where $\boldsymbol{\omega} = (\omega_1, \ldots, \omega_K)'$ is a vector of weights such that $\|\boldsymbol{\omega}\| = 1$ with $\|\cdot\|$ denoting the Euclidean norm. The weight ω_k controls the amount of contribution of the τ_k-th QR. For convenience, we denote by $\tilde{\boldsymbol{\beta}}_1$ the minimizer of $\boldsymbol{\beta}$ for (II.4) and refer to it as "the WCQR estimator". In general, given K, one can use the equally spaced quantiles at $\tau_k = k/(K+1)$ for $k = 1, 2, \ldots, K$. In practice, one can choose $K = 10$ to gain efficiency for most of situations. See Table I for details.

In order to derive the asymptotic property of the proposed estimator, in the following we introduce some notations and conditions. Let $\boldsymbol{\beta}^*$ be the true value of $\boldsymbol{\beta}$, $b_{\tau_k}^*$ be the τ_k-th quantile of ε, and $\mathbf{b}^* = (b_{\tau_1}^*, \ldots, b_{\tau_K}^*)'$. Denote by $f_i^* = f(\mathbf{x}_i, \boldsymbol{\beta}^*)$, $\boldsymbol{\nabla} f_i^* = [\partial f(\mathbf{x}_i, \boldsymbol{\beta})/\partial \boldsymbol{\beta}] \mid_{\boldsymbol{\beta}=\boldsymbol{\beta}^*}$, and $\boldsymbol{\nabla}^2 f_i^* = [\partial^2 f(\mathbf{x}_i, \boldsymbol{\beta})/\partial\boldsymbol{\beta}\partial\boldsymbol{\beta}']|_{\boldsymbol{\beta}=\boldsymbol{\beta}^*}$. Assume that

(a) $\mathbf{G} = \text{var}(\boldsymbol{\nabla} f_1^*) > 0$.

(b) The error ε_i has the distribution function $G(\cdot)$ and density function $g(\cdot)$ which is positive and continuous at the τ_k-th quantiles $b_{\tau_k}^*$.

(c) There is a large enough open subset $\Omega \in \mathbf{R}^p$ which contains the true parameter point $\boldsymbol{\beta}^*$, such that for all $\mathbf{x}_i$ the second derivative matrix $\nabla^2 f(\mathbf{x}_i, \boldsymbol{\beta})$ of $f(\mathbf{x}_i, \boldsymbol{\beta})$ with respect to $\boldsymbol{\beta}$ satisfies that

$$\|\nabla^2 f(\mathbf{x}_i, \boldsymbol{\beta}_1) - \nabla^2 f(\mathbf{x}_i, \boldsymbol{\beta}_2)\| \leq M(\mathbf{x}_i)\|\boldsymbol{\beta}_1 - \boldsymbol{\beta}_2\|$$

$$|\partial^2 f(\mathbf{x}_i, \boldsymbol{\beta})/(\partial\beta_j\partial\beta_k)| \leq N_{jk}(\mathbf{x}_i)$$

for all $\boldsymbol{\beta}_i \in \Omega$, where $E[M^2(\mathbf{x}_i)] < \infty$ and $E[N_{jk}^2(\mathbf{x}_i)] < C_1 < \infty$ for all j, k.

Under these mild conditions, we have the following asymptotic normality result.

Theorem 1: Let

$$S(\boldsymbol{\omega}) = \sum_{k,k'=1}^{K} \omega_k\omega_{k'} \min(\tau_k, \tau_{k'})(1 - \max(\tau_k, \tau_{k'})).$$

Under the conditions (a)-(c),

$$\sqrt{n}(\tilde{\boldsymbol{\beta}}_1 - \boldsymbol{\beta}^*) \xrightarrow{\mathcal{D}} \mathcal{N}\Big(\mathbf{0}, \sigma^2(\boldsymbol{\omega})\mathbf{G}^{-1}\Big),$$

where

$$\sigma^2(\boldsymbol{\omega}) = S(\boldsymbol{\omega})\{\sum_{k=1}^{K}\omega_k g(b_{\tau_k}^*)\}^{-2}$$

For model (II.2), $\mathbf{G} = \text{var}(\mathbf{x}_1)$. If all ω_k are equal, then Theorem 1 reduces to the asymptotic normality of the CQR estimators in Zou and Yuan (2008). When $K = 1$ and $\tau_1 = \tau$, it follows from the above theorem that the τ-th QR estimate of $\boldsymbol{\beta}$ is $\sqrt{n}$-consistent and asymptotically normal with mean zero and variance $g^{-2}(b_\tau^*)\tau(1-\tau)\mathbf{G}^{-1}$.

Since $\mathbf{G}$ does not involve $\boldsymbol{\omega}$, the weights should be selected to minimize $\sigma^2(\boldsymbol{\omega})$. Let $\mathbf{g} = (g(b_{\tau_1}^*), \ldots, g(b_{\tau_K}^*))'$, and let $\boldsymbol{\Omega}$ be a $K \times K$ matrix with the (k, k') element being $\Omega_{kk'} = \min(\tau_k, \tau_{k'})(1 - \max(\tau_k, \tau_{k'}))$. Then the optimal weight $\boldsymbol{\omega}_{opt}$, which minimizes $\sigma^2(\boldsymbol{\omega})$, can be shown as

$$\boldsymbol{\omega}_{opt} = (\mathbf{g}'\boldsymbol{\Omega}^{-2}\mathbf{g})^{-1/2}\boldsymbol{\Omega}^{-1}\mathbf{g}.$$

The optimal weight components can be very different, and some of them may even be negative. In fact, in our simulations we also experience such a scenario. This reflects the necessity to use a data-driven weighting scheme. The usual nonparametric density estimation methods such as kernel smoothing based on estimated residuals $\hat{\varepsilon}_i$ can provide a consistent estimation $\hat{g}(\cdot)$ of $g(\cdot)$. Let the resulting estimate of $\mathbf{g}$ be $\hat{\mathbf{g}}$. Then $\hat{\boldsymbol{\omega}} = (\hat{\mathbf{g}}'\boldsymbol{\Omega}^{-2}\hat{\mathbf{g}})^{-1/2}\boldsymbol{\Omega}^{-1}\hat{\mathbf{g}}$ provides a nonparametric estimator of $\boldsymbol{\omega}$. This leads to an adaptive estimator of $\boldsymbol{\beta}$ by minimizing

$$\sum_{k=1}^{K}\hat{\omega}_k \sum_{i=1}^{n} \rho_{\tau_k}(y_i - f(\mathbf{x}_i; \boldsymbol{\beta}) - b_{\tau_k}) \tag{II.5}$$

over b_{τ_k} and $\boldsymbol{\beta}$, where $\hat{\omega}_k$ is the k-th component of $\hat{\boldsymbol{\omega}}$. Let the resulting estimator of $\boldsymbol{\beta}$ be $\tilde{\boldsymbol{\beta}}_2$. Then $\tilde{\boldsymbol{\beta}}_2$ is asymptotically normal from the following theorem.

Theorem 2: Under the same conditions as in Theorem 1,

$$\sqrt{n}(\tilde{\boldsymbol{\beta}}_2 - \boldsymbol{\beta}^*) \xrightarrow{\mathcal{D}} \mathcal{N}\Big(\mathbf{0}, (\mathbf{g}'\boldsymbol{\Omega}^{-1}\mathbf{g})^{-1}\mathbf{G}^{-1}\Big).$$

Since $\sigma^2(\boldsymbol{\omega}_{opt}) = (\mathbf{g}'\boldsymbol{\Omega}^{-1}\mathbf{g})^{-1}$, $\tilde{\boldsymbol{\beta}}_2$ has the same asymptotic variance matrix as $\tilde{\boldsymbol{\beta}}_1$ as if $\boldsymbol{\omega}_{opt}$ were known. That is, the estimator $\tilde{\boldsymbol{\beta}}_2$ is adaptive. Therefore, $\hat{\boldsymbol{\omega}}$ is called the adaptive weight vector. By Theorem 2, the asymptotic relative efficiency (ARE) of the adaptive WCQR estimation with respect to the least squares (LS) estimation is $e(WCQR, LS) = \sigma^2\mathbf{g}'\boldsymbol{\Omega}^{-1}\mathbf{g}$. Since the oracle maximum likelihood (ML) estimator of $\boldsymbol{\beta}$ has asymptotic variance matrix $I_g^{-1}\mathbf{G}^{-1}$, where $I_g = \int[g'(t)]^2/g(t)\,dt$ is the Fisher information. Therefore, the relative efficiency of the adaptive WCQR estimation with respect to the oracle maximum likelihood estimation (MLE) is

$$e(WCQR, ML) = I_g^{-1}\mathbf{g}'\boldsymbol{\Omega}^{-1}\mathbf{g}.$$

For equally spaced $\{\tau_k\}_{k=1}^{K}$, the adaptive estimator $\tilde{\boldsymbol{\beta}}_2$ is nearly efficient as the oracle MLEs for various error distributions (see Theorem 4 in Jiang,X., et.al. (2012)), which is a great advantage of the proposed methodology.

For each K, the AREs of the adaptive estimator $\tilde{\boldsymbol{\beta}}_2$ with respect to some common estimators can be calculated. To appreciate how much efficiency is gained in practice, we investigate the performance of common estimators. Table I reports those AREs for linear models with various error distributions, which demonstrates that $\tilde{\boldsymbol{\beta}}_2$ is highly efficient for all the distributions under consideration. For linear models, Leng (2009) demonstrated that his regularized rank regression estimator (R^2) was quite efficient and robust. Table I indicates that the proposed adaptive estimate dominates the

ISBN: 978-988-19253-3-6
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

TABLE I
THE RELATIVE EFFICIENCY OF ESTIMATORS. LAD- LEAST ABSOLUTE DEVIATION.

	K	$e(R^2)$	$e(ML)$	$e(LS)$	$e(LAD)$
Normal	10	1.009	0.964	0.964	1.514
	100	1.045	0.998	0.998	1.567
	1000	1.047	1.000	1.000	1.571
Mixed Normal	10	1.003	0.961	1.378	1.380
	100	1.041	0.998	1.430	1.432
	1000	1.044	1.000	1.434	1.436
t_3	10	1.036	0.984	1.967	1.214
	100	1.052	0.999	1.998	1.233
	1000	1.053	1.000	2.000	1.234
$\chi^2(6)$	10	1.387	0.585	1.755	2.913
	100	1.904	0.803	2.410	4.001
	1000	2.154	0.909	2.726	4.525
Cauchy	10	1.601	0.973	inf	1.201
	100	1.644	1.000	inf	1.233
	1000	1.645	1.000	inf	1.234

R^2 for all error distributions and is much more efficient than it when the error follows the Cauchy or chi-squared distribution. It also suggests that typically one could choose $K = 10$ in practice and efficiency is much gained, as shown in simulations.

III. NUMERICAL IMPLEMENTATION

A. Extended interior algorithm

Minimization in (II.4) involves a complicate nonlinear optimization problem. We use the interior point algorithm (see Jiang, X., et. al. (2012) to solve the problem. Matlab codes are available upon request for the proposed methods.

Consider the equivalent problem of (II.4):

$$\min_{\boldsymbol{\theta}} \sum_{k=1}^{K} \omega_k \sum_{i=1}^{n} \rho_{\tau_k}(y_i - l_{ik}(\boldsymbol{\theta})), \tag{III.6}$$

where $l_{ik}(\boldsymbol{\theta}) = f(\mathbf{x}_i, \boldsymbol{\beta}) + b_{\tau_k}$, and $\boldsymbol{\theta} = (b_{\tau_1}, \dots, b_{\tau_K}, \boldsymbol{\beta}')'$. Following Osborne and Watson (1971), we solve the problem (III.6) using the following algorithm:

(1) Given the current value, $\theta^{(j)}$, of θ, calculate t to minimize

$$\sum_{k=1}^{K} \omega_k \sum_{i=1}^{n} \rho_{\tau_k}\{y_i - l_{ik}(\theta^{(j)}) - \nabla l_{ik}(\theta^{(j)})t\}, \tag{III.7}$$

where $\nabla l_{ik}(\theta^{(j)}) = (dl_{ik}(\theta)/d\theta')|_{\theta=\theta^{(j)}}$. Let the minimum be $s^{(j)}$, and let the minimizer be $t = t^{(j)}$.

(2) Calculate λ to minimize

$$\sum_{k=1}^{K} \omega_k \sum_{i=1}^{n} \rho_{\tau_k}\{y_i - l_{ik}(\theta^{(j)} + \lambda t^{(j)})\}. \tag{III.8}$$

Let the minimum be $\bar{s}^{(j+1)}$ with $\lambda = \lambda^{(j)}$.

(3) Set $\theta^{(j+1)} = \theta^{(j)} + \lambda^{(j)} t^{(j)}$. Update the current value of θ by $\theta^{(j+1)}$, and repeat the above procedure until the new iterate fails to improve the objective function by a specified tolerance such as 10^{-4}.

In the above method, the problem (III.8) can be easily solved by line search in the resulting direction $t = t^{(j)}$, but one has to solve a succession of linearized weighted quantile regression problems in (III.7). Let $y_{ik}^* = y_i - l_{ik}(\theta^{(j)})$ and $x_{ik}' = \nabla l_{ik}(\theta^{(j)})$. Then the problem (III.7) becomes

$$\min_{t} \sum_{k=1}^{K} \omega_k \sum_{i=1}^{n} \rho_{\tau_k}(y_{ik}^* - x_{ik}' t), \tag{III.9}$$

whose dual problem can be shown as

$$\max\{\tilde{y}'d \,|\, d = ({d_1}', \dots, {d_K}')',$$

$$d_k \in [\omega_k(\tau_k - 1), \omega_k \tau_k]^n, \mathbf{X}'d = 0\}, \tag{III.10}$$

where $\tilde{y} = ({y_1^*}' \dots, {y_K^*}')'$, $\mathbf{X} = (X_1', \dots, X_n')'$, $X_k = (x_{1k}, \dots, x_{nk})'$, and $y_k^* = (y_{1k}, \dots, y_{nk})'$. The interior point method can be used to solve the dual problem (III.10), the algorithm details are listed as follows:

1. For any initial feasible point d, e.g., $d = 0$, following Meketon (1986), set $D_k = \text{diag}(\min\{\omega_k\tau_k - d_{ik}, d_{ik} - \omega_k(\tau_k - 1)\})$, $D = \text{diag}(D_1, \dots, D_K)$, $s = D^2(I - X(X'D^2X)^{-1}X'D^2)\tilde{y}$, and $t = (X'D^2X)^{-1}X'D^2\tilde{y}$.
2. Set $d_k^* = d_k + (\eta/\gamma_k)s_k$, where $\gamma_k = \max_i(\max\{s_{ik}/(\omega_k\tau_k - d_{ik}), -s_{ik}/(d_{ik} - \omega_k(\tau_k - 1))\})$, and $\eta \in (0, 1)$ is the constant chosen to ensure feasibility. As suggested by Koenker and Park, $\eta = 0.97$.
3. Set $d = d^*$. Updating D, s, and the new d^* continues the iteration.

After solving (III.10) using the above interior point algorithm, we arrive at the next loop which uses the current value $\theta = \theta^{(j+1)}$ for the primal problem in (III.9). This leads to the updated dual problem (III.10) with $y_{ik}^* = y_i - f_{ik}(\theta^{(j+1)})$ and $x_{ik}' = \nabla f_{ik}(\theta^{(j+1)})$. The current d should be adjusted to ensure that it is feasible for the new value of X. Similar to Koenker and Park (1996), we project the current d onto the null space of the new $\mathbf{X}$, i.e. $\hat{d} = (I - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}')d$ and then shrinking it to insure that d_k lies in $[\omega_k(\tau_k - 1), \omega_k\tau_k]^n$, so the adjusted d becomes

$$d_k = \Big(\max_i\Big\{\max\Big(\frac{\hat{d}_{ik}}{\omega_k(\tau_k - 1)}, \frac{\hat{d}_{ik}}{\omega_k\tau_k}\Big) + \delta\Big\}\Big)^{-1}\hat{d}_k$$

for some tolerance parameter $\delta > 0$.

B. Simulations

In this section, we report on simulations to investigate the advantages of the WCQR estimation. An exponential regression model was used:

$$y = 1 + b\exp(\mathbf{c}'\mathbf{x}) + \varepsilon,$$

where b and $\mathbf{c} = (c_1, c_2, c_3)'$ are parameters, ε is the error. The true values of parameters were set as $b = 1.5$, and $\mathbf{c} = (-0.6, -0.8, -0.7)'$.

We draw from the working model 400 samples of sizes 200 and 400. In each simulation, the components of $\mathbf{x}$ were jointly normal distributed with the pairwise correlation coefficient 0.5 and standard normal as marginals. We considered three sets of errors: $N(0, 1)$, $t(4)$, and $\chi^2(4)$. All of them were centralized and scaled so that the medians of the absolute errors were ones. As before, we used $K = 10$ and equally spaced τ_k on $(0, 1)$.

We compared four estimation methods: the L_1, CQR, WCQR estimation, and the oracle MLE (maximum likelihood estimation) of known error distribution. In each simulation the "root of mean squared errors (RMSE)" for different coefficient estimators were calculated, and their average over simulations is reported in Tables II-IV, where Σ denotes the sum of RMSE for all components in $\boldsymbol{\beta}$. Therefore, better estimators should have smaller Σ values. As expected, the

ISBN: 978-988-19253-3-6
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

TABLE II
RMSE(MULTIPLIED BY 10^3) FOR DIFFERENT QR ESTIMATORS UNDER THE NORMAL ERROR.

	$n = 400$				
Estimates	$\hat{b}$	$\hat{c}_1$	$\hat{c}_2$	$\hat{c}_3$	Σ
L_1	145	47	33	29	255
CQR	125	39	28	25	217
WCQR	124	38	28	24	214
OML	121	37	28	24	210

TABLE III
RMSE(MULTIPLIED BY 10^3) FOR DIFFERENT QR ESTIMATORS UNDER THE STUDENT T ERROR.

	$n = 400$				
Estimates	$\hat{b}$	$\hat{c}_1$	$\hat{c}_2$	$\hat{c}_3$	Σ
L_1	144	42	34	31	250
CQR	127	37	30	28	222
WCQR	126	37	30	27	221
OML	126	37	29	27	219

oracle MLE performs the best, the WCQR performs better than the CQR and L_1, and the L_1 is the worst. In terms of overall performance (the value of Σ), WCQR estimators uniformly dominate the L_1 and CQR estimators and are comparable to the oracle MLE under all kind of errors we considered. This is in accordance with our previous theoretical results.

IV. A REAL EXAMPLE

We study the daily S&P 500 Composite index from January 3, 2000 to July 27, 2011. This index represents the bulk of the daily value in the US equity market. The return series y_t is defined as the difference of the log-price. We are interested in the asymmetry of the conditional mean and conditional variance. The double-threshold ARCH (DTARCH) model [see Li and Li (1996), Hui, Y.V. and Jiang, J. (2005), Jiang, J., et.al. (2013)] can be used to describe this kind of asymmetry.

The proposed estimation approaches are applied to the data set with 2908 observations. We set the threshold parameter $r = 0$ and the delay parameter $d = 1$ in DTARCH models, which is consistent with observations in the stock market. Theorefore, We consider the following DTARCH model for the return series y_t:

$$y_t = \begin{cases} \alpha_1^{(1)} y_{t-1} + \varepsilon_t, & \text{if } y_{t-1} \le 0, \\ \alpha_1^{(2)} y_{t-1} + \varepsilon_t, & \text{if } y_{t-1} > 0, \end{cases}$$

where $\varepsilon_t = h_t u_t$, with

$$h_t = \begin{cases} \beta_0^{(1)} + \beta_1^{(1)} |\varepsilon_{t-1}|, & \text{if } y_{t-1} \le 0, \\ \beta_0^{(2)} + \beta_1^{(2)} |\varepsilon_{t-1}|, & \text{if } y_{t-1} > 0. \end{cases}$$

For comparison, we apply the WCQR estimate with equal weights ω_k, labeled as “CQR estimate”, and the QR estimates

TABLE IV
RMSE(MULTIPLIED BY 10^3) FOR DIFFERENT QR ESTIMATORS UNDER CHI-SQURE ERROR.

	$n = 400$				
Estimates	$\hat{b}$	$\hat{c}_1$	$\hat{c}_2$	$\hat{c}_3$	Σ
L_1	147	42	34	30	254
CQR	112	33	27	24	196
WCQR	86	26	21	19	151
OML	62	22	16	15	115

TABLE V
ESTIMATES OF PARAMETERS WITH ESTIMATED STANDARD ERRORS IN PARENTHESES (MULTIPLIED BY 10^2).

Method	$QR(K = 1, \tau_K = 0.75)$	CQR	WCQR
$\alpha_1^{(1)}$	-11.30 (2.13)	-11.60 (3.03)	-11.02 (2.09)
$\alpha_1^{(2)}$	-4.86 (2.57)	-4.64 (2.96)	-5.08 (2.39)
$\beta_0^{(1)}$	0.50 (0.06)	0.50 (0.03)	0.51 (0.03)
$\beta_1^{(1)}$	19.04 (4.81)	19.20 (2.42)	19.31 (2.42)
$\beta_0^{(2)}$	0.51 (0.06)	0.51 (0.03)	0.51 (0.03)
$\beta_1^{(2)}$	6.39 (5.33)	6.36 (2.67)	6.42 (2.67)

(i.e. $K = 1$) at $\tau_K = 0.5$ and 0.75 to fit the model. We calculate the estimated parameters and their standard errors. The results are reported in Tables V. For $\tau_K = 0.5$, it corresponds to the L_1-estimate. Since the result from L_1-estimation is poor, we omit it there.

It is observed that all the estimation approaches identify positive beta’s values. This is desired because the volatility coefficients, β’s, are nonnegative. Negative $\alpha_1^{(1)}$ and $\alpha_1^{(2)}$ indicate that the future mean return will be forecasted as positive (negative) if the current return is negative (positive). This is expected in an efficient market. On the other hand, negative returns in this model have about 3 times the effect of positive returns on future conditional scales. This indicates that the volatility is significantly higher when prices are falling. That is, volatility tends to be higher in bear markets, an asymmetric volatility effect described by Nelson’s EGARCH model (Nelson (1991)).

Among the three estimation methods, the WCQR estimator is the best for all estimators because it has the smallest standard deviation.

ACKNOWLEDGMENT

This research was supported by NSF Grant 11101432 and NSF Grant 11361013. It was also supported in part by the Fundamental Research Funds for the Central Universities of China (No: 31541111215) for Xuejun Jiang at Zhongnan University of Economics and Law, Wuhan, China. Correspondence should be addressed to Xuejun, Jiang, Department of Financial Mathematics and Financial Engineering, South University of Science and Technology, Shenzhen, China, E-mail: jiang.xj@sustc.edu.cn.

REFERENCES

[1] G. W. Bassett and R. W. Koenker. (1992). A Note on Recent Proposals for Computing L_1 Estimates. *Computational Statistics & Data Analysis*. **14**, 207-211.
[2] P. Billingsley (1995). *Probability and Measure*, 3rd ed. New York: Wiley.
[3] P. Chaudhuri, K. Doksum and A. Samarov (1997). On average derivative quantile regression. *Annals of Statistics*, **25**, 715-744.
[4] R. A. Davis and W. T. M. Dunsmuir (1997). Least absolute deviation estimation for regression with ARMA errors. *J. Theor. Prob.* **10**, 481-497.
[5] Y.V. Hui and J. Jiang, J. (2005). “Robust modeling of DTARCH models.” Econometrics Journal **8**: 143-158.
[6] J. Jiang, X. Jiang, X. and X. Song (2013). Weighted composite quantile regression of DTARCH models. The Econometrics Journal, To appear.
[7] X. Jiang, J. Jiang and X. Song (2012) Oracle model selection for nonlinear models based on weighted composite quantile regression. Statistica Sinica, 22(4), 1479-1506.
[8] K. Knight (1998). Limiting distributions for l_1 regression estimators under general conditions. *Ann. Statist.* **26**, 755-770.

[9] R. Koenker and G. Bassett (1978). Regression quantiles. *Econometrica.* **46**, 33-50.
[10] R. Koenker and B. J. Park (1996). An interior point algorithm for nonlinear quantile regression. *Journal of Econometrics.* **71**, 265-283.
[11] R. Koenkerb and Q. Zhao (1996). Conditional quantile estimation and inference for ARCH models. *Econometric Theory*, **12**, 793-813.
[12] Li, C. W. and Li, W. K. (1996). On a double-threshold autoregressive heteroscedastic time series model. *J. Appl. Econometrics*, **11**, 253–274.
[13] Leng, C. (2009). Variable selection and coefficient estimation via regularized rank regression. *Statistica Sinica*, to appear.
[14] Nelson, D. (1991). Conditional heteroscedasticity in asset returns: A new approach. *Econometrica*, **59**, 347-370.
[15] H. Zou and M. Yuan (2008). Composite quantile regression and the oracle model selection theory. *Ann. Statist.* **3**, 1108-1126.

ISBN: 978-988-19253-3-6
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)