Journal of Statistical Computation and Simulation

ISSN: 0094-9655 (Print) 1563-5163 (Online) Journal homepage: http://www.tandfonline.com/loi/gscs20

Adaptive composite quantile regressions and their asymptotic relative efficiency

Ke Yang, Liping Zhu & Wangli Xu

JOURNAL OF STATISTICAL COMPUTATION AND SIMULATION, 2017
https://doi.org/10.1080/00949655.2017.1408112

Adaptive composite quantile regressions and their asymptotic relative efficiency

Ke Yang[a], Liping Zhu[b] and Wangli Xu[a]

[a]Center for Applied Statistics, School of Statistics, Renmin University of China, Beijing, People's Republic of China; [b]Institute of Statistics and Big Data, Renmin University of China, Beijing, People's Republic of China

ABSTRACT

The composite quantile regression (CQR for short) provides an efficient and robust estimation for regression coefficients. In this paper we introduce two adaptive CQR methods. We make two contributions to the quantile regression literature. The first is that, both adaptive estimators treat the quantile levels as realizations of a random variable. This is quite different from the classic CQR in which the quantile levels are typically equally spaced, or generally, are treated as fixed values. Because the asymptotic variances of the adaptive estimators depend upon the generic distribution of the quantile levels, it has the potential to enhance estimation efficiency of the classic CQR. We compare the asymptotic variance of the estimator obtained by the CQR with that obtained by quantile regressions at each single quantile level. The second contribution is that, in terms of relative efficiency, the two adaptive estimators can be asymptotically equivalent to the CQR method as long as we choose the generic distribution of the quantile levels properly. This observation is useful in that it allows to perform parallel distributed computing when the computational complexity issue arises for the CQR method. We compare the relative efficiency of the adaptive methods with respect to some existing approaches through comprehensive simulations and an application to a real-world problem.

ARTICLE HISTORY
Received 4 July 2017
Accepted 18 November 2017

KEYWORDS
Adaptiveness; composite quantile regression; quantile regression; robustness; relative efficiency

1. Introduction

Let $\mathbf{x} = (X_1, \dots, X_p)^{\mathrm{T}} \in \mathbb{R}^p$ be a p-vector of covariates and $Y \in \mathbb{R}^1$ be a univariate response variable. We consider estimation of $\boldsymbol{\beta} \in \mathbb{R}^p$ for linear regression

$$Y = \mathbf{x}^{\mathrm{T}}\boldsymbol{\beta} + \varepsilon, \tag{1}$$

where ε is an independent error. Without loss of generality we assume that both the covariates and the response are centred to have zero mean so that model (1) does not necessarily contain an intercept. We assume that ε in model (1) has the cumulative distribution function $F(\cdot)$ and the density function $f(\cdot)$. Ideally, if the error distribution is known, one can estimate $\boldsymbol{\beta}$ with the maximum likelihood estimation (MLE for short) method. In reality,

CONTACT Wangli Xu xwlbnu@163.com, wxu.stat@gmail.com Center for Applied Statistics School of Statistics, Renmin University of China, Beijing 100872, People's Republic of China

however, the error distribution is quite often unknown. When we are lack of information about the error distribution or density function, we may have to turn to the least squares estimation (LSE for short) method. The LSE is perhaps the first and one of the most popular ways to estimate $\boldsymbol{\beta}$. For the LSE to be valid, the error term is assumed implicitly to have a finite variance. If the error variance σ^2 is infinite, $\widehat{\boldsymbol{\beta}}_{\mathrm{LSE}}$ is no longer root-n consistent and can not serve as an ideal method for coefficient estimation. Such situations occur, for example, when the error term follows student-t distribution with one or two degrees of freedom.

In the seminal work of Koenker and Bassett [1] quantile regression was proposed which can be used to estimate the slope vector $\boldsymbol{\beta}$ without moment restrictions on ε. To be precise, let $\rho_{\tau_k}(u) = u\{\tau_k - I(u < 0)\}$ for $\tau_k \in (0, 1)$. Define

$$(\widehat{b}_{\tau_k}, \widehat{\boldsymbol{\beta}}_{\tau_k}) \stackrel{\text{def}}{=} \arg\min_{b,\boldsymbol{\beta}} \sum_{i=1}^{n} \{\rho_{\tau_k}(Y_i - b - \mathbf{x}_i^{\mathrm{T}}\boldsymbol{\beta})\}. \tag{2}$$

Because ε is independent of $\mathbf{x}$, $\widehat{\boldsymbol{\beta}}_{\tau_k}$ converges in probability to $\boldsymbol{\beta}$ for all $\tau_k \in (0, 1)$. This motivates Koenker [2] to propose the following two weighted estimators. The first estimator, denoted by $\widehat{\boldsymbol{\beta}}_{W,1}$, is simply defined as

$$\widehat{\boldsymbol{\beta}}_{W,1} \stackrel{\text{def}}{=} \sum_{k=1}^{K} w_k \widehat{\boldsymbol{\beta}}_{\tau_k}, \tag{3}$$

where $\tau_1, \tau_2, \ldots, \tau_K$ are K distinct quantile levels and the weights w_ks satisfy

$$\sum_{k=1}^{K} w_k = 1.$$

We assume $\boldsymbol{\Sigma} \stackrel{\text{def}}{=} \mathrm{var}(\mathbf{x})$ is positive definite. Under certain regular conditions, Koenker [2] and Zhao and Lian [3] show that

$$n^{1/2}(\widehat{\boldsymbol{\beta}}_{W,1} - \boldsymbol{\beta}) \xrightarrow{d} \mathcal{N}(0, \boldsymbol{\Sigma}^{-1} W_{1,K}), \tag{4}$$

where $\xrightarrow{d}$ stands for 'convergence in distribution' and

$$W_{1,K} \stackrel{\text{def}}{=} \sum_{k=1}^{K} \sum_{k'=1}^{K} w_k w_{k'} (\min(\tau_k, \tau_k')\{1 - \max(\tau_k, \tau_k')\}/[f\{F^{-1}(\tau_k) f\{F^{-1}(\tau_k')\}]). \tag{5}$$

The second estimator, denoted by $\widehat{\boldsymbol{\beta}}_{W,2}$, is defined as

$$(\widehat{b}_{\tau_1}, \ldots, \widehat{b}_{\tau_K}, \widehat{\boldsymbol{\beta}}_{W,2}) \stackrel{\text{def}}{=} \arg \min_{b_1,\ldots,b_K,\boldsymbol{\beta}} \sum_{k=1}^{K} \sum_{i=1}^{n} w_k \{\rho_{\tau_k}(Y_i - b_k - \mathbf{x}_i^{\mathrm{T}}\boldsymbol{\beta})\}, \tag{6}$$

Under certain regular conditions, Koenker [2] and Zhao and Lian [3] show that

$$n^{1/2}(\widehat{\boldsymbol{\beta}}_{W,2} - \boldsymbol{\beta}) \xrightarrow{d} \mathcal{N}(0, \boldsymbol{\Sigma}^{-1} W_{2,K}), \tag{7}$$

where

$$W_{2,K} \stackrel{\text{def}}{=} \sum_{k=1}^{K}\sum_{k'=1}^{K} w_k w_{k'} \min(\tau_k, \tau_k')\{1-\max(\tau_k, \tau_k')\} \Bigg/ \left[\sum_{k=1}^{K} w_k f\{F^{-1}(\tau_k)\}\right]^2. \tag{8}$$

Koenker [2, p.324] and Zhao and Lian [3, Proposition 1 in page 1337] showed that, in terms of their asymptotic variances, $\widehat{\boldsymbol{\beta}}_{W,1}$ and $\widehat{\boldsymbol{\beta}}_{W,2}$ are equivalent if we choose the weights w_ks delicately. Both the weights w_ks and the quantile levels τ_ks can be user-specified. Zou and Yuan [4] set $\tau_k = k/(K+1)$ and left the weights w_ks unspecified. In this case the optimization problem (6) corresponds to the composite quantile regression (CQR). Noting that the CQR method has the potential to be asymptotically more efficient than the LSE method Zou and Yuan [4], Jiang et al. [5] considered variable selection with the weighted CQR method (6) for high dimensional parametric models, and [6] adapted the weighted CQR methods (3) and (6) to fully nonparametric models. If one opts to set $\tau_k = k/(K+1)$ and $w_k = 1/K$, the CQR method can be adapted to linear, fully nonparametric, varying coefficient, single index and partially linear additive models, respectively. See, for example, Zou and Yuan [4], Kai et al. [7,8], Jiang et al. [9] and Guo et al. [10]. In all these existing works, the quantile levels τ_ks are treated as fixed values in $(0, 1)$.

In this paper we revisit the asymptotic relative efficiency issue of the CQR method from a quite different perspective: We simply set $w_k = 1/K$ and treat the quantile levels τ_ks as realizations of a random variable which has support over (0,1), whereas the CQR method uses equally spaced quantile levels. We draw K quantile levels independently from a generic distribution. Denote these randomly drawn quantile levels by $\tau_1, \tau_2, \ldots, \tau_K$. We estimate $\boldsymbol{\beta}$ with Equation (2) at each quantile level τ_k and define

$$\widehat{\boldsymbol{\beta}}_{\text{AQR}} \stackrel{\text{def}}{=} K^{-1}\sum_{k=1}^{K}\widehat{\boldsymbol{\beta}}_{\tau_k} \tag{9}$$

as the adaptive quantile regression estimator. We refer to the above procedure as the adaptive quantile regression (AQR for short) method. At these quantile levels, we may also implement the CQR method:

$$(\widehat{b}_{\tau_1}, \ldots, \widehat{b}_{\tau_K}, \widehat{\boldsymbol{\beta}}_{\text{ACQR}}) \stackrel{\text{def}}{=} \arg\min_{b_1,\ldots,b_K,\boldsymbol{\beta}} \sum_{k=1}^{K}\sum_{i=1}^{n}\{\rho_{\tau_k}(Y_i - b_k - \mathbf{x}_i^{\mathrm{T}}\boldsymbol{\beta})\}. \tag{10}$$

We refer to the above procedure as the adaptive composite quantile regression (ACQR for short) method. Fix $w_k = 1/K$, for $k = 1, \ldots, K$. Then the distinctions between the weighted and the adaptive estimates are that, in the weighted estimates, $\widehat{\boldsymbol{\beta}}_{W,1}$ and $\widehat{\boldsymbol{\beta}}_{W,2}$, the quantile levels are treated as fixed values, where in the adaptive estimates, $\widehat{\boldsymbol{\beta}}_{\text{AQR}}$ and $\widehat{\boldsymbol{\beta}}_{\text{ACQR}}$, the quantile levels are randomly drawn from a generic distribution.

We study the asymptotic relative efficiency issue of the above two adaptive estimators under different situations. To ease subsequent illustration, we assume that the quantile level τ is a random variable with density function $g(\cdot)$ which has support on the interval $(0, 1)$. We make the following interesting observations.

(1) If we draw τ from uniform distribution, namely, $g(\tau) \equiv 1$, and the error variance is finite, then the AQR method is asymptotically equivalent to the LSE method.

(2) If we choose $g(\tau)$ to be a function proportional to $f(b_\tau)$, that is, $g(\tau) \propto f\{F^{-1}(\tau)\}$, then the AQR method is asymptotically equivalent to the CQR method.
(3) If we draw τ from uniform distribution, namely, $g(\tau) \equiv 1$, the ACQR method is asymptotically equivalent to the CQR method.
(4) For continuous density function $g(\tau)$, if we set $w_k = K^{-1}g\{k/(K+1)\}$ in weighted CQR with $K \to \infty$, then the AQR method is asymptotically equivalent to $\widehat{\boldsymbol{\beta}}_{W,1}$ estimator and the ACQR method is asymptotically equivalent to $\widehat{\boldsymbol{\beta}}_{W,2}$ estimator.

Because the efficiency of the resultant adaptive estimators depends upon the generic distribution of this random variable, the adaptive estimators have the potential to be more efficient than existing competitors such as the LSE and the CQR method.

This paper is organized as follows. In Section 2, we study the asymptotic properties of the AQR and the ACQR methods thoroughly. We also compare their asymptotic relative efficiency with the LSE and the CQR methods. In Section 3, we illustrate the relative efficiency of the AQR and the ACQR methods through comprehensive simulations and an application to a real world problem. We conclude this paper with a brief discussion in Section 4. All technical proofs are given in the appendix.

2. The adaptive quantile regressions

2.1. Literature review

The LSE is perhaps the first and one of the most popular methods to estimate the regression coefficients in linear models. Suppose $\{(\mathbf{x}_i, Y_i), i = 1, \ldots, n\}$ are n independent copies of $(\mathbf{x}, Y)$. We write $\mathbf{X} \overset{\text{def}}{=} (\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n)^{\mathrm{T}} \in \mathbb{R}^{n\times p}$ and $\mathbf{y} \overset{\text{def}}{=} (Y_1, Y_2, \ldots, Y_n)^{\mathrm{T}} \in \mathbb{R}^{n\times 1}$. The LSE in model (1) has the form of $\widehat{\boldsymbol{\beta}}_{\mathrm{LSE}} = (\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}(\mathbf{X}^{\mathrm{T}}\mathbf{y})$. If $\sigma^2 \overset{\text{def}}{=} \mathrm{var}(\varepsilon) < \infty$, $\widehat{\boldsymbol{\beta}}_{\mathrm{LSE}}$ is root-n consistent and asymptotically normal. To be precise,

$$n^{1/2}(\widehat{\boldsymbol{\beta}}_{\mathrm{LSE}} - \boldsymbol{\beta}) \xrightarrow{d} \mathcal{N}(0, \boldsymbol{\Sigma}^{-1}\sigma^2), \tag{11}$$

where $\xrightarrow{d}$ stands for 'convergence in distribution'. Apparently, the LSE method is no longer root-n consistent if the error variance is infinite.

In model (1), for all $\tau \in (0,1)$, the conditional $100 \times \tau\%$ quantile function of $(Y \mid \mathbf{x})$ is $(b_\tau + \mathbf{x}^{\mathrm{T}}\boldsymbol{\beta})$, where b_τ is the $100 \times \tau\%$ quantile of ε. Zou and Yuan [4] proposed a CQR method to estimate $\boldsymbol{\beta}$. The CQR method does not require σ^2 to be finite, thus overcomes the pitfall of the LSE method. Denote the CQR estimator of $\boldsymbol{\beta}$ by $\widehat{\boldsymbol{\beta}}_{\mathrm{CQR}}$. Under certain regularity conditions, Zou and Yuan [4] showed that

$$n^{1/2}(\widehat{\boldsymbol{\beta}}_{\mathrm{CQR}} - \boldsymbol{\beta}) \xrightarrow{d} \mathcal{N}(0, \boldsymbol{\Sigma}^{-1}D_K), \tag{12}$$

where

$$D_K \overset{\text{def}}{=} \sum_{k=1}^{K}\sum_{k'=1}^{K} \min(\tau_k, \tau_{k'}')\{1 - \max(\tau_k, \tau_{k'}')\} \Bigg/ \left[\sum_{k=1}^{K} f\{F^{-1}(\tau_k)\}\right]^2,$$

and $f(\cdot)$ and $F(\cdot)$ stand, respectively, for the density and the distribution functions of ε. In particular, the CQR method reduces to the classical quantile regression (QR for short) method when $K = 1$ [11]. We denote $\widehat{\boldsymbol{\beta}}_{\mathrm{QR}}$ the classical QR estimator. As a special case of (12), the QR estimator satisfies the following asymptotic normality under mild conditions:

$$n^{1/2}(\widehat{\boldsymbol{\beta}}_{\mathrm{QR}} - \boldsymbol{\beta}) \xrightarrow{d} \mathcal{N}(0, \boldsymbol{\Sigma}^{-1}D_1), \tag{13}$$

where $D_1 \stackrel{\text{def}}{=} \tau_1(1-\tau_1)/\{f(b_{\tau_1})\}^2$. For the classical QR method, we specify the quantile level τ_1 of interest. For the classical CQR method, Zou and Yuan [4] suggest to use equally spaced quantile levels $\tau_k = k/(K+1)$, for $k = 1, \ldots, K$ and K is a user-specified number. In practice, we prefer to choose large K values to attain asymptotic efficiency as long as the CQR method is computationally feasible. Let $M \stackrel{\text{def}}{=} E\{f(\varepsilon)\}$. If the number of quantile levels are sufficiently large,

$$D_\infty \stackrel{\text{def}}{=} \lim_{K\to\infty} D_K = 12^{-1}M^{-2}.$$

Zou and Yuan [4] gave a few examples to show that D_∞ could be smaller than the error variance σ^2 even when σ^2 is finite. In other words, the CQR method has the potential to be more efficient than the LSE method.

2.2. The adaptive estimates

The adaptive estimators treat the quantile levels as realizations of a random variable. We assume the quantile level τ is a random variable with density function $g(\cdot)$ which has support on the interval (0,1). We draw $\tau_1, \tau_2, \ldots, \tau_K$, independently, from $g(\cdot)$, then estimate $\boldsymbol{\beta}$ with either the QR or the CQR method at these randomly generated quantile levels. Denote the resultant estimators by $\widehat{\boldsymbol{\beta}}_{\mathrm{AQR}}$ and $\widehat{\boldsymbol{\beta}}_{\mathrm{ACQR}}$, respectively, which are given in Equations (9) and (10).

Before establishing the asymptotic properties of $\widehat{\boldsymbol{\beta}}_{\mathrm{AQR}}$ and $\widehat{\boldsymbol{\beta}}_{\mathrm{ACQR}}$, we first introduce the following regularity conditions.

(A1) The covariance matrix of $\mathbf{x}$, denoted $\boldsymbol{\Sigma} \stackrel{\text{def}}{=} \mathrm{var}(\mathbf{x})$, is positive definite.
(A2) For all $\tau \in (0,1), f\{F^{-1}(\tau)\}$ is uniformly bounded away from zero and infinity. Here $F(\cdot)$ and $f(\cdot)$ are cumulative distribution function and density function for ε in model (1).
(A3) Assume that, for all p-vector $\mathbf{u}$,

$$\lim_{n\to\infty}\left[n^{-1}\sum_{i=1}^{n}\int_0^{u_0+\mathbf{x}_i^{\mathrm{T}}\mathbf{u}} n^{1/2}\{F(a+t/n^{1/2}) - F(a)\}\,\mathrm{d}t\right]$$
$$= f(a)(u_0, \mathbf{u}^{\mathrm{T}})\begin{pmatrix} 1 & 0 \\ 0 & \boldsymbol{\Sigma}\end{pmatrix}(u_0, \mathbf{u}^{\mathrm{T}})^{\mathrm{T}}/2.$$

Conditions (A1) and (A2) are commonly assumed to derive the asymptotic normality of the QR estimator. See, for example, Koenker [11]. Conditions (A1) and (A3) are used to derive the asymptotic normality of the CQR estimator [4]. Because our proposed adaptive estimators are built upon the QR and CQR methods, we also assume these conditions to establish the desired asymptotic properties.

Define

$$B \stackrel{\text{def}}{=} 2\left[\int_0^1 \int_0^{\tau_2} \frac{\tau_1(1-\tau_2)}{f\{F^{-1}(\tau_1)\}f\{F^{-1}(\tau_2)\}} g(\tau_1)g(\tau_2)\,\mathrm{d}\tau_1\,\mathrm{d}\tau_2\right].$$

The asymptotic property of $\widehat{\boldsymbol{\beta}}_{\mathrm{AQR}}$ is given as follows.

Theorem 2.1: *Assume conditions* (A1) *and* (A2) *and the density function* $g(\cdot)$ *has support over* $(0, 1)$. *As n diverges to infinity followed by K,*

$$n^{1/2}(\widehat{\boldsymbol{\beta}}_{\mathrm{AQR}} - \boldsymbol{\beta}) \xrightarrow{d} \mathcal{N}(0, \boldsymbol{\Sigma}^{-1}B). \tag{14}$$

The following corollary states that, the AQR method is asymptotically equivalent to the LSE method if we draw τ from uniform distribution and the error variance is finite; if we choose $g(\tau)$ to be a function proportional to $f\{F^{-1}(\tau)\}$, the AQR method is asymptotically equivalent to the CQR method.

Corollary 2.2:

1. *If* $g(\tau) \equiv 1$ *and* $\sigma^2 < \infty$, *then* $B = \sigma^2$.
2. *If* $g(\tau) \propto f\{F^{-1}(\tau)\}$, *then* $B = D_\infty$.

We can also choose $g(\cdot)$ to be some other functions such that the AQR method is more efficient than both the CQR and the LSE methods. At least, $\widehat{\boldsymbol{\beta}}_{\mathrm{AQR}}$ has the potential to be asymptotically equivalent to $\widehat{\boldsymbol{\beta}}_{\mathrm{LSE}}$ and $\widehat{\boldsymbol{\beta}}_{\mathrm{CQR}}$, as long as we choose the generic density function $g(\cdot)$ properly. This surprising observation is very useful in that it allows us to perform parallel distributed computing when the computational complexity issue arises for the CQR method. When the computational issue arises, we can simply estimate $\boldsymbol{\beta}$ separately with quantile regressions at different quantile levels and average these estimates to produce an asymptotically equivalent one.

Define

$$A \stackrel{\text{def}}{=} 2\left\{\int_0^1 \int_0^{\tau_2} \tau_1(1-\tau_2)g(\tau_1)g(\tau_2)\,\mathrm{d}\tau_1\,\mathrm{d}\tau_2\right\} \Big/ \left\{\int_0^1 f(b_\tau)g(\tau)\,\mathrm{d}\tau\right\}^2.$$

The asymptotic normality of $\widehat{\boldsymbol{\beta}}_{\mathrm{ACQR}}$ is given as follows.

Theorem 2.3: *Assume conditions* (A1) *and* (A3) *and the density function* $g(\cdot)$ *has support over* $(0, 1)$. *As n diverges to infinity followed by K,*

$$n^{1/2}(\widehat{\boldsymbol{\beta}}_{\mathrm{ACQR}} - \boldsymbol{\beta}) \xrightarrow{d} \mathcal{N}(0, \boldsymbol{\Sigma}^{-1}A). \tag{15}$$

The asymptotic variance of $\widehat{\boldsymbol{\beta}}_{\mathrm{ACQR}}$ depends on $g(\cdot)$. If we choose $g(\cdot)$ properly, the ACQR method has the potential to be more efficient than the CQR method. The following corollary states that, if we draw sufficiently many realizations from uniform distribution and treat these realizations as the quantile levels, the ACQR method can be asymptotically equivalent to the CQR method.

Corollary 2.4:

1. *If* $g(\tau) \equiv 1$, *then* $A = D_\infty$.

Corollary 2.2 and 2.4 present the relationship between our adaptive methods and classical LSE and CQR estimators. Besides, our proposed estimators also connect to weighting schemes proposed by Koenker [2].

Corollary 2.5:

(1) *For continuous density function* $g(\tau)$, *if we set* $w_k = K^{-1}g\{k/(K+1)\}$ *to derive* $\boldsymbol{\beta}_{W,1}$ *defined in Equation* (3) *with* $K \to \infty$, *then* $W_{1,\infty} = B$. *Here* $W_{1,\infty}$ *is* $W_{1,K}$ *defined in Equation* (5) *with* $K = \infty$.
(2) *For continuous density function* $g(\tau)$, *if we set* $w_k = K^{-1}g\{k/(K+1)\}$ *to derive* $\boldsymbol{\beta}_{W,2}$ *defined in Equation* (6) *with* $K \to \infty$, *then* $W_{2,\infty} = A$. *Here* $W_{2,\infty}$ *is* $W_{2,K}$ *defined in Equation* (8) *with* $K = \infty$.

We define the optimal choice of $g(\cdot)$ as the density function which minimizes either B defined in Equation (14) or A defined in Equation (15). However, how to find such an optimal $g(\cdot)$ is not straightforward in general. In the following section, we will give a few simulated examples to compare their asymptotic relative efficiency. In view of Equation (13), we anticipate that $\widehat{\boldsymbol{\beta}}_{\mathrm{QR}}$ has smaller asymptotic variance if τ is closer to 0.5. In general, if τ is close to 0 or 1, $\widehat{\boldsymbol{\beta}}_{\mathrm{QR}}$ is not very stable and tends to have large asymptotic variance. If this is the case, we can choose $g(\tau)$ such that $g(\tau)$ attains its maximum when $\tau = 0.5$ and decays when τ is away from 0.5. This strategy often helps us to reduce the asymptotic variances of $\widehat{\boldsymbol{\beta}}_{\mathrm{AQR}}$ and $\widehat{\boldsymbol{\beta}}_{\mathrm{ACQR}}$.

3. Numerical studies

3.1. Simulations

In this section, we compare the relative efficiency of our proposed ACQR and AQR methods with the LSE, CQR and QR methods. In model (1), we vary n from 100 to 500 and set $p=6$ and $\boldsymbol{\beta} = (1, 0.8, 0.6, 0.4, 0.2, 0)^{\mathrm{T}}$. Let $\boldsymbol{\Sigma} = (\sigma_{k,l})_{p\times p}$ and $\sigma_{k,l} = 0.5^{|k-l|}$, for $k, l = 1, \ldots, p$. Define $\mathbf{x}_i \stackrel{\text{def}}{=} \boldsymbol{\Sigma}^{1/2}\mathbf{z}_i$, $i = 1, \ldots, n$, where $\mathbf{z}_i$s are drawn independently from uniform distribution with support over $[-3, 3]$. Following Zou and Yuan [4], we consider the following five error distributions.

(E1) The error term ε follows standard normal distribution.
(E2) The error term ε follows Laplace distribution with location parameter 0 and scale parameter 1.
(E3) The error term ε follows logistic distribution with location parameter 0 and scale parameter 1.

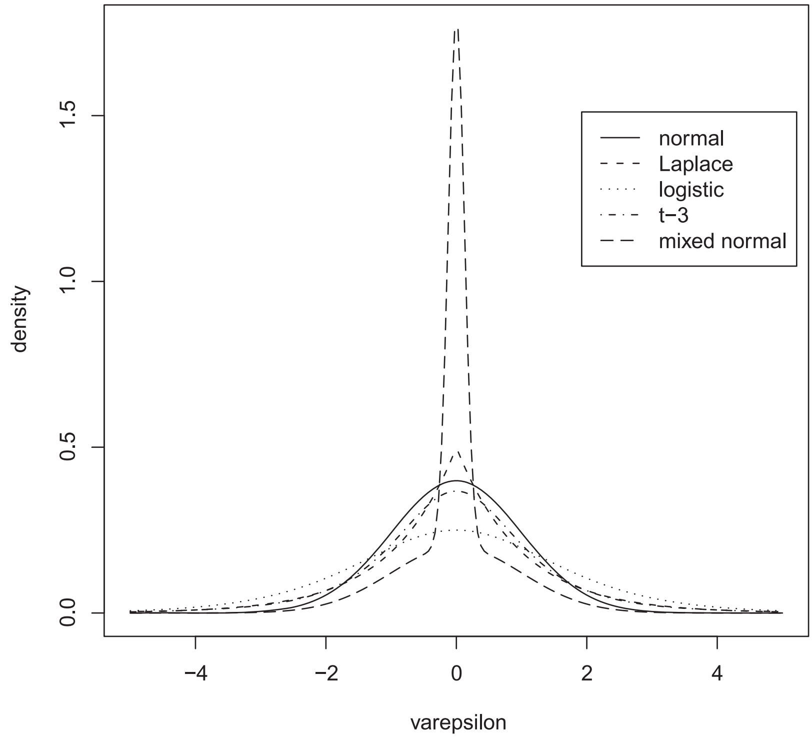

Figure 1. The error density functions of ε.

(E4) The error term ε follows student-t distribution with three degrees of freedom.

(E5) The error term ε follows a mixture of normal distribution $0.5N(0,1) + 0.5N(0,0.125^2)$.

We present these density functions in Figure 1, which facilitates us to understand the shapes and the kurtosis of these density functions intuitively. All these densities are symmetric and the Laplace and the mixture normal distributions have the largest kurtosis.

For the QR method, we only consider $\tau_1 = 0.5$. For the CQR method, we use equally spaced quantile levels for a fixed K. To implement both the AQR and the ACQR methods, we must specify the generic distribution of the quantile level τ. We consider the following ten scenarios.

(G1) We draw τ from Beta$(1,1)$, or equivalently, $g(\tau) \equiv 1$. Denote the resultant estimators by $\widehat{\boldsymbol{\beta}}_{\text{ACQR},1}$ and $\widehat{\boldsymbol{\beta}}_{\text{AQR},1}$, respectively.

(G2) We draw τ from Beta$(2,2)$. Denote the resultant estimators by $\widehat{\boldsymbol{\beta}}_{\text{ACQR},2}$ and $\widehat{\boldsymbol{\beta}}_{\text{AQR},2}$, respectively.

(G3) We draw τ from Beta$(5,5)$. Denote the resultant estimators by $\widehat{\boldsymbol{\beta}}_{\text{ACQR},3}$ and $\widehat{\boldsymbol{\beta}}_{\text{AQR},3}$, respectively.

(G4) We draw τ from Beta$(1,1.5)$. Denote the resultant estimators by $\widehat{\boldsymbol{\beta}}_{\text{ACQR},4}$ and $\widehat{\boldsymbol{\beta}}_{\text{AQR},4}$, respectively.

(G5) We draw τ from Beta$(1, 5)$. Denote the resultant estimators by $\widehat{\boldsymbol{\beta}}_{\text{ACQR},5}$ and $\widehat{\boldsymbol{\beta}}_{\text{AQR},5}$, respectively.

(G6) We draw τ from Beta$(1.5, 1)$. Denote the resultant estimators by $\widehat{\boldsymbol{\beta}}_{\text{ACQR},6}$ and $\widehat{\boldsymbol{\beta}}_{\text{AQR},6}$, respectively.

(G7) We draw τ from Beta$(5, 1)$. Denote the resultant estimators by $\widehat{\boldsymbol{\beta}}_{\text{ACQR},7}$ and $\widehat{\boldsymbol{\beta}}_{\text{AQR},7}$, respectively.

(G8) We draw τ from the density function $g(\tau) = 0.5\pi \sin(\pi\tau)$. Denote the resultant estimators by $\widehat{\boldsymbol{\beta}}_{\text{ACQR},8}$ and $\widehat{\boldsymbol{\beta}}_{\text{AQR},8}$, respectively.

(G9) We draw τ from the density function $g(\tau) = 0.5 + 3\tau - 3\tau^2$. Denote the resultant estimators by $\widehat{\boldsymbol{\beta}}_{\text{ACQR},9}$ and $\widehat{\boldsymbol{\beta}}_{\text{AQR},9}$, respectively.

(G10) We draw τ from the density function $g(\tau) \propto f\{F^{-1}(\tau)\}$. Denote the resultant estimators by $\widehat{\boldsymbol{\beta}}_{\text{ACQR},10}$ and $\widehat{\boldsymbol{\beta}}_{\text{AQR},10}$, respectively. They are used as benchmarks for comparison. We also use this scenario to illustrate the theoretical properties in Corollary (2.2). In practice, the error distribution is usually unknown and has to be estimated from observations.

In the first three scenarios (G1)–(G3), the density functions given in Figure2(A) are all symmetric. In the fourth through the seventh scenarios (G4)–(G7), the density functions given in Figure2(B) are all skewed. In the eighth and the ninth scenarios (G8)-(G9), the density functions given in Figure2(C) are convex and symmetric. In the last scenario (G10), the distributions of the quantile level depend upon the five error distributions specified in (E1)–(E5), which are not shown here for brevity.

Comparing the asymptotic relative efficiency of the QR and the CQR methods with their adaptive versions requires specifying a proper K. Zou and Yuan [4] suggested $K = 19$ suffices for many purposes for the CQR method. In our simulations we choose $K = 19$ when $n = 100$ and $K = 99$ when $n = 500$ for the CQR, the AQR and the ACQR methods. We use such a large K when $n = 500$ to minimize the effect of K on the asymptotic variances of the CQR and the ACQR estimators.

We repeat each experiment 400 times and report the average of the biases and the variances of all estimators. The simulation results are charted in Tables 1–5, corresponding to (E1)–(E5), respectively.

It can be clearly seen that, the averages of the biases of all estimates are trivially small, especially when $n = 500$, indicating that all estimators are asymptotically consistent. The

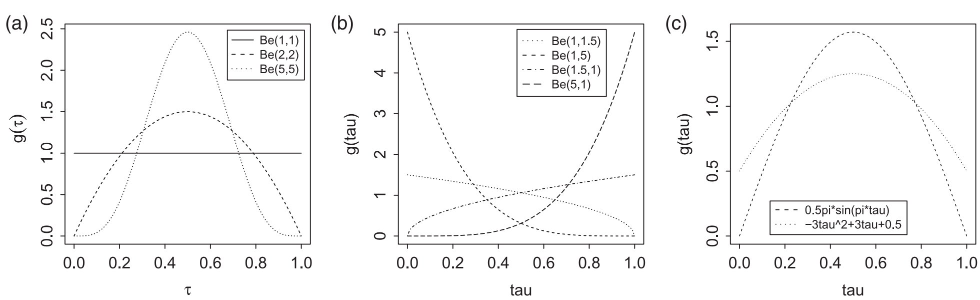

Figure 2. (A): The density functions of τ for scenarios (G1)–(G3). (B): The density functions of τ for scenarios (G4)–(G7). (C): The density functions of τ for scenarios (G8)–(G9).

Table 1. The average of the biases and the variance (in the parentheses) of different estimates when ε follows standard normal distribution.

	β_1	β_2	β_3	β_4	β_5	β_6
			$n = 100, K = 19$			
LSE	−0.53 (3.55)	−3.93 (5.81)	7.47 (5.53)	1.38 (6.48)	−5.28 (6.85)	−2.40 (5.62)
QR	−1.51 (5.61)	0.63 (9.06)	8.73 (8.41)	−4.05 (9.09)	−7.63 (8.84)	−0.33 (8.85)
CQR	0.23 (3.72)	−2.84 (6.23)	8.66 (5.87)	1.53 (6.76)	−6.21 (6.90)	−3.00 (6.03)
ACQR,1	0.28 (3.84)	−3.15 (6.45)	7.36 (6.15)	1.73 (7.17)	−6.39 (7.28)	−2.57 (6.27)
ACQR,2	1.07 (4.08)	−3.62 (6.83)	9.66 (6.46)	0.11 (7.06)	−6.18 (7.22)	−3.34 (6.76)
ACQR,3	−0.18 (4.25)	−2.11 (7.25)	10.35 (6.89)	−0.21 (7.41)	−7.84 (7.65)	−2.84 (7.10)
ACQR,4	2.18 (4.20)	−4.85 (6.94)	8.53 (6.34)	1.88 (7.36)	−5.63 (7.36)	−4.17 (6.26)
ACQR,5	4.97 (5.49)	−6.18 (9.10)	6.89 (9.46)	1.11 (10.72)	−4.28 (9.80)	−8.58 (8.71)
ACQR,6	−1.90 (4.02)	−1.21 (6.64)	8.70 (6.20)	2.48 (7.45)	−6.94 (7.53)	−1.05 (6.48)
ACQR,7	−2.63 (5.10)	−0.65 (9.38)	7.75 (8.26)	4.52 (10.73)	−7.56 (11.08)	1.86 (8.81)
ACQR,8	−0.44 (4.04)	−2.31 (6.61)	10.16 (6.21)	−0.24 (7.08)	−5.69 (7.18)	−3.49 (6.69)
ACQR,9	1.67 (3.86)	−3.64 (6.72)	8.14 (6.35)	1.85 (7.29)	−5.69 (7.27)	−3.99 (6.23)
ACQR,10	2.80 (3.77)	−1.04 (6.43)	−1.06 (5.41)	−3.09 (7.28)	−0.72 (6.39)	−5.29 (6.58)
AQR,1	0.77 (3.95)	−3.44 (6.81)	6.54 (6.10)	−0.55 (7.65)	−5.19 (7.41)	−2.44 (6.60)
AQR,2	0.73 (4.09)	−2.73 (7.07)	7.41 (6.57)	−0.91 (7.39)	−6.50 (7.38)	−3.39 (6.52)
AQR,3	−0.99 (4.30)	−1.55 (7.12)	10.20 (6.75)	−1.16 (7.53)	−7.26 (7.71)	−2.70 (6.94)
AQR,4	2.34 (4.64)	−3.90 (7.42)	5.81 (6.68)	0.40 (7.92)	−3.22 (7.65)	−6.22 (6.85)
AQR,5	5.91 (7.68)	−6.33 (11.74)	3.36 (12.34)	0.11 (14.17)	−3.20 (12.14)	−10.06 (12.01)
AQR,6	−2.31 (4.26)	−1.17 (7.18)	7.74 (6.45)	0.14 (8.11)	−6.78 (7.92)	1.39 (6.96)
AQR,7	−3.01 (6.62)	−3.83 (11.98)	7.68 (10.93)	1.86 (13.22)	−3.08 (12.80)	5.74 (11.64)
AQR,8	−0.37 (4.00)	−2.06 (6.46)	9.18 (6.30)	−1.76 (7.42)	−5.35 (7.35)	−3.77 (6.65)
AQR,9	1.22 (4.02)	−4.28 (6.71)	7.35 (6.55)	−0.07 (7.33)	−4.65 (7.27)	−5.24 (6.63)
AQR,10	0.16 (3.93)	−2.70 (6.81)	7.18 (6.38)	0.08 (7.29)	−5.98 (7.34)	−2.89 (6.47)
			$n = 500, K = 99$			
LSE	−0.20 (0.69)	1.61 (1.11)	−1.11 (1.10)	1.78 (1.01)	1.00 (1.12)	−0.02 (1.21)
QR	−0.05 (1.02)	2.84 (1.63)	−3.43 (1.76)	3.49 (1.61)	0.10 (1.72)	0.05 (1.79)
CQR	0.25 (0.72)	1.56 (1.11)	−1.17 (1.17)	2.32 (1.04)	0.62 (1.18)	−0.07 (1.27)
ACQR,1	0.22 (0.72)	1.76 (1.13)	−1.06 (1.19)	2.20 (1.05)	0.43 (1.19)	−0.03 (1.30)
ACQR,2	0.45 (0.76)	1.72 (1.16)	−1.43 (1.24)	2.60 (1.09)	0.14 (1.25)	0.11 (1.36)
ACQR,3	0.40 (0.82)	1.85 (1.24)	−1.52 (1.37)	2.73 (1.20)	0.22 (1.36)	−0.27 (1.49)
ACQR,4	0.57 (0.74)	1.55 (1.18)	−1.47 (1.23)	2.12 (1.11)	1.00 (1.22)	−0.13 (1.34)
ACQR,5	0.67 (0.99)	−0.40 (1.64)	0.03 (1.60)	1.28 (1.62)	2.48 (1.61)	0.39 (1.85)
ACQR,6	0.01 (0.74)	1.87 (1.14)	−1.29 (1.19)	2.55 (1.05)	0.34 (1.22)	−0.38 (1.28)
ACQR,7	−0.58 (1.00)	3.08 (1.57)	−1.45 (1.45)	2.15 (1.42)	−0.29 (1.64)	−0.44 (1.69)
ACQR,8	0.57 (0.76)	1.61 (1.15)	−1.44 (1.25)	2.63 (1.10)	0.25 (1.27)	0.10 (1.38)
ACQR,9	0.24 (0.74)	1.70 (1.13)	−1.24 (1.22)	2.42 (1.08)	0.38 (1.22)	−0.13 (1.32)
ACQR,10	−0.63 (0.79)	−0.20 (1.29)	−2.99 (1.33)	2.14 (1.19)	−0.95 (1.24)	1.68 (1.22)
AQR,1	−0.36 (0.72)	2.04 (1.18)	−0.82 (1.18)	2.05 (1.04)	0.74 (1.15)	0.47 (1.26)
AQR,2	0.21 (0.73)	1.72 (1.15)	−1.16 (1.20)	2.48 (1.06)	0.10 (1.22)	0.44 (1.30)
AQR,3	0.40 (0.80)	1.65 (1.25)	−1.49 (1.36)	2.96 (1.18)	0.36 (1.33)	−0.21 (1.45)
AQR,4	0.28 (0.74)	1.28 (1.24)	−1.34 (1.23)	1.65 (1.13)	1.54 (1.24)	0.23 (1.36)
AQR,5	−0.46 (1.31)	−0.19 (2.19)	−0.41 (2.04)	0.90 (2.07)	3.17 (2.14)	0.79 (2.42)
AQR,6	−0.51 (0.76)	2.24 (1.22)	−1.51 (1.22)	2.15 (1.06)	0.70 (1.20)	−0.18 (1.29)
AQR,7	−2.24 (1.29)	4.46 (2.19)	−2.16 (2.00)	0.91 (1.99)	0.70 (2.04)	−0.53 (2.21)
AQR,8	0.33 (0.74)	1.59 (1.14)	−1.17 (1.21)	2.64 (1.07)	0.35 (1.21)	0.50 (1.32)
AQR,9	−0.30 (0.72)	1.79 (1.14)	−1.01 (1.17)	2.08 (1.02)	0.80 (1.16)	0.05 (1.25)
AQR,10	0.25 (0.72)	1.68 (1.14)	−1.08 (1.17)	2.61 (1.04)	0.31 (1.19)	0.35 (1.29)

Note: All numbers reported below are multiplied by 1000.

biases of many estimators are smaller when n increases. For example, in Table 2 when ε follows Laplace distribution, the bias of $\beta_{3,\mathrm{LSE}}$ is 5.95×10^{-3} when $n = 100$ and is 0.57×10^{-3} when $n = 500$. In addition, the variances of all estimators when $n = 100$ are also roughly five times larger than the variances when $n = 500$. For example, in Table 2, the variance of $\beta_{3,\mathrm{LSE}}$ is 11.54×10^{-3} when $n = 100$ and is 2.35×10^{-3} when $n = 500$; In Table 4 when ε

Table 2. The average of the biases and the variance (in the parentheses) of different estimates when ε follows Laplace distribution with location parameter 0 and scale parameter 1.

	β_1	β_2	β_3	β_4	β_5	β_6
			$n=100, K=19$			
LSE	10.73 (7.26)	−10.68 (12.24)	5.95 (11.54)	−3.94 (11.92)	0.71 (11.86)	1.15 (10.84)
QR	10.00 (5.10)	−11.20 (9.76)	4.67 (8.96)	−5.12 (8.30)	3.71 (7.76)	−2.31 (7.84)
CQR	9.71 (5.08)	−9.42 (9.06)	7.90 (8.29)	−5.15 (8.19)	3.43 (8.23)	−2.35 (7.62)
ACQR,1	9.16 (5.52)	−8.06 (9.86)	6.46 (8.98)	−4.07 (8.43)	1.75 (8.55)	−1.21 (7.91)
ACQR,2	10.34 (4.88)	−8.82 (9.01)	8.07 (8.25)	−5.72 (8.08)	3.98 (7.97)	−2.96 (7.36)
ACQR,3	8.46 (4.66)	−8.48 (8.52)	8.19 (7.98)	−6.28 (7.67)	4.89 (7.46)	−3.38 (7.02)
ACQR,4	10.14 (5.28)	−10.40 (9.43)	8.58 (8.75)	−3.92 (8.81)	1.29 (9.30)	−1.98 (8.17)
ACQR,5	14.97 (9.78)	−11.61 (18.30)	4.26 (15.74)	−4.22 (18.56)	−0.91 (17.86)	−0.07 (15.74)
ACQR,6	9.69 (5.92)	−10.53 (9.31)	10.08 (9.31)	−5.59 (8.79)	3.41 (8.66)	−0.36 (8.18)
ACQR,7	4.81 (12.00)	−7.09 (16.99)	6.24 (15.80)	−6.30 (15.77)	10.18 (14.53)	−2.58 (15.80)
ACQR,8	8.94 (4.85)	−9.03 (8.58)	7.29 (8.46)	−4.82 (7.80)	4.16 (8.12)	−2.85 (7.77)
ACQR,9	8.40 (5.17)	−8.81 (8.93)	6.91 (8.27)	−4.79 (8.24)	2.76 (8.21)	−1.67 (7.78)
ACQR,10	−0.63 (5.42)	−2.36 (8.59)	−4.61 (8.24)	4.00 (8.62)	10.16 (8.27)	−2.69 (8.10)
AQR,1	11.63 (7.27)	−11.36 (13.04)	4.70 (12.29)	−4.70 (12.60)	2.15 (11.58)	2.03 (11.61)
AQR,2	10.99 (5.43)	−9.65 (10.10)	8.50 (9.15)	−6.14 (8.87)	3.81 (9.26)	−3.80 (7.92)
AQR,3	8.79 (4.80)	−8.21 (8.75)	7.94 (7.84)	−5.37 (7.85)	5.45 (7.67)	−3.46 (7.24)
AQR,4	13.07 (6.98)	−11.98 (13.08)	7.65 (11.59)	−2.76 (12.89)	−0.80 (13.48)	1.57 (11.68)
AQR,5	14.76 (18.01)	−8.07 (32.25)	−0.15 (30.21)	3.07 (35.26)	−9.21 (31.09)	10.34 (31.99)
AQR,6	10.01 (8.56)	−12.75 (12.39)	7.81 (13.11)	−5.10 (12.65)	7.18 (11.94)	−1.81 (11.53)
AQR,7	5.87 (21.58)	−11.95 (31.03)	4.76 (31.21)	−6.37 (32.05)	10.99 (27.72)	−0.10 (29.34)
AQR,8	8.89 (5.25)	−9.62 (9.86)	8.39 (9.16)	−5.84 (8.78)	6.56 (8.74)	−3.62 (8.34)
AQR,9	9.35 (6.57)	−10.25 (10.89)	6.50 (9.80)	−5.61 (10.65)	5.41 (10.33)	−0.88 (9.88)
AQR,10	9.95 (5.36)	−8.56 (9.26)	6.61 (8.66)	−4.67 (8.86)	2.37 (8.66)	−2.33 (8.08)
			$n=500, K=99$			
LSE	−1.18 (1.37)	−1.09 (2.22)	0.57 (2.35)	−1.50 (2.37)	−0.91 (2.49)	−0.34 (2.42)
QR	−0.99 (0.81)	−1.77 (1.36)	1.63 (1.39)	−0.81 (1.46)	0.38 (1.47)	−0.61 (1.46)
CQR	−1.43 (0.92)	−0.96 (1.55)	−0.17 (1.60)	−0.58 (1.59)	0.12 (1.60)	−1.06 (1.68)
ACQR,1	−1.53 (0.94)	−0.78 (1.58)	−0.51 (1.59)	−0.57 (1.60)	0.12 (1.60)	−0.88 (1.71)
ACQR,2	−1.42 (0.88)	−0.88 (1.49)	−0.15 (1.50)	−0.25 (1.49)	0.24 (1.49)	−1.20 (1.61)
ACQR,3	−1.29 (0.83)	−0.87 (1.38)	0.23 (1.36)	−0.47 (1.39)	0.29 (1.41)	−0.91 (1.52)
ACQR,4	−1.27 (0.95)	−1.33 (1.58)	0.08 (1.61)	−0.33 (1.62)	−0.11 (1.66)	−0.62 (1.69)
ACQR,5	−0.83 (1.67)	−1.91 (2.77)	1.56 (2.70)	−0.57 (2.66)	−1.94 (3.10)	0.12 (2.83)
ACQR,6	−1.67 (0.93)	−0.93 (1.60)	−0.31 (1.64)	−0.89 (1.62)	0.77 (1.57)	−1.23 (1.74)
ACQR,7	−3.06 (1.68)	1.33 (3.00)	−2.05 (2.95)	−1.95 (2.88)	1.60 (2.98)	−2.47 (3.06)
ACQR,8	−1.36 (0.86)	−1.02 (1.45)	−0.30 (1.45)	−0.37 (1.47)	0.27 (1.48)	−1.23 (1.59)
ACQR,9	−1.47 (0.90)	−1.00 (1.51)	−0.19 (1.55)	−0.40 (1.54)	−0.03 (1.56)	−0.90 (1.64)
ACQR,10	0.96 (0.89)	−0.27 (1.36)	−0.91 (1.33)	−0.85 (1.39)	2.79 (1.48)	−2.90 (1.35)
AQR,1	−1.95 (1.44)	0.26 (2.24)	0.61 (2.32)	−2.72 (2.30)	−0.35 (2.43)	−0.91 (2.41)
AQR,2	−1.62 (1.00)	−0.57 (1.67)	0.04 (1.69)	−0.53 (1.70)	−0.33 (1.71)	−1.34 (1.82)
AQR,3	−1.47 (0.85)	−0.71 (1.44)	0.31 (1.42)	−0.61 (1.45)	0.12 (1.45)	−1.11 (1.57)
AQR,4	−1.63 (1.45)	0.33 (2.17)	0.91 (2.28)	−0.27 (2.35)	−2.89 (2.58)	0.31 (2.42)
AQR,5	−0.72 (4.05)	3.41 (5.96)	2.38 (6.08)	1.47 (6.29)	−7.58 (7.33)	2.67 (6.55)
AQR,6	−2.26 (1.45)	−0.65 (2.47)	0.33 (2.53)	−3.25 (2.46)	0.55 (2.50)	−1.26 (2.66)
AQR,7	−3.46 (4.27)	−0.60 (7.13)	1.05 (7.37)	−8.65 (6.96)	3.27 (7.72)	−2.73 (7.25)
AQR,8	−1.71 (0.94)	−0.76 (1.60)	−0.03 (1.63)	−0.82 (1.63)	−0.12 (1.65)	−1.44 (1.78)
AQR,9	−2.09 (1.16)	−0.14 (1.85)	0.28 (1.98)	−1.11 (2.01)	−0.75 (2.05)	−0.82 (2.09)
AQR,10	−1.76 (0.91)	−0.47 (1.54)	0.14 (1.59)	−0.90 (1.59)	0.45 (1.63)	−1.32 (1.72)

Note: All numbers reported below are multiplied by 1000.

follows student-t distribution with three degrees of freedom, the variance of $\beta_{1,\mathrm{ACQR},1}$ is 5.30×10^{-3} when $n=100$ and is 1.02×10^{-3} when $n=500$.

Recall that, for $\widehat{\boldsymbol{\beta}}_{\mathrm{AQR},1}$, we treat the quantile levels as realizations of a uniform random variable. Our theoretical analyses indicate that the asymptotic variance of $\widehat{\boldsymbol{\beta}}_{\mathrm{AQR},1}$ is expected to be very close to that of $\widehat{\boldsymbol{\beta}}_{\mathrm{LSE}}$. The simulation results are in line with our theoretical investigations. For example, in Table 1 when ε follows normal distribution and

Table 3. The average of the biases and the variance (in the parentheses) of different estimates when ε follows logistic distribution with location parameter 0 and scale parameter 1.

	β_1	β_2	β_3	β_4	β_5	β_6
			$n=100, K=19$			
LSE	−2.72 (11.65)	−0.09 (18.23)	−4.03 (18.39)	11.22 (18.98)	−0.28 (20.08)	0.57 (20.09)
QR	−6.35 (14.76)	6.51 (22.11)	−6.44 (21.48)	8.84 (23.66)	−1.08 (24.03)	0.82 (26.21)
CQR	−2.54 (10.84)	1.61 (16.71)	−4.65 (16.63)	10.32 (17.40)	1.12 (17.96)	−1.34 (18.46)
ACQR,1	−3.09 (11.29)	0.45 (18.34)	−4.23 (16.94)	11.11 (17.98)	0.21 (18.83)	0.70 (19.00)
ACQR,2	−3.04 (11.56)	1.14 (17.17)	−5.66 (17.12)	9.70 (18.80)	2.87 (19.69)	−1.54 (19.64)
ACQR,3	−3.29 (12.26)	3.58 (18.26)	−5.28 (17.41)	8.51 (19.77)	2.96 (20.02)	−1.65 (20.28)
ACQR,4	−3.12 (11.07)	3.37 (18.23)	−5.17 (16.88)	11.58 (19.17)	2.92 (19.22)	−1.43 (19.73)
ACQR,5	−2.20 (16.60)	3.82 (30.77)	−8.02 (26.41)	14.13 (29.37)	−3.07 (29.33)	0.54 (28.71)
ACQR,6	−2.37 (11.75)	−1.42 (18.18)	−1.74 (18.82)	9.81 (18.05)	−0.32 (19.41)	0.59 (19.75)
ACQR,7	−1.24 (18.55)	−1.22 (27.45)	−2.15 (29.71)	5.84 (26.41)	2.68 (27.82)	2.09 (29.54)
ACQR,8	−5.47 (11.42)	5.44 (17.73)	−2.90 (16.89)	8.83 (18.49)	−0.21 (18.62)	−1.02 (19.13)
ACQR,9	−3.02 (11.03)	−0.37 (17.38)	−4.28 (17.69)	9.93 (18.88)	2.70 (19.18)	−3.33 (19.55)
ACQR,10	11.09 (12.76)	−0.83 (17.16)	−0.30 (18.93)	6.65 (17.06)	11.00 (17.38)	−7.22 (19.03)
AQR,1	−4.35 (13.16)	3.47 (21.20)	−2.42 (20.25)	9.21 (20.37)	1.70 (22.40)	2.28 (20.95)
AQR,2	−5.71 (11.52)	5.44 (17.22)	−5.22 (17.43)	8.99 (19.66)	1.85 (20.72)	−1.57 (19.34)
AQR,3	−3.51 (12.01)	5.02 (17.67)	−5.06 (17.55)	8.17 (19.36)	2.98 (20.06)	−1.69 (20.10)
AQR,4	−3.30 (12.60)	7.83 (21.51)	−7.94 (19.57)	9.86 (22.25)	3.38 (22.85)	0.46 (21.96)
AQR,5	−1.72 (26.64)	10.80 (48.45)	−14.33 (42.84)	15.56 (45.60)	−1.51 (43.03)	2.86 (45.24)
AQR,6	−4.81 (14.38)	0.72 (20.66)	−0.96 (21.94)	8.08 (20.22)	3.07 (22.95)	0.99 (22.48)
AQR,7	−1.87 (26.65)	−2.51 (40.00)	0.17 (42.65)	5.11 (38.45)	9.90 (42.70)	−2.30 (46.23)
AQR,8	−5.98 (11.21)	5.68 (17.29)	−1.22 (17.55)	7.81 (19.31)	0.36 (19.26)	−0.57 (19.07)
AQR,9	−3.44 (11.71)	1.49 (19.09)	−6.35 (19.42)	9.55 (20.17)	2.61 (21.81)	−2.88 (20.99)
AQR,10	−3.32 (11.73)	3.38 (16.71)	−4.51 (17.88)	9.01 (18.87)	1.88 (20.62)	−0.66 (19.87)
			$n=500, K=99$			
LSE	1.05 (2.08)	−4.18 (3.61)	0.68 (3.57)	−1.75 (3.70)	−0.22 (3.86)	−0.57 (3.65)
QR	2.55 (2.67)	−7.11 (5.02)	2.09 (4.46)	0.41 (4.32)	−1.19 (4.50)	0.51 (4.27)
CQR	1.83 (1.99)	−5.49 (3.51)	−0.33 (3.22)	0.07 (3.24)	0.17 (3.47)	−0.23 (3.27)
ACQR,1	2.05 (2.01)	−5.59 (3.51)	−0.21 (3.22)	0.39 (3.27)	0.01 (3.52)	−0.07 (3.33)
ACQR,2	1.78 (2.06)	−5.85 (3.66)	−0.50 (3.26)	0.52 (3.28)	0.19 (3.59)	−0.45 (3.33)
ACQR,3	1.76 (2.14)	−5.75 (3.86)	−0.65 (3.33)	0.45 (3.31)	0.21 (3.69)	−0.10 (3.42)
ACQR,4	1.84 (2.07)	−5.51 (3.63)	−0.20 (3.23)	0.76 (3.23)	−0.61 (3.60)	−0.12 (3.38)
ACQR,5	1.20 (3.00)	−4.44 (5.04)	0.93 (4.92)	1.95 (4.67)	−3.80 (5.16)	−0.76 (4.84)
ACQR,6	1.49 (2.06)	−5.08 (3.65)	−0.72 (3.39)	−0.34 (3.40)	1.53 (3.53)	−0.28 (3.38)
ACQR,7	0.87 (2.82)	−3.51 (4.91)	−0.84 (4.96)	−3.81 (5.52)	4.72 (5.18)	−0.11 (5.10)
ACQR,8	1.89 (2.06)	−6.03 (3.68)	−0.06 (3.26)	0.47 (3.24)	0.40 (3.56)	−0.57 (3.36)
ACQR,9	2.29 (2.01)	−6.08 (3.63)	−0.45 (3.27)	0.44 (3.27)	0.06 (3.49)	−0.27 (3.32)
ACQR,10	1.69 (1.96)	0.04 (3.62)	−0.01 (3.33)	0.89 (3.44)	−0.43 (3.45)	−2.03 (3.62)
AQR,1	1.11 (2.13)	−4.87 (3.57)	1.28 (3.66)	−1.46 (3.85)	−0.41 (3.88)	−0.53 (3.62)
AQR,2	1.26 (2.08)	−5.34 (3.67)	0.08 (3.33)	−0.16 (3.40)	0.45 (3.58)	−0.15 (3.29)
AQR,3	1.71 (2.13)	−5.87 (3.89)	−0.67 (3.39)	0.61 (3.35)	0.59 (3.66)	−0.15 (3.40)
AQR,4	−0.39 (2.29)	−4.04 (4.02)	0.72 (3.93)	0.34 (3.77)	−1.26 (4.04)	−0.46 (3.98)
AQR,5	−2.20 (4.81)	−1.62 (8.30)	3.27 (9.46)	0.89 (8.26)	−8.65 (8.46)	−0.73 (8.89)
AQR,6	1.51 (2.24)	−4.29 (3.93)	−0.76 (3.97)	−2.68 (4.18)	2.73 (4.09)	−0.18 (3.83)
AQR,7	2.77 (5.01)	−3.18 (7.28)	1.31 (8.71)	−10.60 (9.56)	6.05 (9.00)	0.93 (8.69)
AQR,8	1.43 (2.05)	−5.68 (3.67)	0.10 (3.35)	0.14 (3.31)	0.85 (3.52)	−0.39 (3.30)
AQR,9	1.46 (2.05)	−5.56 (3.66)	0.60 (3.51)	−0.40 (3.58)	−0.55 (3.68)	0.35 (3.51)
AQR,10	1.20 (2.02)	−5.16 (3.59)	0.12 (3.33)	−0.53 (3.33)	0.65 (3.53)	−0.40 (3.26)

Note: All numbers reported below are multiplied by 1000.

$n=500$, the variances of $\widehat{\boldsymbol{\beta}}_{\mathrm{AQR},1}$ are 0.72, 1.18, 1.18, 1.04, 1.15 and 1.26 times 10^{-3}, respectively, whereas the variances of $\widehat{\boldsymbol{\beta}}_{\mathrm{LSE}}$ are 0.69, 1.11, 1.10, 1.01, 1.12 and 1.21 times 10^{-3}, respectively. In Table 5 when ε follows mixture normal distribution and $n=500$, the variances of $\widehat{\boldsymbol{\beta}}_{\mathrm{AQR},1}$ are 0.31, 0.53, 0.53, 0.59 0.50 and 0.53 times 10^{-3}, respectively, whereas the variances of $\widehat{\boldsymbol{\beta}}_{\mathrm{LSE}}$ are 0.31, 0.55, 0.54, 0.59, 0.54 and 0.56 times 10^{-3}, respectively.

Table 4. The average of the biases and the variance (in the parentheses) of different estimates when ε follows student-*t* distribution with three degrees of freedom.

	β_1	β_2	β_3	β_4	β_5	β_6
			$n = 100, K = 19$			
LSE	0.75 (8.94)	−6.82 (14.96)	−2.07 (17.68)	12.39 (18.89)	−14.29 (17.08)	1.63 (16.01)
QR	−2.31 (6.62)	−3.96 (12.23)	8.78 (11.63)	0.20 (12.81)	−10.69 (11.05)	−0.04 (11.93)
CQR	0.14 (5.12)	−4.11 (8.82)	5.96 (9.59)	2.08 (10.75)	−12.86 (9.77)	−0.94 (9.03)
ACQR,1	0.76 (5.30)	−5.31 (9.48)	5.83 (9.95)	4.42 (11.20)	−12.15 (10.49)	−2.83 (9.46)
ACQR,2	−1.28 (5.31)	−3.33 (9.47)	7.06 (9.90)	2.35 (11.49)	−12.06 (10.14)	−3.34 (9.33)
ACQR,3	−0.48 (5.52)	−3.67 (9.70)	8.82 (9.91)	0.16 (11.33)	−12.03 (9.78)	0.15 (9.67)
ACQR,4	−0.18 (5.59)	−3.62 (10.08)	5.19 (10.06)	1.54 (10.81)	−10.58 (11.03)	0.35 (10.29)
ACQR,5	−0.18 (9.35)	−3.85 (18.36)	4.20 (17.80)	−1.64 (18.81)	−13.02 (17.60)	−0.61 (18.01)
ACQR,6	−0.34 (5.50)	−3.69 (9.28)	6.01 (10.30)	3.13 (11.17)	−13.33 (10.35)	0.23 (9.08)
ACQR,7	2.18 (9.73)	−9.74 (15.10)	6.74 (17.83)	8.32 (19.55)	−18.37 (17.82)	2.70 (15.92)
ACQR,8	−0.86 (5.23)	−3.56 (8.91)	7.54 (9.27)	1.88 (11.01)	−12.86 (9.55)	−0.15 (9.38)
ACQR,9	−0.64 (5.30)	−4.35 (9.20)	7.64 (10.38)	1.30 (11.08)	−14.54 (9.75)	0.85 (9.29)
ACQR,10	7.60 (5.12)	0.79 (10.20)	−5.21 (8.61)	−0.04 (8.54)	4.27 (9.19)	0.57 (9.77)
AQR,1	−1.54 (7.42)	−4.57 (13.06)	5.56 (14.99)	6.06 (15.60)	−9.42 (15.37)	−2.71 (15.98)
AQR,2	0.12 (6.10)	−5.89 (10.22)	7.14 (10.85)	0.92 (12.29)	−9.79 (11.11)	−3.89 (10.23)
AQR,3	−0.38 (5.63)	−3.81 (9.63)	8.93 (9.63)	1.07 (11.47)	−11.65 (9.99)	−0.43 (9.60)
AQR,4	−2.11 (8.09)	−3.43 (14.51)	5.96 (15.79)	−0.97 (15.61)	−10.71 (16.67)	−5.05 (14.47)
AQR,5	−0.14 (20.61)	−9.55 (39.64)	2.89 (41.74)	−6.86 (40.94)	−11.59 (41.17)	−3.75 (36.25)
AQR,6	−0.38 (10.34)	−11.31 (15.61)	10.25 (17.16)	6.10 (16.45)	−8.38 (19.39)	−5.54 (21.17)
AQR,7	3.96 (25.27)	−13.54 (35.31)	2.80 (42.38)	12.52 (38.48)	−9.34 (42.91)	4.50 (43.53)
AQR,8	−0.16 (5.60)	−7.14 (9.57)	8.96 (10.31)	1.55 (11.69)	−11.48 (10.77)	−2.68 (9.83)
AQR,9	−0.91 (6.50)	−5.76 (11.43)	9.07 (12.65)	3.30 (13.11)	−12.98 (12.56)	−0.65 (11.85)
AQR,10	0.43 (5.52)	−5.64 (9.75)	9.05 (10.31)	2.46 (11.64)	−13.39 (10.36)	−1.07 (10.53)
			$n = 500, K = 99$			
LSE	0.34 (2.02)	−2.02 (3.21)	4.39 (3.15)	0.53 (3.15)	−3.88 (3.37)	4.54 (3.40)
QR	0.47 (1.16)	−2.03 (1.72)	4.36 (1.95)	1.18 (2.03)	−2.00 (1.84)	2.15 (1.85)
CQR	0.36 (1.01)	−3.55 (1.53)	4.24 (1.63)	1.22 (1.75)	−2.09 (1.67)	3.34 (1.69)
ACQR,1	0.38 (1.02)	−3.18 (1.52)	3.81 (1.63)	1.22 (1.76)	−1.91 (1.68)	3.46 (1.69)
ACQR,2	0.53 (0.99)	−3.71 (1.48)	4.41 (1.60)	1.46 (1.68)	−1.95 (1.59)	2.68 (1.65)
ACQR,3	0.30 (0.97)	−3.22 (1.45)	3.77 (1.58)	1.96 (1.70)	−2.17 (1.58)	2.84 (1.64)
ACQR,4	0.77 (1.04)	−4.31 (1.59)	3.86 (1.72)	1.50 (1.76)	−1.75 (1.68)	2.95 (1.67)
ACQR,5	0.94 (1.76)	−5.45 (2.90)	3.10 (3.05)	1.72 (2.93)	−2.45 (2.84)	3.24 (2.88)
ACQR,6	0.29 (1.05)	−3.28 (1.58)	5.06 (1.63)	0.94 (1.82)	−1.64 (1.74)	3.16 (1.75)
ACQR,7	−0.49 (1.81)	−1.74 (2.87)	6.11 (2.75)	−0.04 (3.27)	−1.73 (3.25)	4.51 (3.10)
ACQR,8	0.43 (0.98)	−3.59 (1.47)	4.17 (1.59)	1.55 (1.67)	−2.05 (1.59)	2.89 (1.65)
ACQR,9	0.23 (0.99)	−3.57 (1.51)	4.45 (1.58)	1.04 (1.71)	−1.93 (1.63)	3.24 (1.69)
ACQR,10	2.68 (0.99)	−1.15 (1.65)	−2.54 (1.85)	5.12 (1.69)	−2.91 (1.95)	−1.59 (1.92)
AQR,1	1.48 (1.66)	−3.32 (2.79)	2.74 (2.82)	1.04 (2.91)	−2.00 (2.90)	4.81 (2.67)
AQR,2	0.32 (1.04)	−4.03 (1.62)	4.81 (1.67)	1.22 (1.83)	−1.95 (1.79)	3.41 (1.73)
AQR,3	0.04 (0.97)	−3.42 (1.47)	4.21 (1.58)	1.63 (1.70)	−2.03 (1.59)	3.31 (1.64)
AQR,4	0.17 (1.85)	−4.53 (2.76)	4.72 (3.03)	1.47 (2.82)	−4.16 (3.17)	4.08 (2.94)
AQR,5	−0.16 (6.41)	−5.42 (10.74)	5.21 (10.84)	1.39 (9.81)	−7.40 (10.13)	4.81 (10.02)
AQR,6	0.31 (1.77)	−1.82 (3.10)	3.84 (2.81)	1.07 (3.25)	−0.87 (3.24)	3.74 (2.83)
AQR,7	1.09 (5.77)	3.10 (11.58)	−0.76 (9.38)	1.73 (11.37)	−1.37 (11.53)	5.19 (10.46)
AQR,8	0.41 (1.03)	−4.01 (1.59)	4.41 (1.68)	1.04 (1.79)	−1.95 (1.75)	3.60 (1.72)
AQR,9	−0.01 (1.28)	−4.11 (2.10)	4.96 (2.06)	0.36 (2.18)	−1.83 (2.35)	3.38 (2.19)
AQR,10	−0.20 (1.03)	−3.70 (1.55)	4.87 (1.61)	0.82 (1.80)	−1.75 (1.69)	3.61 (1.69)

Note: All numbers reported below are multiplied by 1000.

Throughout all simulations, the variances of $\widehat{\boldsymbol{\beta}}_{\mathrm{CQR}}$, $\widehat{\boldsymbol{\beta}}_{\mathrm{ACQR,1}}$ and $\widehat{\boldsymbol{\beta}}_{\mathrm{AQR,10}}$ are also quite similar, especially when $n = 500$. For example, in Table 3 when ε follows logistic distribution and $n = 500$, the variances of $\widehat{\boldsymbol{\beta}}_{\mathrm{CQR}}$ are 1.99, 3.51, 3.22, 3.24, 3.47 and 3.27 times 10^{-3}, the variances of $\widehat{\boldsymbol{\beta}}_{\mathrm{ACQR,1}}$ are 2.01, 3.51, 3.22, 3.27, 3.52 and 3.33 times 10^{-3}, and the variances of $\widehat{\boldsymbol{\beta}}_{\mathrm{AQR,10}}$ are 2.02, 3.59, 3.33, 3.33, 3.53 and 3.26 times 10^{-3}, respectively. The corresponding variances of $\widehat{\boldsymbol{\beta}}_{\mathrm{CQR}}$, $\widehat{\boldsymbol{\beta}}_{\mathrm{ACQR,1}}$ with $g(\tau) \equiv 1$ and $\widehat{\boldsymbol{\beta}}_{\mathrm{AQR,10}}$ with $g(\tau) \propto$

Table 5. The average of the biases and the variance (in the parentheses) of different estimates when ε follows a mixture of normal distribution $0.5N(0,1) + 0.5N(0, 0.125^2)$.

	β_1	β_2	β_3	β_4	β_5	β_6
			$n = 100, K = 19$			
LSE	−1.74 (1.74)	1.14 (3.04)	−0.01 (3.05)	−2.13 (2.88)	2.49 (3.24)	0.18 (2.61)
QR	−0.37 (0.37)	0.31 (0.57)	0.38 (0.62)	−0.69 (0.68)	0.86 (0.68)	0.50 (0.55)
CQR	−0.82 (0.58)	0.18 (1.03)	0.02 (1.10)	−1.50 (1.05)	1.24 (1.15)	0.61 (0.89)
ACQR,1	−0.88 (0.71)	0.36 (1.20)	0.34 (1.18)	−1.65 (1.16)	0.93 (1.23)	0.53 (0.94)
ACQR,2	−0.92 (0.52)	0.05 (0.91)	−0.03 (1.01)	−1.17 (0.95)	1.00 (0.99)	0.31 (0.78)
ACQR,3	−0.59 (0.40)	0.18 (0.67)	−0.25 (0.72)	−1.37 (0.70)	1.49 (0.76)	0.32 (0.58)
ACQR,4	−0.38 (0.63)	−0.40 (1.10)	0.20 (1.15)	−1.39 (1.08)	0.89 (1.27)	0.47 (0.91)
ACQR,5	−0.33 (2.09)	−2.40 (3.58)	0.45 (3.51)	0.03 (3.56)	1.69 (3.83)	−0.79 (3.69)
ACQR,6	−1.11 (0.62)	0.49 (1.15)	−0.38 (1.25)	−1.12 (1.12)	1.25 (1.31)	0.18 (1.04)
ACQR,7	−4.31 (2.17)	3.11 (4.03)	0.25 (4.13)	−2.69 (3.36)	1.34 (3.80)	0.88 (3.15)
ACQR,8	−0.45 (0.52)	0.04 (0.85)	−0.11 (0.97)	−1.24 (0.91)	1.25 (0.99)	0.51 (0.77)
ACQR,9	−0.67 (0.55)	−0.17 (0.98)	0.12 (1.03)	−2.13 (1.03)	1.23 (1.09)	0.31 (0.82)
ACQR,10	−2.08 (0.39)	0.29 (0.69)	−0.54 (0.67)	−0.09 (0.73)	2.49 (0.88)	−3.38 (0.86)
AQR,1	−2.60 (1.70)	0.68 (2.96)	−2.44 (2.77)	−0.48 (2.56)	1.21 (3.09)	−1.24 (2.36)
AQR,2	−0.92 (0.88)	−0.58 (1.54)	0.04 (1.67)	−1.04 (1.53)	1.22 (1.65)	0.77 (1.25)
AQR,3	−0.88 (0.49)	−0.25 (0.82)	0.31 (0.89)	−1.45 (0.87)	1.40 (0.92)	0.26 (0.69)
AQR,4	−0.74 (1.57)	−0.42 (2.78)	−0.08 (2.73)	−2.43 (2.75)	2.40 (2.78)	−0.52 (2.46)
AQR,5	0.58 (5.18)	−5.79 (8.57)	0.07 (8.25)	−1.87 (9.04)	3.91 (8.78)	−2.28 (8.82)
AQR,6	−3.30 (1.57)	1.93 (3.08)	−2.83 (2.98)	0.04 (2.47)	1.71 (3.05)	0.23 (2.42)
AQR,7	−8.14 (5.37)	4.85 (9.83)	−3.24 (10.25)	−2.00 (7.85)	−0.71 (9.11)	1.59 (7.74)
AQR,8	−1.67 (0.83)	−0.00 (1.45)	0.38 (1.53)	−1.31 (1.37)	1.70 (1.48)	0.24 (1.20)
AQR,9	−1.36 (1.11)	−0.54 (2.11)	0.04 (2.14)	−3.98 (2.00)	1.78 (2.07)	0.56 (1.81)
AQR,10	−0.77 (0.54)	−0.04 (0.96)	−0.49 (1.04)	−1.22 (0.94)	0.93 (1.05)	0.74 (0.80)
			$n = 500, K = 99$			
LSE	−1.17 (0.31)	1.33 (0.55)	−1.06 (0.54)	0.16 (0.59)	−1.13 (0.54)	0.60 (0.56)
QR	−0.30 (0.05)	0.20 (0.08)	−0.74 (0.10)	0.46 (0.09)	−0.34 (0.10)	0.55 (0.09)
CQR	−0.44 (0.07)	0.49 (0.13)	−1.00 (0.14)	0.37 (0.15)	−0.28 (0.15)	0.54 (0.14)
ACQR,1	−0.49 (0.08)	0.49 (0.12)	−0.98 (0.14)	0.45 (0.15)	−0.30 (0.15)	0.70 (0.14)
ACQR,2	−0.36 (0.06)	0.39 (0.10)	−0.91 (0.12)	0.29 (0.12)	−0.18 (0.13)	0.58 (0.12)
ACQR,3	−0.24 (0.05)	0.27 (0.08)	−0.95 (0.09)	0.31 (0.09)	−0.17 (0.11)	0.60 (0.09)
ACQR,4	−0.51 (0.08)	0.50 (0.13)	−0.90 (0.14)	0.55 (0.15)	−0.44 (0.15)	0.54 (0.14)
ACQR,5	−1.30 (0.23)	1.09 (0.43)	−1.07 (0.35)	1.14 (0.44)	−1.51 (0.42)	0.49 (0.39)
ACQR,6	−0.36 (0.08)	0.37 (0.13)	−0.93 (0.15)	0.16 (0.15)	−0.07 (0.15)	0.48 (0.14)
ACQR,7	−0.16 (0.24)	0.44 (0.38)	−1.12 (0.46)	−0.32 (0.45)	0.32 (0.42)	0.50 (0.46)
ACQR,8	−0.30 (0.06)	0.36 (0.10)	−0.97 (0.12)	0.43 (0.12)	−0.29 (0.12)	0.55 (0.11)
ACQR,9	−0.33 (0.07)	0.43 (0.11)	−1.02 (0.13)	0.39 (0.13)	−0.27 (0.14)	0.54 (0.13)
ACQR,10	0.16 (0.06)	0.34 (0.10)	−0.70 (0.10)	0.62 (0.10)	0.21 (0.11)	−0.43 (0.10)
AQR,1	−1.07 (0.31)	0.82 (0.53)	−0.62 (0.53)	0.17 (0.59)	−1.38 (0.50)	0.52 (0.53)
AQR,2	−0.51 (0.14)	0.57 (0.25)	−0.89 (0.26)	0.28 (0.28)	−0.60 (0.26)	0.34 (0.26)
AQR,3	−0.34 (0.06)	0.25 (0.10)	−0.98 (0.12)	0.35 (0.13)	−0.15 (0.13)	0.62 (0.12)
AQR,4	−1.41 (0.31)	1.47 (0.56)	−0.90 (0.50)	0.86 (0.59)	−1.85 (0.54)	0.08 (0.55)
AQR,5	−2.85 (1.22)	2.68 (2.14)	−0.68 (1.85)	1.04 (2.29)	−5.24 (2.07)	0.53 (2.02)
AQR,6	−0.85 (0.32)	0.47 (0.55)	−0.35 (0.58)	−0.82 (0.60)	−0.09 (0.52)	0.52 (0.56)
AQR,7	0.42 (1.25)	−0.09 (1.99)	0.06 (2.21)	−2.93 (2.22)	1.10 (2.09)	0.48 (2.23)
AQR,8	−0.74 (0.14)	0.60 (0.23)	−1.06 (0.25)	0.47 (0.26)	−0.62 (0.25)	0.29 (0.25)
AQR,9	−0.64 (0.22)	0.66 (0.37)	−0.84 (0.36)	0.08 (0.40)	−0.95 (0.37)	0.23 (0.38)
AQR,10	−0.30 (0.07)	0.25 (0.12)	−0.98 (0.14)	0.23 (0.14)	−0.19 (0.15)	0.58 (0.14)

Note: All numbers reported below are multiplied by 1000.

$f\{F^{-1}(\tau)\}$ match each other pretty well, which again confirms our theoretical investigation in Section 2.

Among the five error distributions, the Laplace and the mixture normal distributions, which correspond to scenario (G2) in Table 2 and scenario (G5) in Table 5, respectively, have the largest kurtosis, both $\widehat{\boldsymbol{\beta}}_{\mathrm{ACQR},3}$ and $\widehat{\boldsymbol{\beta}}_{\mathrm{AQR},3}$ have smallest variances. This indicates that, if the error distribution has larger kurtosis, we better generate τ from distributions with larger kurtosis to improve efficiency as well.

Throughout we generate the errors εs from symmetric density functions. For the errors generated from symmetric distributions, if $g(\tau)$ is more skewed, the resultant ACQR and the AQR estimators have larger variances. In particular, $\widehat{\boldsymbol{\beta}}_{\text{ACQR},5}$, $\widehat{\boldsymbol{\beta}}_{\text{ACQR},7}$, $\widehat{\boldsymbol{\beta}}_{\text{AQR},5}$ and $\widehat{\boldsymbol{\beta}}_{\text{AQR},7}$ are less efficient than $\widehat{\boldsymbol{\beta}}_{\text{ACQR},4}$, $\widehat{\boldsymbol{\beta}}_{\text{ACQR},6}$, $\widehat{\boldsymbol{\beta}}_{\text{AQR},4}$ and $\widehat{\boldsymbol{\beta}}_{\text{AQR},6}$.

Because $g(\tau) = 0.5\pi\,\sin(\pi\tau)$ has very similar shape as $g(\tau) = \text{Beta}(2, 2)$, the ACQR estimators $\widehat{\boldsymbol{\beta}}_{\text{ACQR},2}$ and $\widehat{\boldsymbol{\beta}}_{\text{ACQR},8}$, along with the AQR estimators $\widehat{\boldsymbol{\beta}}_{\text{AQR},2}$ and $\widehat{\boldsymbol{\beta}}_{\text{AQR},8}$, have roughly equal variances even when n is as small as 100.

In most scenarios, both $\widehat{\boldsymbol{\beta}}_{\text{ACQR},10}$ and $\widehat{\boldsymbol{\beta}}_{\text{AQR},10}$ have the smallest variances, indicating that choosing $g(\tau) \propto f\{F^{-1}(\tau)\}$ is perhaps a safe choice which usually yields comparatively more efficient estimators when the error distribution is known. In practice, however, the error density $f\{F^{-1}(\tau)\}$ is often unknown but can be estimated using nonparametric smoothers easily.

In general, the AQR estimators are not as stable as the ACQR estimators. The ACQR is more computationally extensive, however, our limited experiences show that it usually yields comparatively more efficient estimator than the AQR method.

3.2. *The prostate data*

In this section we compare the relative efficiency of several estimators using the Prostate data. This data set was collected by Stamey et al. [12] and is available in the R package 'lasso2'. It aims to study the relation between the level of prostate specific antigen and a number of clinical measures in men who were about to receive a radical prostatectomy. There are 97 observations and each observation contains one response variable and 8 numerical covariates. The response variable lpsa stands for the log of prostate specific antigen. Numeric covariates include the log of cancer volume(lcavol), the log of prostate weight(lweight), age, the log of benign prostatic hyperplasia amount(lbph), seminal vesicle invasion(svi), the log of capsular penetration(lcp), Gleason score(gleason) and percentage Gleason scores 4 or 5(pgg45). We first centralize both the response variable and the covariates so that there is no intercept in the linear model.

In this example the sample size $n = 97$ is relatively small. We set $K = 19$ in the CQR, the AQR and the ACQR methods. For the adaptive methods, we generate τ from uniform distributions for simplicity. To compare the relative efficiency of various estimators, we

Table 6. The estimated coefficients along with their standard deviations for the Prostate data set.

		lcavol	lweight	age	lbph	svi	lcp	gleason	pgg45
LSE	$\widehat{\beta}$	58.70	45.45	−1.96	10.71	76.62	−10.55	4.51	0.45
	std	9.88	16.40	1.29	5.80	25.65	10.10	15.21	0.40
QR	$\widehat{\beta}$	53.26	55.54	−2.68	15.98	80.01	−13.09	20.32	0.41
	std	10.90	18.09	1.42	6.40	28.29	11.14	16.78	0.44
CQR	$\widehat{\beta}$	57.44	48.00	−2.11	13.00	78.58	−13.69	9.04	0.55
	std	9.50	15.77	1.24	5.58	24.67	9.71	14.63	0.38
ACQR,1	$\widehat{\beta}$	56.61	49.95	−2.15	12.87	76.27	−11.65	11.29	0.47
	std	9.48	15.74	1.24	5.57	24.63	9.70	14.61	0.38
AQR,1	$\widehat{\beta}$	60.15	38.54	−1.49	10.99	78.53	−9.23	1.86	0.51
	std	9.89	16.41	1.29	5.81	25.67	10.11	15.22	0.40

Note: All numbers reported below are multiplied by 100.

use bootstrap to estimate their variances. The estimated regression coefficients and their standard deviations are summarized in Table 6,

In this example, both the LSE and the AQR methods have very similar variances. So do the CQR and the ACQR methods. These observations again confirm our theoretical investigations in Section 2. The ACQR and the CQR methods are the most efficient whereas the QR method is the least efficient. The AQR method is comparable with both the ACQR and the CQR methods. All these estimates indicate that age, lcp, gleaso and pgg45 are not predictive for the response variable, while lcavol, lweight and lbph and svi are important covariates.

4. A brief discussion

In this paper we propose two adaptive methods to estimate regression coefficients in linear models. We demonstrate that when the quantile level τ follows uniform distribution, the ACQR method is asymptotically as efficient as the CQR method, and the AQR method is asymptotically equivalent to the LSE method. By treating the distribution function of the quantile level as a random variable, the adaptive methods have the potential to improve the efficiency of the CQR and the LSE methods. However, how to choose an optimal distribution function for the quantile level to minimize the asymptotic variance is not straightforward. This is beyond the scope of the present work and future research along this direction is warranted.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work is supported by National Natural Science Foundation of China [11371236, 11422107 and 11471335], Project of Key Research Institute of Humanities and Social Sciences at Universities [16JJD910002] and Henry Fok Education Foundation Fund of Young College Teachers [141002], Ministry of Education, China, and National Youth Top-notch Talent Support Program, China.

References

[1] Koenker R, Bassett Jr. G. Regression quantiles. Econometrika. 1978;1:33–50.
[2] Koenker R. A note on l-estimates for linear models. Statist Probab Lett. 1984;2(6):323–325.
[3] Zhao K, Lian H. A note on the efficiency of composite quantile regression. J Stat Comput Simul. 2016;86(7):1334–1341.
[4] Zou H, Yuan M. Composite quantile regression and the oracle model selection theory. Ann Statist. 2008;36(3):1108–1126.
[5] Jiang X, Jiang J, Song X. Oracle model selection for nonlinear models based on weighted composite quantile regression. Statist Sinica. 2012;22:1479–1506.
[6] Sun J, Gai Y, Lin L. Weighted local linear composite quantile estimation for the case of general error distributions. J Statist Plann Inference. 2013;143(6):1049–1063.
[7] Kai B, Li R, Zou H. Local composite quantile regression smoothing: an efficient and safe alternative to local polynomial regression. J R Stat Soc Ser B Statist Methodol. 2010;72(1):49–69.
[8] Kai B, Li R, Zou H. New efficient estimation and variable selection methods for semiparametric varying-coefficient partially linear models. Ann Statist. 2011;39(1):305–332.

[9] Jiang R, Zhou Z, Qian W, et al. Two step composite quantile regression for single-index models. Comput Statist Data Anal. 2013;64:180–191.
[10] Guo J, Tang M, Tian M, et al. Variable selection in high-dimensional partially linear additive models for composite quantile regression. Comput Statist Data Anal. 2013;65:56–67.
[11] Koenker R. Quantile regression. New York: Cambridge University Press; 2005.
[12] Stamey TA, Kabalin JN, McNeal JE, et al. Prostate specific antigen in the diagnosis and treatment of adenocarcinoma of the prostate. ii. radical prostatectomy treated patients. J Urol. 1989;141(5):1076–1083.

Appendix. Proof of main results

In what follows we present the technical proofs for Theorem 2.1, Corollary 2.2, Theorem 2.3, Corollary 2.4 and Corollary 2.5. We remark here that the result of Theorem 2.3 is built upon [4]. The difficulty here is that we treat the quantile levels as realizations of a random variable.

Appendix 1. Proof of Theorem 2.1

Under conditions (A1)–(A2), for a given quantile τ_k,

$$E(\widehat{\boldsymbol{\beta}}_{\tau_k} \mid \tau_k) = \boldsymbol{\beta} + O(n^{-1/2}),$$
$$\mathrm{var}(\widehat{\boldsymbol{\beta}}_{\tau_k} \mid \tau_k) = E\{(\widehat{\boldsymbol{\beta}}_{\tau_k} - \boldsymbol{\beta})^2 \mid \tau_k\} = n^{-1}\boldsymbol{\Sigma}^{-1}\tau_k(1-\tau_k)/f^2\{F^{-1}(\tau_k)\} + O(n^{-1}).$$

See, for example, Koenker [11] for details about the above two results. It is also straightforward to verify that, for two independent quantiles τ_k and τ_j,

$$\begin{aligned}\mathrm{cov}(\widehat{\boldsymbol{\beta}}_{\tau_k}, \widehat{\boldsymbol{\beta}}_{\tau_j} \mid \tau_k, \tau_j) &= E\{(\widehat{\boldsymbol{\beta}}_{\tau_k} - \boldsymbol{\beta})(\widehat{\boldsymbol{\beta}}_{\tau_j} - \boldsymbol{\beta}) \mid \tau_k, \tau_j\} \\ &= n^{-1}\boldsymbol{\Sigma}^{-1}(\tau_k \wedge \tau_j - \tau_k\tau_j)/[f\{F^{-1}(\tau_k)\}f\{F^{-1}(\tau_j)\}] + O(n^{-1}).\end{aligned}$$

In addition, the conditional distributions of $\widehat{\boldsymbol{\beta}}_{\tau_k}, k = 1, \ldots, K$ and the joint conditional distributions of $\widehat{\boldsymbol{\beta}}_{\tau_k}, \widehat{\boldsymbol{\beta}}_{\tau_j}, k, j = 1, \ldots, K$ are all enjoy asymptotic normality. Thus, when given τ_k, $K^{-1}\sum_{k=1}^{K}\widehat{\boldsymbol{\beta}}_{\tau_k}$ also asymptotically follows normal distribution. The mean and variance of conditional $K^{-1}\sum_{k=1}^{K}\widehat{\boldsymbol{\beta}}_{\tau_k}$, also denoted conditional $\widehat{\boldsymbol{\beta}}_{\mathrm{AQR}}$, is established as follows,

$$E(\widehat{\boldsymbol{\beta}}_{\mathrm{AQR}} \mid \tau_1, \ldots, \tau_K) = K^{-1}\sum_{k=1}^{K} E(\widehat{\boldsymbol{\beta}}_{\tau_k} \mid \tau_k) = \boldsymbol{\beta} + O(n^{-1/2}),$$

$$\begin{aligned}\mathrm{var}(\widehat{\boldsymbol{\beta}}_{\mathrm{AQR}} \mid \tau_1, \ldots, \tau_K) &= K^{-2}\sum_{k=1}^{K}\sum_{j=1}^{K}\mathrm{cov}(\widehat{\boldsymbol{\beta}}_k, \widehat{\boldsymbol{\beta}}_j \mid \tau_k, \tau_j) \\ &= n^{-1}\boldsymbol{\Sigma}^{-1}K^{-2}\sum_{k=1}^{K}\sum_{j=1}^{K}\frac{(\tau_k \wedge \tau_j - \tau_k\tau_j)}{f\{F^{-1}(\tau_k)\}f\{F^{-1}(\tau_j)\}} + O(n^{-1}).\end{aligned}$$

As $K \to \infty$, by the law of large numbers, we have, in probability,

$$\begin{aligned}K^{-2}\sum_{k=1}^{K}\sum_{j=1}^{K}\frac{(\tau_k \wedge \tau_j - \tau_k\tau_j)}{f\{F^{-1}(\tau_k)\}f\{F^{-1}(\tau_j)\}} &\longrightarrow E\left[\frac{(\tau_k \wedge \tau_j - \tau_k\tau_j)}{f\{F^{-1}(\tau_k)\}f\{F^{-1}(\tau_j)\}}\right] \\ &= 2\left[\int_0^1\int_0^{\tau_2}\frac{\tau_1(1-\tau_2)}{f\{F^{-1}(\tau_1)\}f\{F^{-1}(\tau_2)\}}g(\tau_1)g(\tau_2)\,\mathrm{d}\tau_1\,\mathrm{d}\tau_2\right].\end{aligned}$$

Define

$$B \overset{\mathrm{def}}{=} 2\left[\int_0^1\int_0^{\tau_2}\frac{\tau_1(1-\tau_2)}{f\{F^{-1}(\tau_1)\}f\{F^{-1}(\tau_2)\}}g(\tau_1)g(\tau_2)\,\mathrm{d}\tau_1\,\mathrm{d}\tau_2\right].$$

Then, when K goes to infinity, density function of $\widehat{\boldsymbol{\beta}}_{\mathrm{AQR}}$ given $\tau_2, \ldots, \tau_K$ is as follows,

$$f(\widehat{\boldsymbol{\beta}}_{\mathrm{AQR}} \mid \tau_1, \ldots, \tau_K) \to \{2\pi n^{-1} \det(\boldsymbol{\Sigma}^{-1}B)\}^{-n/2}$$
$$\exp\{-1/2(\widehat{\boldsymbol{\beta}}_{\mathrm{AQR}} - \boldsymbol{\beta})^{\mathrm{T}}(n^{-1}\boldsymbol{\Sigma}^{-1}B)^{-1}(\widehat{\boldsymbol{\beta}}_{\mathrm{AQR}} - \boldsymbol{\beta})\}$$

By law of total probability, the unconditional density function of $\widehat{\boldsymbol{\beta}}_{\mathrm{AQR}}$ is

$$\begin{aligned} f(\widehat{\boldsymbol{\beta}}_{\mathrm{AQR}}) &= \int_0^1 \cdots \int_0^1 f(\widehat{\boldsymbol{\beta}}_{\mathrm{AQR}} \mid \tau_1, \ldots, \tau_K) g(\tau_1) \ldots g(\tau_K)\, \mathrm{d}\tau_1 \ldots \mathrm{d}\tau_K \\ &\to \{2\pi n^{-1} \det(\boldsymbol{\Sigma}^{-1}B)\}^{-n/2} \exp\{-1/2(\widehat{\boldsymbol{\beta}}_{\mathrm{AQR}} - \boldsymbol{\beta})^{\mathrm{T}}(n^{-1}\boldsymbol{\Sigma}^{-1}B)^{-1}(\widehat{\boldsymbol{\beta}}_{\mathrm{AQR}} - \boldsymbol{\beta})\} \end{aligned}$$

Thus, under conditions (A1) and (A2), as $K \to \infty$,

$$n^{1/2}(\widehat{\boldsymbol{\beta}}_{\mathrm{AQR}} - \boldsymbol{\beta}) \xrightarrow{d} \mathcal{N}(0, B\boldsymbol{\Sigma}^{-1}).$$

This completes the proof of Theorem 2.1. ■

Appendix 2. Proof of Corollary 2.2

When σ^2 is finite and $g(\tau) \equiv 1$, B boils down to

$$\begin{aligned} 2\int_0^1 \int_0^{\tau_2} \frac{\tau_1(1-\tau_2)}{f(b_{\tau_1})f(b_{\tau_2})}\, \mathrm{d}\tau_1\, \mathrm{d}\tau_2 &= 2\int_{-\infty}^{\infty} \{1 - F(z_2)\} G(z_2)\, \mathrm{d}z_2 \\ &= 2\int_{-\infty}^{\infty} \int_{z_2}^{\infty} f(t)\, \mathrm{d}t G(z_2)\, \mathrm{d}z_2 \\ &= \int_{-\infty}^{\infty} f(t)\{2\int_{-\infty}^{t} G(z_2)\, \mathrm{d}z_2\} \mathrm{d}t. \end{aligned}$$

The first equality follows by letting $\tau_k = F(z_k)$ for $k = 1, 2$, and

$$G(z_2) \overset{\text{def}}{=} \int_{-\infty}^{z_2} F(z_1)\, \mathrm{d}z_1.$$

Using the fact that

$$\begin{aligned} 2\int_{-\infty}^{t} G(z_2)\, \mathrm{d}z_2 &= 2\int_{-\infty}^{t} F(s) \int_s^t \mathrm{d}z_2\, \mathrm{d}s = 2\int_{-\infty}^{t} (t-s) \int_{-\infty}^{s} f(x)\, \mathrm{d}x\, \mathrm{d}s \\ &= 2\int_{-\infty}^{t} f(x) \int_x^t (t-s)\, \mathrm{d}s\, \mathrm{d}x = \int_{-\infty}^{t} (t-x)^2 f(x)\, \mathrm{d}x, \end{aligned}$$

we can easily obtain that

$$\int_{-\infty}^{\infty} f(t) \left\{2\int_{-\infty}^{t} G(z_2)\, \mathrm{d}z_2\right\} \mathrm{d}t = \int_{-\infty}^{\infty} t^2 f(t)\, \mathrm{d}t - 2\int_{-\infty}^{\infty} tf(t) \int_{-\infty}^{t} xf(x)\, \mathrm{d}x\, \mathrm{d}t. \tag{A1}$$

Note that

$$\int_{-\infty}^{\infty} tf(t) \int_{-\infty}^{t} xf(x)\, \mathrm{d}x\, \mathrm{d}t = \int_{-\infty}^{\infty} tf(t) \int_t^{\infty} xf(x)\, \mathrm{d}x\, \mathrm{d}t = \frac{1}{2}\int_{-\infty}^{\infty} tf(t) \int_{-\infty}^{\infty} xf(x)\, \mathrm{d}x\, \mathrm{d}t.$$

Consequently, the right-hand side of Equation (A1) is equal to $\mathrm{var}(\varepsilon) = \sigma^2$. Accordingly,

$$2\int_0^1 \int_0^{\tau_2} \frac{\tau_1(1-\tau_2)}{f(b_{\tau_1})f(b_{\tau_2})}\, \mathrm{d}\tau_1\, \mathrm{d}\tau_2 = 2\int_{-\infty}^{\infty} f(t) \left\{\int_{-\infty}^{t} G(z_2)\, \mathrm{d}z_2\right\} \mathrm{d}t = \sigma^2.$$

Next we consider the second case: $g(\tau) \propto f\{F^{-1}(\tau)\}$. In other words, $g(\tau) = cf\{F^{-1}(\tau)\}$ for a certain c. Because $g(\cdot)$ is a density function, we use this fact to calculate what value c takes. Note that

$$1 = \int_0^1 g(\tau)\,\mathrm{d}\tau = c\int_0^1 f\{F^{-1}(\tau)\}\,\mathrm{d}\tau = c\int_{-\infty}^{\infty} \{f(x)\}^2\,\mathrm{d}x.$$

This indicates that $c = 1/M$ and $M \overset{\text{def}}{=} E\{f(\varepsilon)\}$. The second statement follows immediately. ■

Appendix 3. Proof of Theorem 2.3

Given a fixed K and the quantile levels $\tau_1, \tau_2, \ldots, \tau_K$, Zou and Yuan [4] showed that, under conditions (A1) and (A3), $n^{1/2}(\widehat{\boldsymbol{\beta}}_{\mathrm{CQR}} - \boldsymbol{\beta}) \xrightarrow{d} \mathcal{N}(0, D_K\boldsymbol{\Sigma}^{-1})$. As $K \to \infty$, by the law of large numbers, we have, in probability,

$$\begin{aligned} K^{-2}\sum_{k=1}^{K}\sum_{k'=1}^{K} \min(\tau_k, \tau_k')\left\{1 - \max(\tau_k, \tau_k')\right\} &\longrightarrow E[\min(\tau_1, \tau_2)\{1 - \max(\tau_1, \tau_2)\}] \\ &= 2\left\{\int_0^1\int_0^{\tau_2} \tau_1(1-\tau_2)g(\tau_1)g(\tau_2)\,\mathrm{d}\tau_1\,\mathrm{d}\tau_2\right\}. \end{aligned}$$

Again by the law of large numbers,

$$\left\{K^{-1}\sum_{k=1}^{K} f(b_{\tau_k})\right\}^2 \longrightarrow \left\{\int_0^1 f(b_\tau)g(\tau)\,\mathrm{d}\tau\right\}^2 \text{ inprobability.}$$

Define

$$A \overset{\text{def}}{=} 2\left\{\int_0^1\int_0^{\tau_2} \tau_1(1-\tau_2)g(\tau_1)g(\tau_2)\,\mathrm{d}\tau_1\,\mathrm{d}\tau_2\right\} \Big/ \left\{\int_0^1 f(b_\tau)g(\tau)\,\mathrm{d}\tau\right\}^2.$$

Then, when K goes to infinity, density function of $\widehat{\boldsymbol{\beta}}_{\mathrm{ACQR}}$ given $\tau_2, \ldots, \tau_K$ is as follows,

$$\begin{aligned} f(\widehat{\boldsymbol{\beta}}_{\mathrm{ACQR}} \mid \tau_1, \ldots, \tau_K) &\to \{2\pi n^{-1}\det(\boldsymbol{\Sigma}^{-1}A)\}^{-n/2} \\ &\exp\{-1/2(\widehat{\boldsymbol{\beta}}_{\mathrm{ACQR}} - \boldsymbol{\beta})^{\mathrm{T}}(n^{-1}\boldsymbol{\Sigma}^{-1}A)^{-1}(\widehat{\boldsymbol{\beta}}_{\mathrm{ACQR}} - \boldsymbol{\beta})\} \end{aligned}$$

By law of total probability, the unconditional density function of $\widehat{\boldsymbol{\beta}}_{\mathrm{ACQR}}$ is

$$\begin{aligned} f(\widehat{\boldsymbol{\beta}}_{\mathrm{ACQR}}) &= \int_0^1 \ldots \int_0^1 f(\widehat{\boldsymbol{\beta}}_{\mathrm{ACQR}} \mid \tau_1, \ldots, \tau_K)g(\tau_1)\ldots g(\tau_K)\,\mathrm{d}\tau_1\ldots\mathrm{d}\tau_K \\ &\to \{2\pi n^{-1}\det(\boldsymbol{\Sigma}^{-1}A)\}^{-n/2}\exp\{-1/2(\widehat{\boldsymbol{\beta}}_{\mathrm{ACQR}} - \boldsymbol{\beta})^{\mathrm{T}}(n^{-1}\boldsymbol{\Sigma}^{-1}A)^{-1}(\widehat{\boldsymbol{\beta}}_{\mathrm{ACQR}} - \boldsymbol{\beta})\} \end{aligned}$$

Thus, under conditions (A1) and (A3), as $K \to \infty$,

$$n^{1/2}(\widehat{\boldsymbol{\beta}}_{\mathrm{ACQR}} - \boldsymbol{\beta}) \xrightarrow{d} \mathcal{N}(0, A\boldsymbol{\Sigma}^{-1}).$$

This completes the proof of Theorem 2.3. ■

Appendix 4. Proof of Corollary 2.4

When $g(\tau) \equiv 1$, A reduces to

$$2\left\{\int_0^1\int_0^{\tau_2} \tau_1(1-\tau_2)\,\mathrm{d}\tau_1\,\mathrm{d}\tau_2\right\} \Big/ \left\{\int_0^1 f(b_\tau)\,\mathrm{d}\tau\right\}^2 = (1/12)\Big/\left[\int_{-\infty}^{\infty}\{f(\varepsilon)\}^2\,\mathrm{d}\varepsilon\right]^2.$$

The above equality follows by letting $\tau = F(z)$, and accordingly, $b_\tau = F^{-1}(\tau)$. ■

Appendix 5. Proof of Corollary 2.5

By the concept of integration, we can rewrite A as

$$A = \frac{\lim_{K\to\infty}\sum_{k,k'=1}^{K+1}[(k\wedge k')/(K+1) - kk'/(K+1)^2]g\{k/(K+1)\}g\{k'/(K+1)\}}{\lim_{K\to\infty}[\sum_{k=1}^{K+1} f[F^{-1}\{k/(K+1)\}]g\{k/(K+1)\}]^2}.$$

Let $k/(K+1) = \tau_k$ and $g\{k/(K+1)\} = Kw_k$ where $w_k = 1/(K+1)$, then

$$\begin{aligned}
A &= \lim_{K\to\infty} \frac{\sum_{k,k'=1}^{K+1}(\tau_k \wedge \tau_{k'} - \tau_k\tau_{k'})w_k w_{k'}}{[\sum_{k=1}^{K+1} f\{F^{-1}(\tau_k)\}w_k]^2} \\
&= \lim_{K\to\infty} \frac{\sum_{k,k'=1}^{K}(\tau_k \wedge \tau_{k'} - \tau_k\tau_{k'})w_k w_{k'} + (\tau_{K+1} - \tau_{K+1}^2)w_{K+1}^2}{[\sum_{k=1}^{K} f\{F^{-1}(\tau_k)\}w_k + f\{F^{-1}(\tau_{K+1})\}w_{K+1}]^2} \\
&= \lim_{K\to\infty} \frac{\sum_{k,k'=1}^{K} w_k w_{k'}(\tau_k \wedge \tau_{k'} - \tau_k\tau_{k'})}{[\sum_{k=1}^{K} w_k f\{F^{-1}(\tau_k)\}]^2} = W_{2,\infty}.
\end{aligned}$$

We can obtain the relationship between B and $W_{1,\infty}$ by similar methods. ∎