Quantile Regression for Count Data with Optimized Jittering

by

©Haoran WANG

A Thesis submitted to the School of Graduate Studies in partial fulfillment
of the requirements for the degree of

Master of Science in Statistics
Department of Mathematics and Statistics

Memorial University of Newfoundland

November 28, 2024

St. John's, Newfoundland and Labrador, Canada

Abstract

Quantile regression (QR) is a natural extension to the classic linear regression. It models the conditional quantiles of the continuous response variable instead of modeling the conditional mean. Often times we encounter discrete response variables, such as counts or other kind of categorical variables. Due to the discontinuity of counts, traditional QR creates systematic bias when applied to count data and hence is not directly applicable. Jittering with uniform random perturbations is one of the options to smooth the discrete response. In this thesis, we propose a new QR model for count data, which improves the existing uniform jittering method. The proposed approach involves artificially adding Tweedie or Beta random perturbations to original count response, generating pseudo-continuous QR response. Through proper selection of perturbation parameters, jittering can provide better parameter estimation of the regression parameters as compared with the existing methods. We employ the Asymmetric Laplace Distribution (ALD) to determine the optimal perturbation parameters through the Monte-Carlo Expectation Maximization (MCEM) and Metropolis-Hasting (MH) sampler. Our proposed method for QR model provides consistent estimators of the QR coefficients. The estimators follow asymptotic normal distribution as sample size goes to infinity. Simulation studies show much improved performance when sample sizes are small to moderate. As an illustration, the proposed method was applied to analyze a fishery data.

Lay Summary

Quantile regression (QR) is a natural extension to the classic linear regression. It models the conditional quantiles of the continuous response variable instead of modeling the conditional mean. Often times we encounter discrete response variables, such as counts or other kind of categorical variables. Direct application of traditional QR to count data creates systematic bias and leads to invalid inference. Jittering is a method used to QR for count data. By adding a small amount of random perturbations (jittering) to response, the QR technique can be applied to count data. In this thesis, we generate random perturbations from Tweedie or Beta distributions to smooth the discrete response. We use the Asymmetric Laplace Distribution (ALD) to form a pseudo likelihood function and then apply Monte-Carlo Expectation Maximization (MCEM), and Metropolis-Hasting (MH) sampler to find the parameters that optimize performance of jittering.

Acknowledgments

I am grateful to the Hong Kong University of Science and Technology for introducing me to the world of mathematics and to the Memorial University of Newfoundland for providing me with the great opportunity to study in Canada and write my first thesis.

I would like to express my deepest gratitude to my supervisor, Dr. Zhaozhi Fan, who introduced me to Quantile Regression, for his unwavering support, patience, guidance, and invaluable insights throughout the course of my research. His expertise and encouragement have been instrumental in the successful completion of this thesis.

My heartfelt thanks go to my Father, Mother, sister, pet dogs, whose constant encouragement and understanding have been a source of strength throughout this journey. Your belief in me has been a driving force in my academic pursuits. I am also deeply grateful to my friends in Newfoundland, who supported me both physically and mentally.

I would like to acknowledge the financial support provided by the School of Graduate Studies, Department of Mathematics and Statistics in the form of Graduate Assistantship and Teaching Assistantship and financial support from Dr. Zhaozhi Fan (NSERC DG) and Dr. Asokan Variyath (AARMS CRG).

To all who have contributed to my work in any way, I extend my sincere thanks.

Haoran (Mark) Wang

Statement of Contribution

This research was a collaborative effort involving the following contributions:

- **Mr. Haoran Wang**: As a primary researcher, I was responsible for providing the proposal of research topic, developing the approach, conducting the numerical simulations, fishing data collection and analysis. I also contributed to writing manuscript and preparing seminars and presentations.

- **Dr. Zhaozhi Fan**: My supervisor provided oversights and guidance throughout the research process. His contributions included conceptualizing the research idea, providing professional feedback, assisting with interpretation of research findings, securing funding. Dr. Zhaozhi Fan also played a key role in research design, reviewing drafts, and scheduling dates of seminars for the department.

The collaborative effort between Dr. Zhaozhi Fan and me was crucial to the successful completion of this research thesis.

Contents

List of Figures

List of Tables

Nomenclature

System of names and terms, and rules for forming mathematical and statistical terms that appear in this thesis are as follows, unless specifically stated otherwise:

- Random variable: X
- Vectors: $\boldsymbol{x}$ (lower case bold letters). The subscript $[n]$ denotes the vector dimension where the notation $\boldsymbol{x}_{[n]}$ is used.
- Random vector: $\boldsymbol{X}$ (upper case bold letters). The subscript $[n]$ denotes the vector dimension where the notation $\boldsymbol{X}_{[n]}$ is used.
- Matrix or random vector: $\boldsymbol{X}_{[n \times p]}$ (upper case bold letters). The subscript $[n \times p]$ denotes the matrix dimensions.
- Transpose operator: T or $'$ (e.g., $\boldsymbol{X}^T$ or $\boldsymbol{X}'$)
- Probability distribution function (PDF): $f_Y(y)$, where the Y subscript denotes the variables on which the function is computed. The shortened notation $f(y)$ is used where there is no risk of ambiguity.
- Cumulative distribution function (CDF): $F_Y(y)$, where the Y subscript denotes the variables on which the function is computed. The shortened notation $F(y)$ is used where there is no risk of ambiguity.
- Quantile level: τ, where $\tau \in (0, 1)$ (e.g., 0.25^{th} quantile is the first quartile $Q1$).

- Quantile function: $Q_Y(\tau)$, where the Y subscript denotes the variables on which the quantile is computed. The shortened notation $Q(\tau)$ is used where there is no risk of ambiguity.
- i-th vector element: x_i
- i-th matrix row: $\boldsymbol{x}_i$
- Null vector: $\mathbf{0}$
- Identity vector: $\mathbf{1}$
- Identity matrix: $\boldsymbol{I}$
- Population size: N
- Sample size: n
- Number of regressors: p
- Mean regression parameter: $\boldsymbol{\beta} = (\beta_1, ..., \beta_p)$.
- Mean regression estimator: $\hat{\boldsymbol{\beta}} = (\hat{\beta}_1, ..., \hat{\beta}_p)$.
- Quantile regression parameter: $\boldsymbol{\beta}_\tau = (\beta_{\tau 1}, ..., \beta_{\tau p})$ or $\boldsymbol{\beta}(\tau) = (\beta_1(\tau), ..., \beta_p(\tau))$.
- Quantile regression estimator: $\hat{\boldsymbol{\beta}}_\tau = (\hat{\beta}_{\tau 1}, ..., \hat{\beta}_{\tau p})$ or $\hat{\boldsymbol{\beta}}(\tau) = (\hat{\beta}_1(\tau), ..., \hat{\beta}_p(\tau))$.
- Check loss function: $\rho_\tau(u)$
- Simple quantile regression model: $Q_Y(\tau|x) = \beta_{\tau 0} + \beta_{\tau 1} x + \varepsilon_\tau$
- Multiple quantile regression model: $Q_Y(\tau|\boldsymbol{X}) = \boldsymbol{X}'\boldsymbol{\beta}_\tau + \boldsymbol{\varepsilon}_\tau$

Chapter 1

Introduction

Regression analysis is one of the most frequently used statistical methods. Linearly modeling the conditional expectation of the response variable is subject to some strict assumptions, such as normality and homogeneity of the error terms. It is not robust to outliers and skewed distributions.If assumptions are violated, the conclusion may be misleading. Koenker and Bassett (1978) proposed quantile regression (QR) method that models the effect of independent variables on the quantiles of the response distribution. QR extends the mean regression model to any arbitrary quantiles of the distribution. It provides a complete picture of the relationship between the response and the independent variables when quantiles are properly selected. It only requires the response variable to have a continuous distribution, and is more flexible to analyze data with non-normal distributions or heterogeneous variances.

In many areas of studies, a relationship between response and independent variables can only be observed at some specific quantiles. For example in ecology, effects of ocean temperature on coral polyps growth can only be seen at the higher temperature, since it causes coral polyps to loose zooxanthellae that live in the polyps' tissues (Ding et al., 2022). QR is robust to outliers and is more flexible in modeling real data. The method has been widely used in diverse areas of studies, such as medicine, sociology, finance, etc., (see Peng, 2021; Perillo et al., 2017; C. W. Chen et al., 2012; Friederichs and Hense, 2007; Hajdu and Hajdu, 2014, among others).

Count data are frequently encountered in areas where events across different sectors were recorded and hence needed to be modeled, such as healthcare, economics and environmental sciences. Poisson distribution is most frequently used to model count data. In QR analysis, continuity of response distribution is an important condition due to the continuous nature of quantiles of a distribution. Naively using traditional QR method designed for continuous data to count data leads to systematic bias of estimation, (see Rampichini et al., 2015; Geraci and Farcomeni, 2022; Harding and Lamarche, 2019; Padellini and Rue, 2018; Machado and Silva, 2005).

Manski (1975) firstly introduced maximum score estimator of normalized parameters that only requires weak distributional assumption for consistency and strong consistency of the estimator was also proved (Manski, 1985). But unfortunately insufficient convergence rate in distribution $n^{-1/3}$ makes the inference invalid in practice (see Kim and Pollard, 1990). One solution was initially contributed by Horowitz (1992), who focused on convergence rate and modified the maximum score estimation to fit median regression on binary and multinomial responses. This modified estimation enhanced the convergence rate to $n^{-2/5} \sim n^{-1/2}$ depending on the strength of assumptions, making statistical inference possible with large sample size. Later, Lee (1992) generalized maximum score estimation to introduce semi-parametric median regression for ordered discrete response.

As discussed by Koenker (2005) and others, the fact that the sample objective function is non-differentiable when estimating the conditional quantile function with counts, makies Taylor expansion not applicable. Discontinuity of objective function due to positive mass makes the asymptotic distributions of estimator complicated to obtain. Lu and Fan (2020) proposed generalized linear quantile mixed model, estimating parameters by using Newton-Raphson algorithms. Another strategy is to smooth the counts directly, leading to a 'pseudo-continuous' response. Machado and Silva (2005) introduced 'jitterng the counts' that involves artificially adding some noises or perturbations to the original counts to smooth the discreteness. Let Y be a non-negative integer. A random variable U following uniform $[0, 1)$ or $[-0.5, 0.5)$ is introduced to integer and make $Z = Y + U$ be the QR dependent variable. Valid statistical inference can still be done with jittering.

In this thesis, we propose to draw random numbers from the Tweedie and Beta distributions for jittering. These distributions involve location and scale parameters. The optimal choice of the parameters of Tweedie (μ, ϕ, p) and Beta (a, b) distributions can be determined by Monte-Carlo Expectation Maximization (MCEM) algorithm. The results observed from intensive numerical study show the QR model for counts with Tweedie and Beta noises provides extra flexibility and maintain a decent performance in parameter estimation and inference.

The remainder of this thesis is organized as follows: in Chapter 2 we briefly introduce the quantile regression, including those models for discrete data. In Chapter 3 we propose the QR method for counts data and statistical algorithms for choosing noise parameters. Consistency, asymptotic normal distribution of QR estimators and corresponding sandwich estimates of variance-covariance matrices are derived in this chapter. In Chapter 4 we study the performance of different QR methods through simulation. The proposed QR model is illustrated by analyzing a fishery data. In Chapter 5 we conclude the thesis with some remarks. Some model developments and future research are given in the this chapter.

Chapter 2

Quantile Regression Models and Methods

In this chapter, we briefly review concepts in mean regression, which then facilities the introduction of QR models more clearly. Section 2.1 introduces cumulative density functions and unconditional quantiles and functions, followed by a technique insight that defines quantile as a solution of minimization problem. Then, Section 2.2 introduces the conditional quantiles and QR, and provides computational methods to obtain QR estimates. Section 2.3 explains the theoretical aspects of QR model including asymptotic properties and relevant statistical inferences. Finally, some existing methods of QR model for discrete data will be briefly outlined in Section 2.7.

2.1 Unconditional Quantiles and Quantile Functions

2.1.1 CDFs and Unconditional Quantiles

Let X be any real-valued random variable. Its cumulative distribution function (CDF) , denoted as $F_X(x)$,

$$F_X(x) = P(X \leq x). \tag{2.1}$$

For continuous random variables the CDF can be expressed as

$$F_X(x) = P(X \leq x) = \int_{-\infty}^{x} f_X(t)dt. \tag{2.2}$$

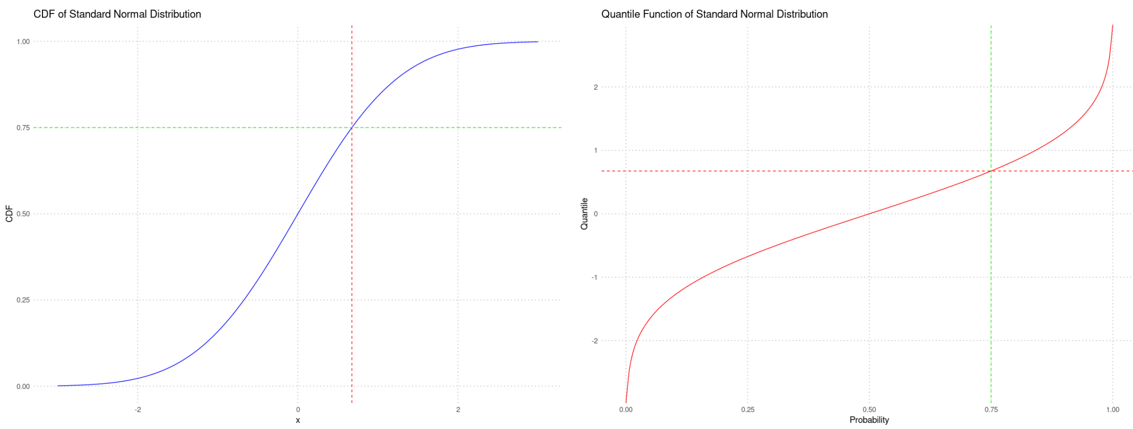
The τ^{th} quantile of the random variable X is the value x such that $P(X \leq x) = \tau$, where $\tau \in (0,1)$. The quantile function is hence the inverse of the CDF

$$Q_X(\tau) \equiv F_X^{-1}(\tau) = \inf\{x : F_X(x) \geq \tau\}. \tag{2.3}$$

Figure 2.1 shows the CDF and quantiles of random variable X following standard normal distribution.

Some properties of quantile function $F^{-1}(\tau)$ are collected:

1. quantile function $F^{-1}(\tau)$ is non-decreasing, if $\tau_1 \leq \tau_2$, then $F^{-1}(\tau_1) \leq F^{-1}(\tau_2)$;
2. it is right-continuous, for any τ, $F^{-1}(\tau) = \lim_{\epsilon \to 0^+} F^{-1}(\tau + \epsilon)$;
3. for $x \in \mathbb{R}$, $F^{-1}(F(x)) \leq x$;
4. for $\tau \in (0,1)$, $F(F^{-1}(\tau)) \geq \tau$;
5. if F is strictly increasing in a neighborhood of $Q(\tau) = F^{-1}(\tau)$, then $F(F^{-1}(\tau))$ and $F^{-1}(F(Q(\tau)) = Q(\tau)$, and
6. $F(x) \geq \tau$ if and only if $x \geq F^{-1}(\tau)$.



(a) CDF

(b) Quantiles

Figure 2.1: CDF and Quantils of Standard Normal Distribution.

2.1.2 Empirical CDFs and Sample Quantiles

Let $(X_1, ..., X_n)$ be an i.i.d. random sample of size n from a CDF F, and the empirical distribution function

$$F_n(x) = \frac{1}{n}\sum_{i=1}^{n} I_{\{X_i \leq x\}}. \tag{2.4}$$

Then, the τ^{th} sample quantile is defined by

$$\hat{Q}_X(\tau) = F_n^{-1}(\tau). \tag{2.5}$$

It is easy to see that $\hat{Q}_X(0) = F_n^{-1}(0) = \min(X_1, ..., X_n)$ is the sample minimum and $\hat{Q}_X(1) = F_n^{-1}(1) = \max(X_1, ..., X_n)$ is the sample maximum. More generally, sample quantiles can be obtained through order statistics.

Let $(X_{(1)}, X_{(2)}, ..., X_{(n)})$ be the order statistics, then the k/n-th sample quantile is the k-th order statistics corresponding to the sample, $(X_1, ..., X_n)$.

For i.i.d. sample $(X_1, ..., X_n)$ drawn from a continuous F with density f, if $f(Q(\tau)) >$

0, then

$$\begin{aligned}\sqrt{n}(\hat{Q}(\tau) - Q(\tau)) \\ = \sqrt{n}(F_n^{-1}(\tau) - F^{-1}(\tau)) \xrightarrow{D} \mathcal{N}(0, \sigma_\tau^2),\end{aligned} \tag{2.6}$$

given stochastic equicontinuity of the empirical processes, see (Andrews, 1986; Walker, 1968). In the limiting distribution,

$$\sigma_\tau^2 = \frac{\tau(1-\tau)}{f(Q(\tau))^2}.$$

2.1.3 Unconditional Quantiles as a Solution of a Minimization Problem

The method of inverse CDF is not applicable in two cases: distributions does not have closed form CDF (e.g., Tweedie), and quantiles of distribution are conditional. Alternatively, quantiles of distribution can be obtained by minimizing a certain objective function.

Let Y be a generic random variable and and μ be the population mean that minimizes the mean squared deviations $E(Y-\mu)^2$. The mean of distribution solves the minimization problem

$$\mu = \arg\min_{c \in \mathbb{R}} E(Y-c)^2, \tag{2.7}$$

which shows that mean of a population, as a location of a distribution, can be found as a solution of a minimization problem.

Use this rationale and apply the multipliers τ and $(\tau-1)$ properly, a generalized method to define quantiles as a certain minimization problem was established by

Koenker and Bassett (1978), who defined the check loss function as follows

$$\rho_\tau(u) = u(\tau - I_{\{u<0\}}) = \begin{cases} \tau \cdot u & \text{if} \quad u \geq 0 \\ (\tau - 1) \cdot u & \text{if} \quad u < 0, \end{cases} \tag{2.8}$$

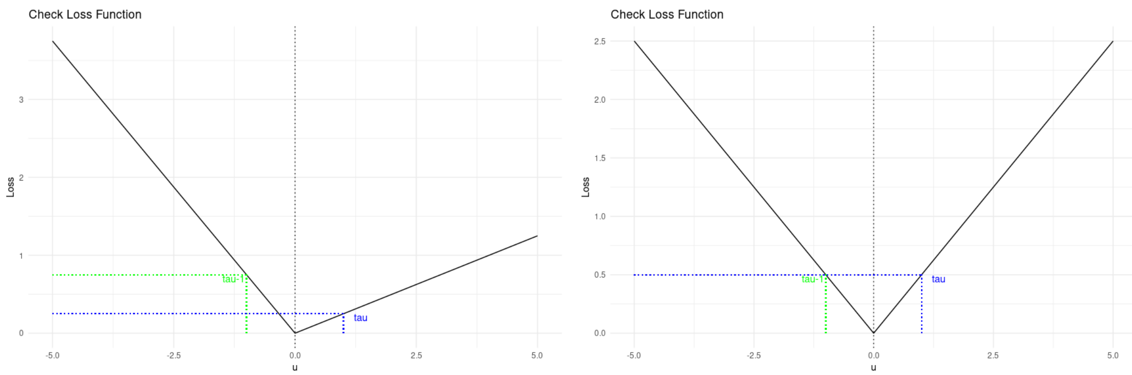
where $\tau \in (0, 1)$ and $I(\cdot)$ is an indicator function that equals 1 if the condition is true, or 0 otherwise. Check loss function assigns τ as a weight to u: if u is zero or positive, the weight is τ; if u is negative, the weight is $\tau - 1$. Let quantile levels τ be 0.25 or 0.50 be examples. For $\tau = 0.25$,

$$\rho_{0.25}(u) = \begin{cases} -0.75u, & \text{if } u < 0 \\ 0.25u, & \text{if } u \geq 0; \end{cases}$$

for $\tau = 0.5$,

$$\begin{aligned} \rho_{0.5}(u) &= \begin{cases} -0.5u, & \text{if } u < 0 \\ 0.5u, & \text{if } u \geq 0 \end{cases} \\ &= 0.5|u|. \end{aligned}$$

Two check loss functions are plotted in Figure 2.8, showing that check loss function is piecewise linear. Let $f_Y(y)$ be the PDF of a continuous variable Y, and τ be quantile



(a) $\tau = 0.25$ (b) $\tau = 0.5$

Figure 2.2: Check Loss Function for Different Quantile levels τ.

level. We are interested in minimizing the following expectation of check loss function

$$
\begin{aligned}
E[\rho_\tau(Y-c)] &= \int_{-\infty}^{\infty} \rho_\tau(y-c) f_Y(y) dy \\
&= (\tau - 1) \int_{-\infty}^{c} (y-c) f_Y(y) dy + \tau \int_{c}^{\infty} (y-c) f_Y(y) dy \\
&= (\tau - 1) \int_{-\infty}^{c} (y-c) dF_Y(y) + \tau \int_{c}^{\infty} (y-c) dF_Y(y) \\
&= (\tau - 1) \int_{-\infty}^{c} y dF_Y(y) + \frac{\partial}{\partial c} \tau \int_{c}^{\infty} y dF_Y(y) \\
&\quad - c\Big[(\tau - 1) \int_{-\infty}^{c} dF_Y(y) + \tau \int_{c}^{\infty} dF_Y(y)\Big] \\
&= \tau E(Y) - \int_{-\infty}^{c} y dF_Y(y) - c\tau - c \int_{-\infty}^{c} dF_Y(y).
\end{aligned}
\tag{2.9}
$$

Partially differentiating the above equation with respect to c, and setting the partial differential equation equal to zero,

$$
\begin{aligned}
\frac{\partial}{\partial c} E[\rho_\tau(Y-c)] &= \frac{\partial}{\partial c}\Big[\tau E(Y) - \int_{-\infty}^{c} y dF_Y(y) - c\tau - c \int_{-\infty}^{c} dF_Y(y)\Big] \\
&= \tau - \int_{-\infty}^{c} dF_Y(y) \\
&= \tau - F_Y(c) \\
&\overset{\text{set}}{=} 0,
\end{aligned}
\tag{2.10}
$$

hence,

$$F_Y(c) = \tau \quad \Rightarrow \quad c = F_Y^{-1}(\tau).$$

$E[\rho_\tau(Y - c)]$ is globally minimized at the point $c = F_Y^{-1}(\tau)$ if the response variable has a continuous distribution.

A special case is the median of distribution, i.e., $\tau = 0.5^{\text{th}}$. As we can see in Figure 2.2b, $\tau = 0.5$ is a special case where the check loss function is symmetric and can be simplifed as an absolute value function. When $\tau = 0.5$, it follows that the check loss function $\rho_{0.5}(Y - c) = 0.5|Y - c|$. Then, the minimization of expected check function $E[\rho_{0.5}(Y - c)]$ can be written as

$$\arg\min_{c\in\mathbb{R}} E[\rho_{0.5}(Y - c)] \equiv \arg\min_{c\in\mathbb{R}} 0.5E|Y - c| \equiv \arg\min_{c\in\mathbb{R}} E|Y - c|.$$

Therefore, median of distribution can be alternatively found by minimizing the expected absolute value of deviation, that is

$$median(Y) = \arg\min_{c\in\mathbb{R}} E|Y - c|. \tag{2.11}$$

Let $Y = (y_1, ..., y_n)$ be a given sample, the τ^{th} sample quantile of Y is the minimizer of the following function,

$$\hat{Q}_Y(\tau) = \arg\min_{c\in\mathbb{R}} \frac{1}{n} \sum_{i=1}^{n} \rho_\tau(y_i - c), \tag{2.12}$$

where $\tau \in (0, 1)$ is the quantile level.

2.2 Quantile Regression Model: Estimation

2.2.1 Conditional Quantiles and QR Model

Let $Y = \{y_1, ..., y_n\}$ be the observations, and $\boldsymbol{X}_{[n\times p]} = \{\boldsymbol{x}_1, ..., \boldsymbol{x}_n\}^T$ be the design matrix. The linear mean regression is defined as

$$y_i = \boldsymbol{x}_i'\boldsymbol{\beta} + \varepsilon_i, \quad i = 1, ..., n,$$

where $\boldsymbol{\beta}$ is a vector of regression coefficients and ε_i is the i-th error. Under the assumptions that $\varepsilon_i \sim N(0, \sigma^2)$ and are iid, by replacing μ in Equation 2.7 with conditional expectation $E[Y|\boldsymbol{X} = \boldsymbol{x}] = \boldsymbol{x}_i'\boldsymbol{\beta}$. The regression coefficients β can be estimated through the least squares method

$$\hat{\beta} = \arg\min_{\beta\in\mathbb{R}} \frac{1}{n}\sum_{i=1}^{n}(y_i - \boldsymbol{x}_i'\boldsymbol{\beta})^2. \tag{2.13}$$

If the squared errors are replaced with errors defined through a check loss function, the regression model will be extended to quantile regression.

Recall that the τ-th quantile of the response Y, denoted as $Q_Y(\tau)$, is estimated through the minimization problem

$$\hat{Q}_Y(\tau) \equiv \arg\min_{q\in\mathbb{R}} \frac{1}{n}\sum_{i=1}^{n}\rho_\tau(Y_i - q). \tag{2.14}$$

Assuming that covariates $\boldsymbol{X} = \boldsymbol{x}$ are collected and the conditional quantile function $Q_{Y_i}(\tau|\boldsymbol{X}_i)$ of Y_i is to be modeled. We have

$$Y_i = Q_{Y_i}(\tau|\boldsymbol{X_i}) + \varepsilon_i,\ i = 1, 2, \cdots, n,$$

where $\varepsilon_i's$ are independent and follow a continuous distribution with the τth quantile equals 0. The conditional quantile function can be estimated by the minimization problem

$$\hat{Q}_Y(\tau|\boldsymbol{x}) \equiv \arg \min_{q_Y(\tau,\boldsymbol{X})\in\mathbb{R}} \frac{1}{n}\sum_{i=1}^{n}\rho_\tau(Y - q_Y(\tau,\boldsymbol{X})), \tag{2.15}$$

where $q_Y(\tau,\boldsymbol{X}) = Q_Y(\tau|\boldsymbol{x})$. By letting conditional quantile function be $Q_Y(\tau|\boldsymbol{x}) = \boldsymbol{x}'\boldsymbol{\beta}_\tau$, the estimated QR coefficient, denoted as $\hat{\boldsymbol{\beta}}_\tau$, solves the following minimization problem

$$\hat{\beta}_\tau \equiv \arg \min_{\beta_\tau\in\mathbb{R}} \frac{1}{n}\sum_{i=1}^{n}\rho_\tau(y_i - \boldsymbol{x}'\boldsymbol{\beta}_\tau), \tag{2.16}$$

on the sample level. The estimated τ^{th} quantile of y_i, denoted as $\hat{Q}_{y_i}(\tau|\boldsymbol{x}_i)$, is

$$\hat{Q}_{y_i}(\tau|\boldsymbol{x}_i) = \boldsymbol{x}_i'\hat{\boldsymbol{\beta}}_{\boldsymbol{\tau}}. \tag{2.17}$$

2.2.2 Estimation of QR coefficients

The estimation of the QR coefficients is going through numerical solutions of relevant minimization problems for most of the practical cases. For a linear QR

$$Q_Y(\tau|\boldsymbol{X}) = \boldsymbol{X}'\boldsymbol{\beta}_\tau, \tag{2.18}$$

the objective function is

$$R(\boldsymbol{\beta}_\tau) = d_\tau(\boldsymbol{y}, \hat{\boldsymbol{y}}(\beta_\tau)) = \sum_{i=1}^{n}\rho_\tau(y_i - \boldsymbol{x}_i'\boldsymbol{\beta}_\tau). \tag{2.19}$$

It is piecewise linear and continuous, but not continuously differentiable. It has directional derivatives on all directions though. For a direction $\boldsymbol{w}$, the directional derivative of $R(\boldsymbol{\beta}_\tau)$ is given by

$$\begin{aligned}
\frac{d}{dt}R(\boldsymbol{\beta}_\tau, \boldsymbol{w}) &= \frac{d}{dt}R(\boldsymbol{\beta}_\tau + t\boldsymbol{w})\Big|_{t=0} \\
&= -\sum_{i=1}^{n}\left[I(y_i - \boldsymbol{x}_i'\boldsymbol{\beta}_\tau - \boldsymbol{x}_i't\boldsymbol{w}) - I(y_i - \boldsymbol{x}_i'\boldsymbol{\beta}_\tau - \boldsymbol{x}_i't\boldsymbol{w} < 0)\right]\Big|_{t=0} \qquad (2.20)\\
&= -\sum_{i=1}^{n}\psi_t^*(y_i - \boldsymbol{x}_i'\boldsymbol{\beta}_\tau, -\boldsymbol{x}_i'\boldsymbol{w})\boldsymbol{x}_i'\boldsymbol{w},
\end{aligned}$$

where

$$\psi_t^*(u, v) = \begin{cases} \tau - I(u < 0) & \text{if } u \neq 0, \\ \tau - I(v < 0) & \text{if } u = 0. \end{cases} \qquad (2.21)$$

If the directional derivatives are all non-negative, where $\frac{d}{dt}R(\hat{\boldsymbol{\beta}}_\tau, \boldsymbol{w}) \geq 0$ for all $\boldsymbol{w} \in \mathbb{R}^p$ with $\|\boldsymbol{w}\| = 1$ at a point $\hat{\boldsymbol{\beta}}_\tau$, then $\hat{\boldsymbol{\beta}}_\tau$ minimizes the objective function $R(\boldsymbol{\beta}_\tau)$.

An alternative computational algorithm to obtain the QR estimate when $\tau = 0.5$ is given by Wagner (1959), who showed that the least absolute deviations can be reformulated as a certain linear programming. Linear programming problem aims to find a vector $\boldsymbol{x}^* \in \mathbb{R}_+^n$ that minimizes or maximizes the value of a given objective function among all $\boldsymbol{x}^* \in \mathbb{R}_+^n$ that satisfies a given system of linear equation or inequality. For other conditional quantiles of a distribution, consider the regression model in Equation 2.18, and minimization problem in Equation 2.19 can be transformed to a linear programming problem. Let us denote the $[c]_+$ as the non-negative part of c. By letting

$$\begin{aligned}
\boldsymbol{u} &= [\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta}_{\boldsymbol{\tau}}]_+ \\
\boldsymbol{v} &= [\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta}_{\boldsymbol{\tau}}]_-,
\end{aligned}$$

Koenker (2005) showed that the QR minimization problem can be reformulated as

$$\min_{\boldsymbol{\beta_\tau}}\{\tau\mathbf{1}^T\boldsymbol{u}+(1-\tau)\mathbf{1}^T\boldsymbol{v}|\boldsymbol{y}=\boldsymbol{X}\boldsymbol{\beta_\tau}+\boldsymbol{u}-\boldsymbol{v},\{\boldsymbol{u},\boldsymbol{v}\}\in\mathbb{R}^n_+\}. \tag{2.22}$$

Furthermore, define

$$B=[\boldsymbol{X}-\boldsymbol{X}\boldsymbol{I},-\boldsymbol{I}],$$

and

$$\psi=\begin{bmatrix}[\boldsymbol{\beta_\tau}]_+\\ [-\boldsymbol{\beta_\tau}]_+\\ [\boldsymbol{y}-\boldsymbol{X}\boldsymbol{\beta_\tau}]_+\\ [\boldsymbol{X}\boldsymbol{\beta_\tau}-\boldsymbol{y}]_+\end{bmatrix},$$

$$\boldsymbol{d}=\begin{bmatrix}\mathbf{0}_{[p]}\\ \mathbf{0}_{[p]}\\ \tau\mathbf{1}_{[n]}\\ (1-\tau)\mathbf{1}_{[n]}\end{bmatrix}.$$

In this standard linear programming problem, the primal formulation is

$$\begin{aligned}\min_{\psi}\quad & \boldsymbol{d}^T\psi\\ \text{subject to}\quad & \boldsymbol{B}\psi=\boldsymbol{y}\\ & \theta\geq 0.\end{aligned}$$

Therefore, its dual counterpart becomes

$$\begin{aligned}\max_{\boldsymbol{d}}\quad & \boldsymbol{y}^T\boldsymbol{z}\\ \text{subject to}\quad & \boldsymbol{B}^T\boldsymbol{z}\leq\boldsymbol{d}.\end{aligned}$$

By a simple rearrangement, the linear programming problem can be reformulated as

$$\max_{\boldsymbol{z}}\{\boldsymbol{y}^T\boldsymbol{z}|\boldsymbol{X}^T\boldsymbol{z}=\mathbf{0},\boldsymbol{z}\in[-1,+1]^n\}.$$

By adding $\boldsymbol{X}^T\mathbf{1}$ and multipytng it by $(1-\tau)$, the above equation $\boldsymbol{X}^T\boldsymbol{z} = \mathbf{0}$ is equivalent to

$$(1-\tau)\boldsymbol{X}^T\boldsymbol{z}+(1-\tau)\boldsymbol{X}^T\mathbf{1}=(1-\tau)\boldsymbol{X}^T\mathbf{1}.$$

It follows that the expression of dual problems is

$$\max_{\boldsymbol{z}}\{\boldsymbol{y}^T\boldsymbol{z}|\boldsymbol{X}^T\boldsymbol{z}=(1-\tau)\boldsymbol{X}^T\mathbf{1},\boldsymbol{z}\in[0,1]^n\}. \tag{2.23}$$

Overall, the formulation shown in Equation 2.22 is the general formulation of linear programming for QR. The formulation shown in Equation 2.23 provides extra computational convenience.

2.3 Quantile Regression Model: Inference

In this section, asymptotic properties of QR estimator, estimated standard errors, confidence intervals, hypothesis tests and their formulas are discussed.

2.3.1 Asymptotic Properties

Under some regularity conditions, the estimator of QR coefficients converges in distribution to a normal random vector,

$$\sqrt{n}(\hat{\boldsymbol{\beta}}_{\boldsymbol{\tau}}-\boldsymbol{\beta}_{\boldsymbol{\tau}})\xrightarrow{D}\mathcal{N}(\mathbf{0},\Sigma_n), \tag{2.24}$$

where Σ_n is the asymptotic covariance matrix. Let $(y_i, X_i), i = 1, 2, ..., n$ be a sample, the QR model can be written in a linear form as

$$Q_{y_i}(\tau|\boldsymbol{X_i}) = \sum_{j=1}^{p} \beta_{\tau j} x_{ij}, \tag{2.25}$$

and

$$y_i = \sum_{j=1}^{p} \beta_{\tau j} x_{ij} + \varepsilon_{\tau j}.$$

where $\varepsilon_{\tau j}$ follows a continuous distribution with the τ^{th} quantile equals to zero.

For i.i.d. errors case, the distribution of ε_τ is strictly positive on the given quantile function $f_{\varepsilon_\tau}(F_{\varepsilon_\tau}^{-1}(\tau)) > 0$. Davino et al. (2013, pp. 66–71) shows the estimated QR coefficients $\hat{\boldsymbol{\beta}}_{\boldsymbol{\tau}}$ is asymptotically distributed as

$$\sqrt{n}(\hat{\boldsymbol{\beta}}_{\boldsymbol{\tau}} - \boldsymbol{\beta}_{\boldsymbol{\tau}}) \xrightarrow{D} \mathcal{N}(\boldsymbol{0}, \omega^2(\tau)\boldsymbol{D}^{-1}), \tag{2.26}$$

where

$$\omega^2(\tau) = \frac{\tau(1-\tau)}{(f_{\varepsilon_\tau}(F_{\varepsilon_\tau}^{-1}(\tau)))^2}$$

is the scale parameter as a function of sparsity function (i.e., $s = 1/f_{\varepsilon_\tau}(F_{\varepsilon_\tau}^{-1}(\tau))$), and the matrix

$$\boldsymbol{D} = \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \boldsymbol{x_i}'\boldsymbol{x_i}$$

is a positive definite matrix. The result yields to the asymptotic covariance matrix of $\hat{\boldsymbol{\beta}}_{\boldsymbol{\tau}}$,

$$\Sigma_{\hat{\beta}_\tau} = \frac{1}{n}\omega^2(\tau)\boldsymbol{D}^{-1}. \tag{2.27}$$

It is worth noting that, however, $f_{\varepsilon_\tau}(F_{\varepsilon_\tau}^{-1}(\tau))$ is the probability density of error term ε_τ at point of quantile $F_{\varepsilon_\tau}^{-1}(\tau)$, which is unknown. To estimate this density term, Siddiqui (1960) adapted the inverse density function, $1/f_{\varepsilon_\tau} = dQ_{\varepsilon_\tau}(\tau)/d\tau$. This function can be estimated by sample quantiles $\hat{Q}_{\varepsilon_\tau}(\tau + h)$ and $\hat{Q}_{\varepsilon_\tau}(\tau - h)$ for sufficiently small h, where the sample quantiles are calculated by residuals $\hat{\varepsilon}_{\tau i} = y_i - \sum_{j=1}^{p} \hat{\beta}_{\tau j} x_{ij}$. This gives the estimated scale parameter $\hat{\omega}$. Thus, the estimated covariance matrix of $\hat{\boldsymbol{\beta}}_{\boldsymbol{\tau}}$ can be obtained as

$$\hat{\Sigma}_{\hat{\beta}_\tau} = \frac{1}{n}\hat{\omega}^2(\tau)\hat{\boldsymbol{D}}^{-1}, \tag{2.28}$$

where $\hat{\boldsymbol{D}} = \frac{1}{n}\sum_{i=1}^{n} \boldsymbol{x_i}'\boldsymbol{x_i}$.

2.3.2 Standard Errors and Confidence Intervals

Statistical inference of QR coefficients starts from standard errors of its estimated coefficients β_τ since confidence intervals and hypothesis tests are based on a valid standard error. Standard errors in QR measure the variability or uncertainty of estimated QR coefficients on different quantiles of the response. The standard errors link directly to the estimated asymptotic covariance matrix of QR estimators. Let the estimated covariance matrix $\hat{\Sigma}_{\hat{\beta}_\tau}$ in Subsection 2.3.1 be a $p \times p$ matrix,

$$\Sigma_{\hat{\beta}_\tau} = \begin{pmatrix} \hat{\sigma}^2_{\hat{\beta}_{\tau 1}} & \hat{\sigma}_{\hat{\beta}_{\tau 1}}\hat{\sigma}_{\hat{\beta}_{\tau 2}} & \cdots & \hat{\sigma}_{\hat{\beta}_{\tau 1}}\hat{\sigma}_{\hat{\beta}_{\tau p}} \\ \hat{\sigma}_{\hat{\beta}_{\tau 2}}\hat{\sigma}_{\hat{\beta}_{\tau 1}} & \hat{\sigma}^2_{\hat{\beta}_{\tau 2}} & \cdots & \hat{\sigma}_{\hat{\beta}_{\tau 2}}\hat{\sigma}_{\hat{\beta}_{\tau p}} \\ \vdots & \vdots & \ddots & \vdots \\ \hat{\sigma}_{\hat{\beta}_{\tau p}}\hat{\sigma}_{\hat{\beta}_{\tau 1}} & \hat{\sigma}_{\hat{\beta}_{\tau p}}\hat{\sigma}_{\hat{\beta}_{\tau 2}} & \cdots & \hat{\sigma}^2_{\hat{\beta}_{\tau p}} \end{pmatrix}. \tag{2.29}$$

Simply taking the square root of diagonal elements of the estimated asymptotic covariance matrix returns the corresponding standard error of QR estimator, that is,

$$s.e.(\hat{\beta}_{\tau j}) = \hat{\sigma}_{\hat{\beta}_{\tau j}}, \quad j = 1, ..., p. \tag{2.30}$$

Based on the standard errors, the significance of co-variates effect on dependent variable can be measured. Confidence interval of a quantile regression coefficient provides a range of value with a confidence level that the true coefficient would be covered, which assesses the statistical significance and reliability of each QR estimator. The confidence interval of QR coefficient can be constructed by standard errors. For a confidence level of $(1-\alpha)$, the corresponding confidence interval is given by

$$C.I.: \hat{\beta}_{\tau j} \pm z_{\alpha/2} s.e.(\hat{\beta}_{\tau j}), \quad j = 1, ..., p. \tag{2.31}$$

2.3.3 Parametric Hypothesis Tests

Parametric hypothesis tests whether the independent variables significantly affect the distribution of dependent variable at different quantiles. We introduce some important hypothesis tests for QR model. Without loss of generality, illustrations are based on the following QR model,

$$Q_{y_i}(\tau|\boldsymbol{x}_i) = \beta_{\tau 0} + \beta_{\tau 1} x_{i1} + ... + \beta_{\tau q} x_{iq} + ... + \beta_{\tau p} x_{ip},$$

where $i = 1, ..., n$ and $\beta_{\tau q}$ is the q-th QR coefficient, $q = 1, ..., p$.

Let null hypothesis be

$$H_0 : \beta_{\tau j} = 0,$$

we reject H_0 if $|(\hat{\beta}_{\tau j} - \beta_0)/s.e.(\hat{\beta}_{\tau j})| > z_{\alpha/2}$ and conclude that $\hat{\beta}_{\tau j}$ is statistically significant.

Next, Davino et al. (2013, pp. 84–86) introduced the likelihood ratio (LR) test to compare the fit of two nested QR models. The LR test evaluates whether including or excluding some predictors will significantly change the performance of the QR model. Instead of obtaining likelihood function of QR, the LR test statistic is estimated by the difference of two sums of the weighted absolute deviations, which is given by

$$LR = 2\omega^{-1}(\tilde{V}(\tau) - \hat{V}(\tau)). \tag{2.32}$$

Inside this LR test statistic, $\tilde{V}(\tau)$ is the absolute deviation of restricted model excluding some dependent variables, and $\hat{V}(\tau)$ is the one that includes all dependent variables. The scale parameter ω is given above. The LR follows asymptotically a χ^2 distribution with degrees freedom that equals the numbers of QR coefficients that are zero. To test whether q QR coefficients equal to zero, let null hypothesis be

$$H_0 : \beta_{\tau 1} = ... = \beta_{\tau q} = 0,$$

and the absolute deviations are obtained as

$$\tilde{V}(\tau) = \sum_{i=1}^{n} |y_i - \sum_{j=q+1}^{p} \hat{\beta}_{\tau j} x_{ij}|,$$

and

$$\hat{V}(\tau) = \sum_{i=1}^{n} |y_i - \sum_{j=1}^{p} \hat{\beta}_{\tau j} x_{ij}|,$$

separately. We reject the null hypothesis if $2\omega^{-1}(\tilde{V}(\tau) - \hat{V}(\tau)) > \chi^2_{(2,\alpha)}$ at the significance level α and conclude that the corresponding independent variable should not be dropped.

The Lagrange multiplier (LM) test is another hypothesis test which is very similar to LR test. It can more accurately determine which independent variables can be excluded when values are very small. The LM test statistic

$$LM = \boldsymbol{g}'[\boldsymbol{D}^{22}]^{-1}\boldsymbol{g}, \tag{2.33}$$

where $\boldsymbol{g}$ is the gradient of the objective function excluding some independent variables, and $\boldsymbol{D}_{22}$ is a sub-matrix of independent variables, both under null hypothesis. The LM statistic asymptotically follows a χ^2 distribution with the degree of freedom equaling the numbers of QR coefficients equal to zero. To test the hypothesis

$$H_0 : \beta_{\tau(q+1)} = ... = \beta_{\tau p} = 0,$$

we firstly obtain the gradient of the objective function with respect to regression coefficient $\boldsymbol{\beta}_\tau$, that is

$$\boldsymbol{g} = \begin{pmatrix} \psi(\hat{\varepsilon}_{\tau 1}) \\ \vdots \\ \psi(\hat{\varepsilon}_{\tau p}) \end{pmatrix} = \begin{pmatrix} \tau - I(\hat{\varepsilon}_{\tau 1} < 0) \\ \vdots \\ \tau - I(\hat{\varepsilon}_{\tau 5} < 0) \end{pmatrix},$$

where $\hat{\varepsilon}_{\tau i} = \sum_{j=1}^{5} y_i - \hat{\beta}_{\tau j} x_{ij}, j = 1, ..., p$. Then, $\boldsymbol{D} = \boldsymbol{X}'\boldsymbol{X}$ is a matrix in quadratic form consisted of the independent variables. Partition the matrix,

$$\boldsymbol{D} = \left(\begin{array}{ccc:ccc} \sum_{i=1}^n x'_{i1}x_{i1} & \cdots & \sum_{i=1}^n x'_{i1}x_{iq} & \sum_{i=1}^n x'_{i1}x_{i(q+1)} & \cdots & \sum_{i=1}^n x'_{i1}x_{ip} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^n x'_{iq}x_{i1} & \cdots & \sum_{i=1}^n x'_{iq}x_{iq} & \sum_{i=1}^n x'_{iq}x_{i(q+1)} & \cdots & \sum_{i=1}^n x'_{iq}x_{ip} \\ \hdashline \sum_{i=1}^n x'_{i(q+1)}x_{i1} & \cdots & \sum_{i=1}^n x'_{iq+1)}x_{iq} & \sum_{i=1}^n x'_{iq+1)}x_{i(q+1)} & \cdots & \sum_{i=1}^n x'_{iq+1)}x_{ip} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^n x'_{ip}x_{i1} & \cdots & \sum_{i=1}^n x'_{ip}x_{iq} & \sum_{i=1}^n x'_{ip}x_{i(q+1)} & \cdots & \sum_{i=1}^n x'_{ip}x_{ip} \end{array}\right)$$

$$= \left(\begin{array}{c:c} \boldsymbol{D}_{11} & \boldsymbol{D}_{12} \\ \hdashline \boldsymbol{D}_{21} & \boldsymbol{D}_{22} \end{array}\right)$$

where sub-matrix $\boldsymbol{D}_{22}$ consists of independent variables with hypothesised value of zero. Its inverse $\boldsymbol{D}^{22} = [\boldsymbol{D}_{22} - \boldsymbol{D}_{21}\boldsymbol{D}_{11}\boldsymbol{D}_{12}]^{-1}$. Then, we reject H_0 if $\boldsymbol{g}'[\boldsymbol{D}^{22}]^{-1}\boldsymbol{g} > \chi^2_{(2,\alpha)}$ at level of significance α, and conclude that these independent variables cannot be dropped.

The Wald test can be also used for variable selection. The advantage is that if the quadratic form $\boldsymbol{D}$ is very close to zero, the Wald test will drop irrelevant independent variables more safely than other tests. The Wald test statistic, denoted by W, is given by

$$W = n\omega^{-2}\hat{\boldsymbol{\beta}}_\tau'[\boldsymbol{D}^{22}]^{-1}\hat{\boldsymbol{\beta}}_\tau. \tag{2.34}$$

Notice that the Wald test also follows a χ^2 distribution with degrees of freedom equaling the numbers of QR coefficients that are hypothesised as in H_0. To test the hypothesis that

$$H_0 : \beta_{\tau 1} = ... = \beta_{\tau q} = 0,$$

we have to estimate the value of QR estimators firstly and then let

$$\hat{\boldsymbol{\beta}}_\tau = \begin{pmatrix} \hat{\beta}_{\tau 1} \\ \vdots \\ \hat{\beta}_{\tau q} \end{pmatrix}$$

and

$$\boldsymbol{D}_{22} = \begin{pmatrix} \sum_{i=1}^n x_{i1}' x_{i1} & \dots & \sum_{i=1}^n x_{i1}' x_{iq} \\ \vdots & \ddots & \vdots \\ \sum_{i=1}^n x_{iq}' x_{i1} & \dots & \sum_{i=1}^n x_{iq}' x_{iq} \end{pmatrix},$$

where the expression of $\boldsymbol{D}_{22}$ is given above. Then, calculate the inverse matrix $(\boldsymbol{D}^{22})^{-1}$, and the Wald statistic can be obtained. Thus, we reject H_0 if $n\omega^{-2}\hat{\boldsymbol{\beta}}_\tau'[\boldsymbol{D}^{22}]^{-1}\hat{\boldsymbol{\beta}}_\tau > \chi^2_{(2,\alpha)}$, at the significance level α, and conclude that these variables cannot be dropped.

2.4 Some Basic Properties of Conditional Quantiles

Equivarience is one of the properties of conditional quantiles. Monotonic transformation of the response or linear transformation of regressors hold some invariance properties. Transformations with equivariance include scale; shift or regression; reparametrization of design; monotonic transformations. Denote $\hat{\beta}_\tau(y, \boldsymbol{X})$ as a QR estimator at the τ^{th} quantile. Some basic properties of equivariance are collected as follows,

Proposition 2.4.1 (Koenker and Bassett, 1978).

Let A be any $p \times p$ non-singular matrix, $\gamma \in \mathbb{R}^p$, and $a > 0$. Then, for any $\tau \in (0,1)$,

1. $\hat{\beta}_\tau(ay, \boldsymbol{X}) = a\hat{\beta}_\tau(y, \boldsymbol{X})$ *(Scale equivariance)*

2. $\hat{\beta}_\tau(-ay, \boldsymbol{X}) = -a\hat{\beta}_{1-\tau}(y, \boldsymbol{X})$ *(Scale equivariance)*

3. $\hat{\beta}_\tau(y + \boldsymbol{X}\gamma, \boldsymbol{X}) = \hat{\beta}_\tau(y, \boldsymbol{X}) + \gamma$ *(shift or regression equivariance)*

4. $\hat{\beta}_\tau(y, \boldsymbol{X}A) = A^{-1}\hat{\beta}_\tau(y, \boldsymbol{X})$ *(equivariance to reparameterization of design).*

Proof.

Let

$$\begin{aligned}\Psi_\tau(\hat{\beta}, y, \boldsymbol{X}) &= \sum_{\{i: y_i \geq \boldsymbol{x}_i^T\hat{\beta}\}} \tau|y_i - \boldsymbol{x}_i^T\hat{\beta}| + \sum_{\{i: y_i < \boldsymbol{x}_i^T\hat{\beta}\}} (1-\tau)|y_i - \boldsymbol{x}_i^T\hat{\beta}| \\ &= \sum_{i=1}^n \left[\tau - 1 + \frac{1}{2} + \frac{1}{2}\text{sgn}(y_i - \boldsymbol{x}_i^T\hat{\beta})\right] |y_i - \boldsymbol{x}_i^T\hat{\beta}|\end{aligned}$$

where $\text{sgn}(u)$ takes value 1, 0, -1 as $u > 0$, $u = 0$, $u < 0$. Now, note that

1. $a\Psi_\tau(\hat{\beta}, y, \boldsymbol{X}) = \Psi_\tau(a\hat{\beta}, ay, \boldsymbol{X})$

2. $a\Psi_\tau(\hat{\beta}, y, \boldsymbol{X}) = \Psi_{1-\tau}(-a\hat{\beta}, -ay, \boldsymbol{X})$

3. $\Psi_\tau(\hat{\beta}, y, \boldsymbol{X}) = \Psi_\tau(\hat{\beta} + \gamma, y + \boldsymbol{X}\gamma, \boldsymbol{X})$

4. $\Psi_\tau(\hat{\beta}, y, \boldsymbol{X}) = \Psi_\tau(A^{-1}\hat{\beta}, y, \boldsymbol{X}A)$.

The proof is complete. □

By Proposition 2.4.1, some transformations of dependent variable can be derived. In order to illustrate the transformations clearly, let us use a simple linear QR model on the τ-th quantile as an example (see Davino et al., 2013, pp. 97–100),

$$Q_{\boldsymbol{y}}(\tau|\boldsymbol{x}) = \beta_{\tau 0} + \beta_{\tau 0}\boldsymbol{x}.$$

The scale equivariance property is useful when modifying the unit of dependent variable, which implies that multiplying dependent variable by positive constant c, the QR coefficients of the transformed QR model can be obtained multiplying the original coefficients by c, that is,

$$Q_{c\boldsymbol{y}}(\tau|\boldsymbol{x}) = cQ_{Y}(\tau|\boldsymbol{x}) = c\beta_{\tau 0} + c\beta_{\tau 1}\boldsymbol{x}. \tag{2.35}$$

The properties holds as well if multiplying the dependent variable by a negative constant d, the QR coefficients of the transformed QR model are complement of original coefficients, that is,

$$Q_{d\boldsymbol{y}}(\tau|\boldsymbol{x}) = dQ_{\boldsymbol{y}}(\tau|\boldsymbol{x}) = d\beta_{(1-\tau),0} + d\beta_{(1-\tau),1}\boldsymbol{x}. \tag{2.36}$$

The shift or regression equivariance property is applied when the expression of dependent variable is a linear combination of independent variables through a certain slope, say γ. Let $\boldsymbol{y}^* = \boldsymbol{y} + \boldsymbol{x}\gamma$, Kuan (2007) shows that

$$Q_{\boldsymbol{y}^*}(\tau|\boldsymbol{x}) = Q_{\boldsymbol{y}^*}(\tau|\boldsymbol{x}) + \boldsymbol{x}\gamma = \beta_{\tau 0} + (\beta_{\tau 1} + \gamma)\boldsymbol{x}. \tag{2.37}$$

The equivariance to reparametrization of design is widely applied in any regression analysis when the matrix of independent variable $\boldsymbol{x}$ is not designed with full column rank. The property states that introducing a non-singular matrix $\boldsymbol{A}$ in the QR model, the new QR coefficients is obtained in a matrix form,

$$Q_{\boldsymbol{y}}(\tau|\boldsymbol{x}\boldsymbol{A}) = \boldsymbol{A}^{-1}\boldsymbol{x}\boldsymbol{\beta}_{\tau}. \tag{2.38}$$

The last property is equivariance of monotonic transformation. The equivariance properties ensure that the statistical inference based on QR model will not be affected by monotone transformation (C. Chen, 2005). The negative skewness can be reduced (Manning, 1998) after transforming the response. Let $h(\cdot)$ be a monotonic function. It follows that for any random variable Y,

$$P(Y \leq y) = P(h(Y) \leq h(y)).$$

This property allows us to converse a linear QR model to non-linear one and vice versa,

$$Q_{h(\boldsymbol{y})}(\tau|\boldsymbol{x}) = h(Q_{\boldsymbol{y}}(\tau|\boldsymbol{x})) = h(\beta_{\tau 0} + \beta_{\tau 1}\boldsymbol{x}). \tag{2.39}$$

For instance, let us consider a very common logarithmic transformation for exponentially growing data, where the non-linear QR model is

$$Q_{\boldsymbol{y}}(\tau|\boldsymbol{x}) = e^{\beta_{\tau 0} + \beta_{\tau 1}\boldsymbol{x}}.$$

Let $h(\cdot) = \log(\cdot)$. The new QR model becomes

$$\begin{aligned} Q_{\log(\boldsymbol{y})}(\tau|\boldsymbol{x}) = \log(Q_{\boldsymbol{y}}(\tau|\boldsymbol{x})) &= \log(e^{\beta_{\tau 0} + \beta_{\tau 1}\boldsymbol{x}}) \\ &= \beta_{\tau 0} + \beta_{\tau 1}\boldsymbol{x}, \end{aligned}$$

the relationship between QR response and predictors is linearized.

2.5 Resampling

Resampling works as a numerical method that obtains the asymptotic inference, alternatively validating the asymptotic theorem in practice. Design matrix bootstrap (Kocherginsky & He, 2007) is a widely used resampling method that involves repeatedly generating a number of samples B with sample size as the original data-set. The bootstrap procedure obtains samples by random sampling with replacement from original data-set. Then, performing QR modeling on bootstrap samples to obtain a set of QR estimate. Consider the j^{th} QR estimator, $j = 1, ..., p$, the average bootstrap QR estimator is given by

$$\overline{\hat{\beta}_j^B(\tau)} = \frac{1}{B}\sum_{b=1}^{B}\hat{\beta}_{b,j}(\tau), \tag{2.40}$$

and the corresponding standard error is

$$s.d.(\hat{\beta}_j^B(\tau)) = \frac{1}{B-1}\sum_{b=1}^{B}(\hat{\beta}_{b,j}(\tau) - \overline{\hat{\beta}_j^B(\tau)}). \tag{2.41}$$

The standard error of bootstrap QR estimators represents the estimate of asymptotic standard error of QR estimators. It is useful for the construction of confidence intervals and hypothesis tests. Moreover, the asymptotic variance of the j^{th} independent variable can be obtained as

$$\hat{V}_j^B = \frac{1}{B}\sum_{b=1}^{B}\big(\hat{\beta}_{b,j}^B(\tau) - \overline{\hat{\beta}_{b,j}^B(\tau)}\big)\big(\hat{\beta}_{b,j}^B(\tau) - \overline{\hat{\beta}_{b,j}^B(\tau)}\big)', \quad j = 1, ..., p. \tag{2.42}$$

Thus, a confidence interval for the j^{th} independent variable is

$$CI : \overline{\hat{\beta}_j^B(\tau)} \pm z_{\alpha/2}s.d.(\hat{\beta}_j^B(\tau)).$$

However, although the resampling methods can provide unbiased estimate of standard error, persistent sample variability and bootstrap resampling variability present, probably misleading the inference (Efron & Tibshirani, 1998).

2.6 Goodness of Fit

Goodness of fit test validates how well the predicted quantile of response corresponds to the observed data. The assessment of goodness of fit for QR model is an extension of R^2 index in mean regression (Koenker & Machado, 1999). For a mean regression model, a bigger value of

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y}_i)^2},$$

implies a better model fitting.

In QR, Gujarati (2003) derived such an index using the ratio of model-related sum of residuals and total sum of residuals. Let

$$\hat{Q}_{y_i}(\tau|\boldsymbol{x}_i) = \sum_{j=1}^{p} \hat{\beta}_j x_{ij},$$

where $\hat{\beta j}$ is the j^{th} QR estimator. The residual absolute sum of weighted residuals is

$$\Theta_{\tau 0} = \sum_{i=1}^{n} \rho_\tau(y_i - \hat{Q}_{y_i}(\tau|\boldsymbol{x}_i)),$$

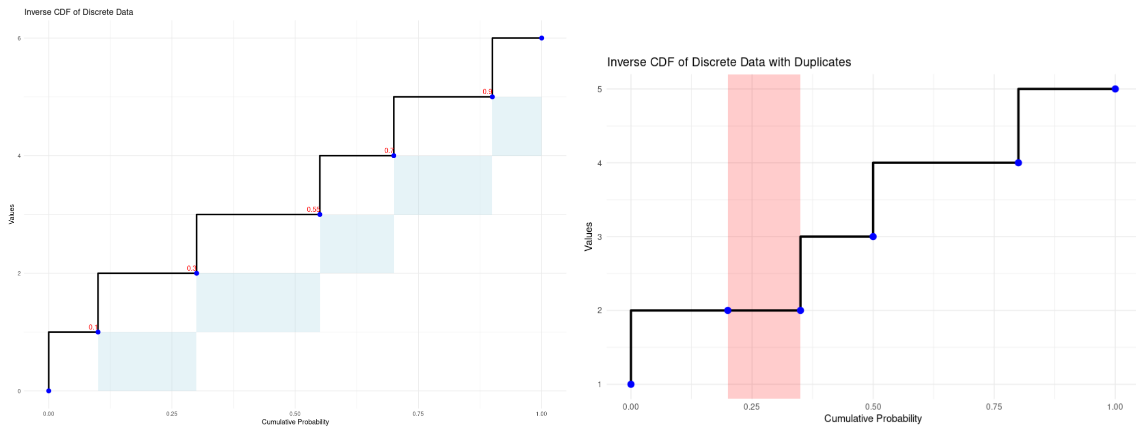
and the total sum of weighted difference between observed value and sample quantile $\hat{Q}_{y_i}(\tau)$ is

$$\Theta_{\tau 1} = \sum_{i=1}^{n} \rho_\tau(y_i - \hat{Q}_{y_i}(\tau)).$$

then, the pseudo R^2 can be obtained by comparing the above two sums of weighted difference, which is given by

$$\text{pseudo } R^2 = 1 - \frac{\Theta_{\tau 0}}{\Theta_{\tau 1}}. \tag{2.43}$$

Notice that since $0 \leq \Theta_{\tau 0} \leq \Theta_{\tau 1}$, pseudo $R^2 \in [0, 1]$. The greater the pseudo R^2, the better the QR model fitting.



(a) Discrete points (b) Duplicated points

Figure 2.3: Inverse CDF of discrete data. Blue and red indicate the undefined quantiles of data

2.7 Existing Smoothing Methods of QR for Discrete Data

The above sections discussed estimating QR model on continuous data. However, traditional QR model is not directly applicable for discrete data due to discontinuity of CDF and potential duplicated values (Jentsch & Leucht, 2015). Figure 2.3 provides two examples for illustration. Figure 2.3a displays the inverse CDF of data $(0, 1, 2, 3, 4, 5)$ with probabilities of $(0.1, 0.2, 0.25, 0.15, 0.2, 0.1)$, respectively. As we can see the inverse CDF of discrete data is step-wise, with horizontal steps and vertical jumps. The blue areas highlight the flat sections of CDF between jumps, which present the CDF does not increase. Within these flat sections, the quantile of data cannot be uniquely defined by the inverse CDF, as any point in the flat section (blue rectangle) is the corresponding quantile of data. Therefore, the vertical jumps do not indicate the behaviours of data between those points uniquely so that quantiles cannot be defined. On the other hand, discrete distribution will violate the one-to-one mapping on quantiles of data points if duplicated values appear. See Figure 2.3b as an example, consider data $(1, 2, 2, 3, 4)$ with probabilities of $(0.2, 0.15, 0.15, 0.3, 0.2)$, respectively, where value of 2 is duplicated. The quantile for cumulative probability between the first occurrence of 2 and the second occurrence of 2 is obscure since the duplicated values across entire range of the cdf, which can be found in red highlighted area. Therefore, a direct application of traditional QR method introduces systematic error into the model, posing challenges on estimating parameters and further inference.

The remainder of this chapter briefly introduces some existing methods of QR for discrete data. The first and second methods smooth the objective function so that the conditional quantile function can be continuous. The third one smooths the counts response directly, which makes the estimation follow traditional QR model.

2.7.1 Asymptotic Maximimun Likelihood Estimator

Efron (1992) proposed an approach of asymptotic maximum likelihood estimator for QR cooping with counts. This method smooths the objective function of QR to define the estimator of parameters. Newey and Powell (1987) introduced the approach of asymptotic maximum likelihood estimator, which estimates the conditional location function for counts that connect to conditional expectiles. Assume that the dependent variable is exponential to independent variables, and let $(y_i, \boldsymbol{x}_i)_{i=1}^n$ be the counts sample. Efron (1992) proposed an asymptotic maximum likelihood estimator by solving an objective function as follows

$$\boldsymbol{\beta}_w^{AML} = \arg\min_{\boldsymbol{\beta}\in\mathbb{R}^p} \sum_{i=1}^{n} \big(y_i\boldsymbol{x}_i'\boldsymbol{\beta} - \exp(\boldsymbol{x}_i'\boldsymbol{\beta}) - \ln(y_i!)\big) w^{I\left(y_i > \exp(\boldsymbol{x}_i'\boldsymbol{\beta})\right)}, \tag{2.44}$$

where $w > 0$ and $I(\cdot)$ is an indicator function equals to 1 if the statement is true or 0 elsewhere. It is worth noting that when $w = 1$, Equation 2.44 becomes the Poisson pseudo-likelihood estimator. Moreover, Efron (1992) found that the τ^{th} regression quantile can be obtained by w such that

$$\tau = \frac{1}{n}\sum_{i=1}^{n} I\big(y_i \le \exp(\boldsymbol{x}_i'\boldsymbol{\beta}_w^{AML})\big). \tag{2.45}$$

The advantage of this method is that the percentile of count data is easy to estimate, since the percentile is not restricted to be integers. Thus, the asymptotic properties of estimator follows the traditional QR estimator. However, since $I\big(y_i \le \exp(\boldsymbol{x}_i'\boldsymbol{\beta}_w^{AML})\big) = 1$ if $y_i = 0$, the asymptotic maximum likelihood regression percentile cannot be calculated for value of τ that smaller than the proportion of zeros in the data. This implies the potential limitation of asymptotic maximum likelihood estimator that cannot estimate the QR model when data is zero-inflated, such as zero-inflated Poisson and zero-inflated negative binomial distributed responses.

2.7.2 Mid-Quantile Regression for Discrete Responses

Geraci and Farcomeni (2022) introduced a QR method to handle discrete data. Their QR model is built on mid-quantiles of discrete responses, which is the same idea as mid-p-value that offers a unifying theorem for quantile estimation (Lancaster, 1961). Their proposed approach introduces a two-step estimator which can be applied to different type of discrete data. Parzen (1993) introduced mid-cumulative distribution function (mid-CDF) that modifies the standard CDF of discrete variable. Let $F_Y(y)$ be the CDF of discrete random variable Y, the mid-CDF is defined as

$$G_Y(y) \equiv \Pr(Y \le y) - 0.5 \cdot \Pr(Y = y), \tag{2.46}$$

which is a step function due to discontinuity of Y. Let $\mathcal{S}_y = (y_1, ..., y_s)$ be a set of possible values of Y, ordered from smallest to largest, with probability of $p_1, ..., p_s$, respectively. Defining the mid-probabilities $\pi_1 = p_1/2$ and $\pi_j = G(y_j) = \sum_{j=1}^{j-1} p_u +$

$p_j/2$,

$$H_Y(p) = \begin{cases} y_1 & \text{if } p < \pi_1, \\ y_j & \text{if } p = \pi_j, j = 1, \ldots, s, \\ (1-\gamma)y_j + \gamma y_{j+1} & \text{if } p = (1-\gamma)\pi_j + \gamma\pi_{j+1}, 0 < \gamma < 1, j = 1, \ldots, s-1, \\ y_s & \text{if } p > \pi_s, \end{cases}$$

is called the mid-quantile function (mid-QF) (Ma et al., 2011). On the sample level, Equation 2.46 can be written as

$$\hat{G}_Y(y) = \hat{F}_Y(y) - 0.5 \cdot \hat{m}_Y(y), \tag{2.47}$$

where $\hat{m}_Y(y) = n^{-1}\sum_{i=1}^{n} I(Y_i = y)$. Two advantages of mid-quantiles are posed by Geraci and Farcomeni (2022). The first one is that mid-quantiles can be considered as fractional order statistics (Stigler, 1977), thus smoothing the discrete data based on the order statistics. Compared to traditional sample quantiles, the approach of mid-quantiles is more sensitive to identify the difference between discrete distributions. Therefore, the impact of duplicated values on quantifying quantile of discrete distributions is significantly reduced. Furthermore the mid-quantiles can be relabeled so that the quantiles can be obtained. Then, based on the mid-CDF and mid-quantile function, conditional mid-CDF can be defined as

$$G_{Y|X}(y|x) \equiv F_{Y|X}(y|x) - 0.5 \cdot m_{Y|X}(y|x), \tag{2.48}$$

where Y is a random variable in $\mathbb{R}$ and X is a p-dimensional vector of covariates. Let $\pi_j = G_{Y|X}(y_j|x)$, the conditional mid-quantile function can be defined as the piece-wise linearly connecting inversion $G^{-1}_{Y|X}(\pi_j|x), j = 1, ..., s$. The quantile-specific regression model can be defined using the above as

$$H_{T(Y)|X}(\tau|X) = \boldsymbol{x}'\boldsymbol{\beta}_\tau, \tag{2.49}$$

where $\boldsymbol{\beta}_\tau$ is a vector of unknown QR coefficients for a given quantile level τ and $T(\cdot)$ is a known monotone and differentiable link function.

The estimation of parameters proceeds in two steps. In the first step, estimating mid-cdf as shown in Equation 2.47 using sample with modified

$$\hat{F}_{Y|X}(y|x) = \frac{n^{-1}\sum_{i=1}^{n} I(Y_i \leq y)K_\lambda(X_i, x)}{\hat{\delta}_X(x)},$$

where K_λ is kernel with bandwidth λ and $\hat{\delta}_X(x)$ is the kernel estimator of the marginal density of X. Then, the second step involves minimizing their proposed objective function to obtain estimator. Defining function $G^c_{Y|X}(y|x)$ as a function interpolating sample points $(z_i, G^c_{Y|X}(z_i|x))$ where z_i denotes the distinct value in $\mathcal{S}_Y$ and the ordinates has been obtained in the first step. To estimate $\boldsymbol{\beta}_\tau$ in Equation 2.49, the proposed objective function is defined as

$$\psi_n(\boldsymbol{\beta};\tau) = \frac{1}{n}\sum_{i=1}^{n}\left\{\tau - G^c_{Y|X}(T^{-1}(\boldsymbol{x}_i'\boldsymbol{\beta}|\boldsymbol{x}_i)\right\}^2. \tag{2.50}$$

The mid-QR estimator is the solution of minimizing the objective function,

$$\hat{\boldsymbol{\beta}}_\tau = \arg\min_{\boldsymbol{\beta}\in\mathbb{R}^p} \psi_n(\boldsymbol{\beta};\tau). \tag{2.51}$$

The approach of mid-QR is an extension of inverse CDF technique for marginal quantiles proposed by De Backer et al. (2020) for fitting QR models for censored data. The asymptotic properties follow the traditional QR model for continuous data (Ma et al., 2011).

2.7.3 Quantile Regression with Jittered Data

Machado and Silva (2005) proposed a QR method for counts called jittering. Jittering is a smoothing process that directly smooths the discrete response. This mathematical approach was initially proposed by Stevens (1950) in the purpose of better data visualization, and its mathematical expression was further specified by Pearson (1950). In order to prevent the data visualizations from over-plotting due to the identical or very similar values of data, adding some reasonably small and random offset to each of the original points makes it easier to view the plot and density of the data. Jittering can be achieved by adding some small random values, negative or positive, to the data points, which is usually called noises or perturbations. One condition is to ensure the one-to-one mapping on manipulated data of the original data. Thus we can see that the idea of jittering is to reach continuity of response and avoid duplicated values by adding random values to counts. Machado and Silva (2005) introduced this smoothing approach into QR models, making the conditional quantile function continuous at any quantile level τ. Since counts are not negative, the additive noise shall be non-negative. Thus, Machado and Santos Silva construct pseudo continuous random variables as new QR responses whose quantile has a one-to-one mapping with the original QR response Y. Let U follow uniform $[0, 1)$, and $Z = Y + U$, where the random variable Z is the continuous response for QR. Under some mild conditions of traditional QR model given by Koenker (2005), let $T(\cdot;\tau)$ be a monotone and differentiable function probably conditional on the quantile of data, and the conditional jittered QR model is defined as

$$Q_{T(Z;\tau)}(\tau|\boldsymbol{X}) = \boldsymbol{X}'\boldsymbol{\beta}_\tau \tag{2.52}$$

where $\boldsymbol{X}$ is vector of predictors, and $\boldsymbol{\beta}_\tau$ is a vector of QR coefficients on the quantile level of $\tau \in (0, 1)$. On the sample level, let $(y_i, \boldsymbol{x}_i)_{i=1}^n$ be a data set with size of

n. The continuous response can be constructed by introducing uniform noises $\boldsymbol{u} = (u_1, ..., u_n)$ into observations and defining $z_i = y_i + u_i$ for $i = 1, ..., n$, where each triple of $(y_i, \boldsymbol{x}_i, u_i)_{i=1}^n$ is called one 'jittered sample'. Moreover, the conditional quantile function of Z follows that

$$Q_Z(\tau|\boldsymbol{x}) = Q_Y(\tau|\boldsymbol{x}) + \frac{\tau - \sum_{y=0}^{Q_Y(\tau|\boldsymbol{x})-1} \Pr(Y = y|\boldsymbol{x})}{\Pr(Y = Q_Y(\tau|\boldsymbol{x})|\boldsymbol{x})}. \tag{2.53}$$

The estimation of QR model with jittered data follows that of traditional one given by Koenker and Bassett (1978). For sample $(z_i, \boldsymbol{x}_i)_{i=1}^n$ consisted of jittered response and predictors, the objective function is defined as

$$S_n(\boldsymbol{\beta}) = \frac{1}{n} \sum_{i=1}^{n} \rho_\tau(T(z_i; \tau) - \boldsymbol{x}_i'\boldsymbol{\beta}). \tag{2.54}$$

The estimated QR parameters can be obtained by minimizing the Equation 2.54, which yields the estimator

$$\hat{\boldsymbol{\beta}}_\tau = \arg\min_{\boldsymbol{\beta} \in \mathbb{R}^p} S_n(\boldsymbol{\beta}). \tag{2.55}$$

To solve the minimization problem, techniques introduced in Subsection 2.2.2 can be applied. Also, Geraci (2014) provided an efficient algorithm for obtaining QR estimates available in R by installing external package 'lqmm' and using function 'lqm.counts' as follows,

```
install.package('lqmm')
library(lqmm)
lqm.counts(formula, data, weights = NULL, offset = NULL,
           contrasts = NULL, tau = 0.5, M = 50, zeta = 1e-05,
           B = 0.999, cn = NULL, alpha = 0.05,
           control = list())
```

For further details on R packages and arguments, see The Comprehensive R Archive

Network (CRAN) at https://cran.r-project.org/web/packages/lqmm/index.html. To further enhance the performance of QR model with jittered data, Machado and Silva (2005) further introduced an efficient technique of averaging-out noises that involves generating many jittered samples and taking averaged estimate in order to reduce the bias and variability of QR estimator. This technique will be fully explained in Chapter 3. Our study focuses on Machado and Santos Silva's uniform jittering approach and aims to enrich the choice of additive noises.

Chapter 3

Proposed QR Model with Optimized Jittering Approach

In this chapter, we investigate the method of jittering and propose a method to optimize the performance of this method.

3.1 Extension of Uniform Jittering Method

Jittering method is widely applied when analyzing count data using QR models. In our simulation studies we found that the jittering with uniform random variable on a narrow support tends to give worse QR estimates, and the performance of jittering would be improved if the support of the jittering variable is wider.

As discussed by Machado and Silva (2005), Beta distribution is an alternative to jitter the counts. If $a = b = 1$, the random variable following $\text{Beta}(1, 1)$ is equivalent to $\text{Uniform}(0, 1)$. Those two parameters of Beta distribution allow us to control the shape of Beta variable, where some specific shape of Beta could lead to a better fit than some others could do, when Beta-jittering is applied. Another option is to choose the jittering term from the Tweedie family of distribution, which is shaped by its parameters. In order to enrich the jittering method but prevent it from under- or over-fitting, we introduce an alternative jittering process that utilizes Beta distribution (a, b) and Tweedie distribution (μ, ϕ, p), where p is power. For the Tweedie distribution, we only consider when $1 < p < 2$ (i.e., compound Poisson-Gamma distribution) for its statistical properties. Parameters of the jittering will be selected such that the jittering QR model is optimized.

In the remainder of this chapter, we present the process and implementation of the jittering approaches. Then, we introduce method and algorithms for determining the distribution parameters. Asymptotic properties of jittering QR estimators using proposed method will be discussed.

3.2 Jittering with Tweedie and Beta Noise

in this section, we review relevant properties of Tweedie and Beta distributions, followed by optimized jittering process and corresponding implementation. For clarification, we name the smoothing processes as 'Tweedie jittering' or 'Beta jittering', respectively according to the distribution of perturbation used.

3.2.1 Properties of Tweedie and Beta Distributions

Tweedie distribution and Beta distribution are non-negative distributions that are chosen for perturbing counts. Both of them provide us with flexibility to generate the additive noises. Their parameters define the shape of the variables, which could lead to different performance when jittering.

Tweedie distribution is a member of the exponential dispersion family (Jorgensen, 1987). This statistical distribution is well known as its flexibility and wide applications in all the statistical modelling. Specifically, it is able to model the compound Poisson-Gamma distributed data, where the data are either continuous or discrete. One key feature worth mentioning is that the Tweedie distribution is able to handle the zero-inflated data, where the occurrence of zeros is more frequent than what standard distributions can accommodate (Saha et al., 2020). This feature provides us extra flexibility to handle more types of data sets.

As discussed by Jorgensen (1987), the Tweedie distribution has three parameters: mean ($\mu > 0$), dispersion parameter ($\phi > 0$), and power parameter (p). For a random variable Y following Tweedie distribution, the PDF of Y is given by

$$f(y; \mu, \phi, p) = a(y; \phi) exp\left[\frac{1}{\phi}\left\{y\frac{\mu^{1-p}}{1-p} - \frac{\mu^{2-p}}{2-p}\right\}\right], \tag{3.1}$$

where

$$a(y; \phi) = \frac{1}{y}\sum_{j=1}^{\infty} \frac{y^{-j\alpha}(p-1)^{j\alpha}}{\phi^{j(1-\alpha)}(2-p)^j j!\Gamma(-j\alpha)}, \text{and } \alpha = \frac{2-p}{1-p}, 1 \leq p \leq 2, \phi > 0. \tag{3.2}$$

Additionally, for $p = 0$ it yields a normal distribution, $p = 1$ gives a quasi Poisson distribution, $p = 2$ yields a Gamma distribution, and $p = 3$ yields an inverse Gaussian distribution. The mean value is $E[Y] = \mu$ and the variance is $Var(Y) = \phi\mu^p$. Therefore, Tweedie distribution is characterized by the mean-variance relationship. Even though there is no closed form CDF for Tweedie distribution, when the power parameter $1 < p < 2$, Tweedie distributed random variable can be written as sum of M gamma variables, where M follows the Poisson distribution (Smyth, 2007).

Proposition 3.2.1 (Compound Poisson-gamma, Smyth, 2007)**.**
Let M be the Poisson distributed random variable with rate parameter λ such that $P(M = k) = \frac{e^{-\lambda}\lambda^k}{k!}, \quad k = 0, 1, 2, \ldots,$ and X be a gamma random variable with shape parameter α and scale parameter β such that

$$f(x_i; \alpha, \beta) = \frac{x_i^{\alpha-1}e^{-x_i/\beta}}{\beta^\alpha\Gamma(\alpha)}, x_i > 0. \tag{3.3}$$

Then, the random variable Y

$$Y = \sum_{i}^{M} X_i \tag{3.4}$$

follows the Tweedie distribution with parameters

$$\mu = \lambda\alpha\beta, p = \frac{\alpha + 2}{\alpha + 1}, \phi = \frac{\lambda^{1-p}(\alpha\beta)^{2-p}}{2-p},$$

or inversely, given the parameters of Tweedie distribution, the parameters of gamma and Poisson distributions can be written as

$$\lambda = \frac{\mu^{2-p}}{\phi(2-p)}, \alpha = \frac{2-p}{p-1}, \beta = \phi(p-1)\mu^{p-1}.$$

It is well-known that the CDF of the sum of n iid gamma variables is $F_Y(y) = \frac{\gamma(n\alpha,\beta y)}{\Gamma(\alpha)}$ if $0 \leq Y < \infty$ or zero elsewhere, where $\gamma(n\alpha, \beta y) = \int_0^y f(t; n\alpha, \beta)dt$ is incomplete gamma distribution (Smyth, 2007). Based on the compound Poisson-gamma distribution, the CDF of a Tweedie variable can be obtained when $1 < p < 2$.

Corollary 3.2.1 (CDF of Tweedie Variable when $1 < p < 2$)**.**
By Proposition 3.2.1, let

$$Y \sim Tw_p(\mu, \phi),$$

by expressing $Y = \sum_{i=1}^{M} X_i$, where X_i follows Beta$(a = (2-p)/(p-1), b = \phi(p-1)\mu^{p-1})$ and M follows Poisson$(\lambda = (\mu^{2-p})/(\phi(2-p))$, the CDF of Y can be written as

$$\begin{aligned} F_Y(y) &= \Pr(Y \leq y) \\ &= \sum_{k=0}^{\infty} \Pr\left(\sum_{i=1}^{k} X_i \leq y \middle| M = k \right) P(M = k) \\ &= \sum_{k=0}^{\infty} \frac{1}{\Gamma(\alpha)} \gamma(k\alpha, \beta y) \frac{e^{-\lambda}\lambda^k}{k!}, \end{aligned} \tag{3.5}$$

if $y \in [0, +\infty)$ or 0 otherwise, where $P(M = k)$ is the probability mass function of the Poisson distribution, and $P(\sum_{i=1}^{k} X_i \leq y | M = k)$ is the conditional probability that the sum of k independent Gamma random variables is less than y.

Dunn (2022) (also see Dunn and Smyth, 2005) provided the R package 'tweedie' and functions 'dTweedie' and 'pTweedie' for computing the pdf and cdf of Tweedie distribution at a certain value, and 'rTweedie' for generating Tweedie random variable, see below,

```
install.package('tweedie')
library(tweedie)
dTweedie(y, mu, phi, p, LOG=TRUE)
pTweedie(q, mu, phi, p)
rTweedie(q, mu, phi, p)
```

For further details on R packages and arguments, see The Comprehensive R Archive Network (CRAN) at
https://search.r-project.org/CRAN/refmans/tweedie/html/dtweedie.html.

Beta distributions are a family of continuous distribution with the support $(0, 1)$. This attribute simplifies the process of updating beliefs in Bayesian analysis. This probability distribution is parameterized by two positive parameters, shape one $(a > 0)$ and shape two $(b > 0)$. Let X be a Beta distributed random variable, and the PDF of X is given by

$$f(x; a, b) = \frac{x^{a-1}(1-x)^{b-1}}{B(a, b)}, a > 0, b > 0, \tag{3.6}$$

where $B(a, b) = \int_0^1 t^{a-1}(1-t)^{b-1} dt$ is the Beta function, which is a normalization constant to ensure the total probability integrates to 1. It is interesting that the Beta distribution can take different shapes including uniform ($a = b = 1$), uni-modal ($a = 2$, $b > 2$, or $a > 2$, $b = 2$), J-shaped ($a < 1$, $b > 2$, or $a > 2$, $b < 1$), and U-shaped

$(1 < a < 2, 1 < b < 2)$ (Johnson et al., 1994 - 1995). The mean and variance of Beta distributed random variables are given by

$$E[X] = \frac{a}{a+b} \text{ and } Var(X) = \frac{ab}{(a+b)^2(a+b+1)}.$$

The CDF of the Beta distribution is the integral of its PDF with respect to the random variable from 0 to x, that is

$$F(x; a, b) = \int_0^x \frac{t^{a-1}(1-t)^{b-1}}{B(a, b)} dt, \tag{3.7}$$

if $x \in (0, 1)$. Another simplified expression of CDF of Beta distribution is given by

$$F_X(x; a, b) = \frac{B(x; a, b)}{B(a, b)}, \tag{3.8}$$

where $B(x; a, b) = \int_0^x t^{a-1}(1-t)^{b-1} dt$ is the incomplete beta function (see Paris, 2011). In R, one may use functions 'dbeta' and 'pbeta' to compute the PDF and CDF of Beta distribution at a certain value, and 'rbeta' to generate beta distributed random variable, see below,

```
dbeta(x, shape1, shape2, ncp = 0, log = FALSE)
pbeta(q, shape1, shape2, ncp = 0, lower.tail = TRUE,
      log.p = FALSE)
rbeta(n, shape1, shape2, ncp = 0)
```

For further details on R packages and arguments, see The Comprehensive R Archive Network (CRAN) at
https://search.r-project.org/CRAN/refmans/fitODBOD/html/pBETA.html.

3.2.2 Tweedie-Jittering and Beta-Jittering

Let Y be a discrete random variable from the set of non-negative integers, and X be a vector of predictors in $\mathbb{R}^p$. The conditional quantile function $Q_Y(\tau|\boldsymbol{x})$ is not a continuous function on the quantile because Y has a discrete distribution. Given the necessary distribution parameters, we choose Tweedie distribution $(\mu, \phi, 1 < p < 2)$ or Beta distribution (a, b), denoted by $f_U(u|\theta)$, to generate additive perturbations, denoted by U, to smooth the counts data Y. From Subsection 2.7.3, we know that one condition is that the additive noise should ensure the one-to-one relationship on quantiles of Z and quantiles of Y. Therefore, we set two primary assumptions for Y and U as follows,

(A1) Y is non-negative random variable, with the support of $\mathbb{N}_0$. $\boldsymbol{X}$ is a predictor variable in $\mathbb{R}^p$, $p \in \mathbb{N}_0$. Let $\{(y_i, \boldsymbol{x}_i')\}_{i=1}^n$ be the sample;

(A2) U is a non-negative continuous random variable following $f_U(u|\boldsymbol{\theta}^{\text{noise}})$. The distribution parameters $\boldsymbol{\theta}^{\text{noise}}$ are predetermined such that the practical range of realization u is $[0, 1)$.

Remark.

Condition A2 is to ensure one-to-one relationship with quantile of jittered response. The practical range $[0, 1)$ *indicates the smallest and largest value of* u *that we can sample from* $f_U(u|\boldsymbol{\theta}^{noise})$ *in practice. The Tweedie variable does not automatically satisfy the condition due to its non-negative support. Therefore, the parameters selection and handling of exceeded values of Tweedie variables are given in Section 3.3.*

Let $Z = Y + U$ be the dependent variable, where Z has a continuous distribution and

$$U \sim \begin{cases} \mathrm{Tw}_p(\mu, \phi) & \text{or,} \\ \mathrm{Beta}(a, b). \end{cases}$$

Suppose that Y is observed, and let $\boldsymbol{y} = (y_1, ..., y_n)$ be a vector of observations, since U is random variable, mathematically the QR estimator $\hat{\boldsymbol{\beta}}_\tau$ can be obtained by minimizing a sample objective function with U integrated out, that is,

$$\hat{\boldsymbol{\beta}}_\tau \equiv \arg\min_{\boldsymbol{\beta} \in \mathbb{R}^p} \frac{1}{n} \sum_{i=1}^{n} \int_{\text{supp}(u)} \rho_\tau(y_i + u - \boldsymbol{x}_i \boldsymbol{\beta}) f_U(u) du, \tag{3.9}$$

where $f_U(\cdot)$ denotes the PDF of U (Machado & Silva, 2005). To avoid computational complexity, the additive noise $\boldsymbol{u} = (u_1, ..., u_n)$ can be generated and directly added to observation. Let $\boldsymbol{X} = \boldsymbol{x}$, then the data set with noises are $(z_i, x_i)_{i=1}^n$, where $z_i = y_i + u_i$. Therefore, the QR model with response Z can be defined as

$$Q_Z(\tau|\boldsymbol{X}) = \boldsymbol{X}' \boldsymbol{\beta}_\tau, \tag{3.10}$$

where $\boldsymbol{\beta}_\tau$ is the QR coefficients at the τ^{th} quantile. Our jittering process can be summarized as

Step 1: Choose distribution parameters, i.e., $\boldsymbol{\theta} = (\mu, \phi, p)$ for Tweedie, or $\boldsymbol{\theta} = (a, b)$ for Beta distribution;

Step 2: Generate a set of additive noises of size n, $\boldsymbol{u} = (u_1, ..., u_n)$, from $f_U(u|\boldsymbol{\theta})$, where $f_U(u|\boldsymbol{\theta})$ is the PDF of either Tweedie or Beta variable upon choice, and $\boldsymbol{\theta}$ is vector of distribution parameters;

Step 3: Let $z_i = y_i + u_i, i = 1, ..., n$. Fit the QR model as Equation 3.10 and obtain the estimated QR parameters $\boldsymbol{\beta_\tau}$ by minimizing the objective function of QR coefficients.

Based on this procedure of jittering method and the distribution of U, it is possible to obtain $Q_Y(\tau|\boldsymbol{x})$ by $Q_Z(\tau|\boldsymbol{x})$.

Lemma 3.2.1.

Under assumption ***A1*** *and* ***A2****, the conditional quantile of count Y follows*

$$Q_Y(\tau|\boldsymbol{x}) = \lceil Q_Z(\tau|\boldsymbol{x}) - 1 \rceil, \tag{3.11}$$

where $\lceil c \rceil$ is the ceiling function that returns the smallest integer greater than or equal to constant c.

Proof. Let U and u be as defined in the step 2 of the jittering process. For non-negative discrete random variable Y conditional on $\boldsymbol{x}$, $Y - 1 \leq Z - 1 < Y$. Since quantile function is non-decreasing, it follows that $Q_Y(\tau|\boldsymbol{x}) - 1 \leq Q_Z(\tau|\boldsymbol{x}) - 1 \leq Q_Y(\tau|\boldsymbol{x})$, where $Q_Y(\tau|\boldsymbol{x})$ is also an integer. □

Theorem 3.2.1 (Relationship between Conditional Quantile Functions)**.**
Suppose that a random variable U follows a valid probability density function, denoted as $f_U(\cdot)$, and a corresponding cumulative distribution function, denoted as $F_U(\cdot)$, where parameters are known. Let $Z = Y + U$, under conditions ***A1*** *and* ***A2****,*

$$\tau = \sum_{y=0}^{Q_Y(\tau|\boldsymbol{x})} P(Y = y|\boldsymbol{x}) \cdot F_U(Q_Z(\tau|\boldsymbol{x}) - y), \tag{3.12}$$

where Y is non-negative discrete variable, $\boldsymbol{x}$ is vector of co-variates, and $0 < \tau < 1$ is the quantile.

Proof.
Let $Z = Y + U$, because U is non-negative, Y is less or equal to Z.

$$\begin{aligned} F_Z(z) &= \Pr(Y \leq z|\boldsymbol{x}) \cdot F_U(z - y) \\ &= \sum_{y=0}^{\lceil z-1 \rceil} \Pr(Y = y|\boldsymbol{x}) \cdot F_U(z - y), \end{aligned}$$

where $\lceil \cdot \rceil$ is the ceiling function. Note that the last equality sums to $Y = \lceil z-1 \rceil$ since when $y = \lceil z \rceil = z$ if z is an integer, $F_U(z - z) = F_U(0) = 0$ for non-negative continuous variable, resulting in the last term of the series being zero. At $z = Q_Z(\tau|\boldsymbol{x})$, the above becomes

$$\tau = \sum_{y=0}^{Q_Y(\tau|\boldsymbol{x})} \Pr(Y = y|\boldsymbol{x}) \cdot F_U(Q_Z(\tau|\boldsymbol{x}) - y),$$

by the facts that $F_Z(Q_Z(\tau|\boldsymbol{x})) = \tau$ and $Q_Y(\tau|\boldsymbol{x}) = \lceil Q_Z(\tau|\boldsymbol{x}) - 1 \rceil$ by Lemma 3.2.1. $\square$

The estimator $\hat{\boldsymbol{\beta}}_\tau$ of quantile regression coefficient in Equation 3.10 is the solution of minimization problem,

$$\hat{\boldsymbol{\beta}}_\tau \equiv \arg\min_{\boldsymbol{\beta} \in \mathbb{R}^p} S(\boldsymbol{\beta}) \equiv \arg\min_{\boldsymbol{\beta} \in \mathbb{R}^p} \frac{1}{n} \sum_{i=1}^{n} \rho_\tau(z_i - \boldsymbol{x}_i'\boldsymbol{\beta}), \tag{3.13}$$

where $\rho_\tau(\cdot)$ is the check loss function on τ^{th} quantile shown in Equation 2.8. Estimating equation can be derived by differentiating the objective function $S(\boldsymbol{\beta})$ with respect to $\boldsymbol{\beta}$ and set the derivative to 0. That is,

$$\begin{aligned} U(\boldsymbol{\beta}) = \frac{\partial S(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}} &= \frac{1}{n} \sum_{i=1}^{N} \rho_\tau'(z_i - \boldsymbol{x}_i'\boldsymbol{\beta}) \\ &= \frac{1}{n} \sum_{i=1}^{N} (\tau - I_{\{z_i - \boldsymbol{x}_i'\boldsymbol{\beta} < 0\}})\boldsymbol{x}_i \\ &= \frac{1}{n} \sum_{i=1}^{N} \psi_\tau(z_i - \boldsymbol{x}_i\boldsymbol{\beta})\boldsymbol{x}_i \\ &\overset{\text{set}}{\equiv} 0, \end{aligned} \tag{3.14}$$

where $\boldsymbol{x_i} = (x_{i1}, x_{i2}, ..., x_{ip})', i = 1, ..., n$, is a $p \times 1$ vector of co-variates, $\boldsymbol{\beta}_\tau = (\beta_{\tau 0}, \beta_{\tau 1}, ..., \beta_{\tau, p-1})'$ is a $p \times 1$ vector of QR coefficidnts, and $\psi_\tau(v) = \rho_\tau'(v) = \tau - I_{\{v<0\}}$ is a discontinuous function. An efficient algorithm to obtain $\hat{\boldsymbol{\beta}}_\tau$ by solving the Equation 3.14 is given by Koenker (2004). The implement is available in statistical software R by loading external package 'quantreg' and using function 'rq' as follows:

```
install.package('quantreg')
library(quantreg)
rq(x, y, tau=-1, alpha=.1, dual=TRUE, int=TRUE, tol=1e-4,
ci = TRUE, method="score", interpolate=TRUE, tcrit=TRUE, hs=TRUE)
```

This package will be used for the estimation and inference of QR model in Chapter 4.

Now, a question arise in step 1: how to choose distribution parameters to ensure the performance of QR estimator and satisfy condition **A2**? We notice that random variables from the Beta distribution can satisfy the assumption easily. However, the support of Tweedie distribution is non-negative, from zero to positive infinity. If the generated value exceeds 1, the one-to-one relationship with quantile of the response may be violated. Therefore, we have to restrict the parameters, μ, ϕ, p so that the generated random variables can be bounded from above by 1. An algorithm to determine the parameters of Tweedie distribution and Beta distribution to enhance the performance of jittering method as well as to satisfy the conditions will be discussed in Subsection 3.3.

3.2.3 Average Jittering Estimator

Average jittering QR estimator can be obtained by the technique of averaging out noises introduced by Machado and Silva (2005), which is also applicable to Beta and Tweedie jittering method. Recall that $Z = Y + U$ is the actual response depending on the additive noise generated from either the Tweedie or Beta distribution. The bias and efficiency of estimators of QR with jittering highly depend on the random noises generated at that time. To reduce the uncertainty in performance of QR model and hence improve the efficiency of the QR estimators, the technique of averaging out the random noises can be applied (Machado & Silva, 2005).

Let $(y_i, \boldsymbol{x}_i)$ be the sample. Using the jittering method, we repeatedly generate m sets of random noises U from either $\text{Tw}_p(\mu, \phi)$ or $\text{Beta}(a, b)$. Let $U^{(l)}, l = 1, 2, 3, ..., m$, be the l^{th} set of random noises generated from the chosen statistical distribution. The l^{th} jittered QR response is

$$Z^{(l)} = Y + U^{(l)}.$$

and

$$\hat{\boldsymbol{\beta}}_\tau^{(l)} \equiv \underset{\boldsymbol{\beta}\in\mathbb{R}^p}{\operatorname{argmin}} \sum_i^n \rho_\tau(z_i^{(l)} - \boldsymbol{x}_i'\boldsymbol{\beta}),$$

where $\hat{\boldsymbol{\beta}}_\tau^{(l)}$ is the QR estimator obtained from the l-th data-set, $(Z^{(l)}, \boldsymbol{X})$. Then, average of m coefficient estimators is

$$\hat{\boldsymbol{\beta}}_\tau^m = \frac{1}{m}\sum_{l=1}^m \hat{\boldsymbol{\beta}}_\tau^{(l)}, \tag{3.15}$$

where $\hat{\boldsymbol{\beta}}_\tau^m$ is denoted as average jittering QR estimator over m jittered samples, e.g., $\hat{\boldsymbol{\beta}}_\tau^{50}$ denotes the average jittering estimator when $m = 50$ jittered samples are used. According to the simulation results, $m \in (10, 50)$ would be a good choice. The averaged QR estimator $\hat{\boldsymbol{\beta}}_\tau^m$ usually provides lower variance than that without averaging out the noises. To distinguish the QR methods, we call it the 'simple jittering method' if noise isn't averaged out, and the 'average jittering method' if it is.

3.2.4 Model Specification and Transformation

The foundation of QR was introduced by Koenker and Bassett (1978) in their seminal work. However, the linear assumption in traditional QR often restricts its ability to capture the nuanced dynamics of real data. The non-linear QR model addresses this restriction by introducing a non-linear function into the QR model. Thus, the non-linear QR model is

$$Q_Z(\tau|\boldsymbol{x}) = g(\boldsymbol{x}'\boldsymbol{\beta}_\tau), \tag{3.16}$$

where $g(\cdot)$ is a non-linear function. This allows flexibility in interpretation of conditional QR models. Denote a transformation or 'link' function as $T(Z;\tau)$ probably depending on the quantile level $T(Z;\tau)$ can be the identity, a linear translate (Ma et al., 2011) and the logarithmic function (Machado & Silva, 2005). Considering a non-linear data-set, by letting $g(\cdot)$ be an exponential function, the QR model is specified as follows

$$Q_Z(\tau|\boldsymbol{X}) = \exp(\boldsymbol{X}'\boldsymbol{\beta}_\tau). \tag{3.17}$$

However, since the random noises are added to the original counts, the conditional quantile function model $Q_Z(\tau|X)$ should be bounded from below by the τ^{th} quantile of the random noise U, denoted as $Q_U(\tau)$. Accordingly, we adjust the QR model by adding $Q_U(\tau)$ to the non-linear quantile function, that is

$$Q_Z(\tau|\boldsymbol{X}) = Q_U(\tau) + \exp(\boldsymbol{X}'\boldsymbol{\beta}_\tau). \tag{3.18}$$

The logarithmic transformation of QR response depending on quantile level τ can be defined as

$$T(Z;\tau) = \begin{cases} \log(Z - Q_\tau(U)) & \text{for } Z > Q_U(\tau), \\ \zeta & \text{for } Z \leq Q_U(\tau), \end{cases} \tag{3.19}$$

where ζ is a small positive number. This transformation is valid due to the property of equivariance of QR model and that $T(Z;\tau)$ is monotone. Alternatively, one may also specify a QR model as

$$Q_Z(\tau|\boldsymbol{X}) = Q_U(\tau) + \boldsymbol{X}'\boldsymbol{\beta}_\tau, \tag{3.20}$$

and the corresponding linear transformation is

$$T(Z;\tau) = \begin{cases} Z - Q_\tau(U) & \text{for } Z > Q_U(\tau), \\ \zeta & \text{for } Z \leq Q_U(\tau). \end{cases} \tag{3.21}$$

Then, after transforming or linearly translating the QR response, the corresponding linear QR model is defined as

$$Q_{T(Z;\tau)}(\tau|\boldsymbol{X}) = \boldsymbol{X}'\boldsymbol{\beta}_\tau, \tag{3.22}$$

where $T(Z;\tau)$ is the linearly transformed jittered QR response of Z, defined as above. The estimate of QR coefficients

$$\hat{\boldsymbol{\beta}}_\tau \equiv \underset{\boldsymbol{\beta}\in\mathbb{R}^p}{\operatorname{argmin}} \sum_{i}^{n} \rho_\tau(T(z_i;\tau) - \boldsymbol{x}_i'\boldsymbol{\beta}). \tag{3.23}$$

Hence, the average jittering QR estimate can be obtained from m jittered samples $(Z^{(l)}, \boldsymbol{X})$, $l = 1, ..., m$, by applying the averaging-out noises technique. Suppose that

$$\hat{\boldsymbol{\beta}}_\tau^{(l)} \equiv \underset{\boldsymbol{\beta}\in\mathbb{R}^p}{\operatorname{argmin}} \sum_{i}^{n} \rho_\tau(T(z_i^{(l)};\tau) - \boldsymbol{x}_i'\boldsymbol{\beta}), \tag{3.24}$$

and the average jittering QR estimates is

$$\hat{\beta}_\tau^m = \frac{1}{m} \sum_{l=1}^{m} \boldsymbol{\beta}_\tau^{(l)}. \tag{3.25}$$

Based on the relationship between variables, transformation of response can be applied, which gives simplification to result interpretations. Moreover, as discussed by C. Chen (2005), statistical inference on QR model is not affected by monotone transformation of response. Thus, asymptotic properties of the model can be derived following Koenker and Bassett (1978). The corresponding variance-covariance matrices of QR estimates and natural sandwich estimators will be discussed in Section 3.4.

3.3 Optimal Choice of Noise Parameters

Geraci and Bottai (2006) initially estimated QR model by using the asymmetric Laplace distribution (ALD). The close connection between ALD and the QR model provides us an option of optimally selecting the jittering distribution.

3.3.1 Asymmetric Laplace Distribution for Independent Observations

The ALD is a generalization of the Laplace distribution that allows different scale parameters on the positive and negative halves of the real line. By varying the distribution parameters, one can customize the shape of ALD to model asymmetric and skewed data.

A random variable Y is said to be asymmetric Laplace distributed if it has the probability density function

$$f(y|\mu, \sigma, \tau) = \frac{\tau(1-\tau)}{\sigma} \exp\left\{-\rho_\tau\left(\frac{y-\mu}{\sigma}\right)\right\}$$

where the support of Y is the real line, $-\infty < \mu < \infty$ is the location parameter, $\sigma > 0$ is the scale parameter, $0 < \tau < 1$ is the skewness parameter, and $\rho_\tau(u) = u\big(\tau - I(u \leq 0)\big)$ is the check loss function and $I(\cdot)$ is the indicator function. For an i.i.d. sample $(y_1, y_2, ..., y_n)$ from an ALD with some specific parameters, the corresponding likelihood function of ALD is obtained as

$$
\begin{aligned}
L(\mu, \sigma, \tau; y) &= \prod_{i}^{n} \frac{\tau(1-\tau)}{\sigma} \exp\left\{-\rho_\tau\left(\frac{y_i - \mu}{\sigma}\right)\right\} \\
&\propto \frac{1}{\sigma} \exp\left\{-\sum_{i}^{n} \rho_\tau\left(\frac{y_i - \mu}{\sigma}\right)\right\},
\end{aligned}
$$

where $\tau \in (0, 1)$.

3.3.2 Estimation of QR Model Using ALD

Consider a data-set in the form of $(y_i, \boldsymbol{x}_i')$, for i = 1, 2, ..., n. Recall that a linear QR model is

$$
Q_{y_i}(\tau|\boldsymbol{x}_i) = \boldsymbol{x}_i'\beta_\tau, \quad i = 1, 2, ..., n \tag{3.26}
$$

where β_τ is the QR coefficient parameter at the τ^{th} quantile. The regression coefficient β_τ minimizes the objective function

$$
\hat{\boldsymbol{\beta}}_\tau = \arg\min_{\beta_\tau \in \mathbb{R}^p} \left\{ \sum_{i \in (i: y_i \geq \boldsymbol{x}_i'\boldsymbol{\beta}_\tau)}^{n} \tau|y_i - \boldsymbol{x}_i'\boldsymbol{\beta}_\tau| + \sum_{i \in (i: y_i < \boldsymbol{x}_i'\boldsymbol{\beta}_\tau)}^{n} (1-\tau)|y_i - \boldsymbol{x}_i'\boldsymbol{\beta}_\tau| \right\}, \tag{3.27}
$$

and the log-likelihood function of ALD takes the form

$$
\begin{aligned}
l(\mu, \sigma, \tau) &= \log\tau + \log(1-\tau) - \log\sigma - \sum_{i}^{n} \rho_\tau\left(\frac{y_i - \mu}{\sigma}\right) \\
&= \log\tau + \log(1-\tau) - \log\sigma \\
&\quad - \left(\sum_{i \in (i: y_i \geq \mu)}^{n} \tau|y_i - \mu| + \sum_{i \in (i: y_i < \mu)}^{n} (1-\tau)|y_i - \mu| \right),
\end{aligned} \tag{3.28}
$$

whose negative logarithm is a linear function of the absolute residuals, and whose mode (peak value) corresponds to the τ^{th} conditional quantile of the dependent variable. Therefore, given that σ is always positive, maximizing the above log-likelihood function is equivalent to minimizing the summation of absolute errors, just with different weights assigned to the positive or negative residuals based on the quantile level τ. Therefore, treating the skewness parameter τ of ALD as the quantile level of QR model, the connection between ALD and QR is established.

Let $\mu_i = \boldsymbol{x_i}'\boldsymbol{\beta}_\tau, i = 1, 2, ..., n$. Then the likelihood function is

$$
\begin{aligned}
L(\boldsymbol{\beta}_\tau, \sigma; \boldsymbol{y}, \tau) &\propto \prod_{i=1}^{n} \sigma^{-1} \exp\left\{-\rho_\tau\left(\frac{y_i - \mu_i}{\sigma}\right)\right\} \\
&\propto \sigma^{-1} \exp\left\{-\sum_{i}^{n} \rho_\tau\left(\frac{y_i - \mu_i}{\sigma}\right)\right\} \\
&\propto \sigma^{-1} \exp\left\{-\sum_{i}^{n} \rho_\tau\left(\frac{y_i - \boldsymbol{x}_i^T \boldsymbol{\beta_\tau}}{\sigma}\right)\right\}
\end{aligned}
\tag{3.29}
$$

and hence the log-likelihood function is

$$
l(\boldsymbol{\beta}_\tau, \sigma; \boldsymbol{y}, \tau) = \log \tau + \log(1-\tau) - \log \sigma - \sum_{i}^{n} \rho_\tau\left(\frac{y_i - \boldsymbol{x}_i^T \boldsymbol{\beta}_\tau}{\sigma}\right).
$$

As the maximization of log-likelihood $l(\boldsymbol{\beta}_\tau, \sigma; \boldsymbol{y}, \tau)$ over a domain $\boldsymbol{\beta}_\tau \in \mathbb{R}^p$ is equivalent to minimizing $-l(\boldsymbol{\beta}_\tau, \sigma; \boldsymbol{y}, \tau)$ over the same domain, we have

$$
\begin{aligned}
\arg\max_{\boldsymbol{\beta}_\tau \in \mathbb{R}^p} l(\boldsymbol{\beta}_\tau, \sigma; \boldsymbol{y}, \tau) &\equiv \arg\min_{\boldsymbol{\beta}_\tau \in \mathbb{R}^p} -l(\boldsymbol{\beta}_\tau, \sigma; \boldsymbol{y}, \tau) \\
&= \arg\min_{\boldsymbol{\beta}_\tau \in \mathbb{R}^p} -\log \tau - \log(1-\tau) + \log \sigma + \sum_{i}^{n} \rho_\tau\left(\frac{y_i - \boldsymbol{x}_i' \boldsymbol{\beta}_\tau}{\sigma}\right) \\
&= \arg\min_{\boldsymbol{\beta}_\tau \in \mathbb{R}^p} \sum_{i}^{n} \rho_\tau\left(\frac{y_i - \boldsymbol{x}_i' \boldsymbol{\beta}_\tau}{\sigma}\right) \\
&= \arg\min_{\boldsymbol{\beta}_\tau \in \mathbb{R}^p} \sum_{i}^{n} \frac{y_i - \boldsymbol{x}_i' \boldsymbol{\beta}_\tau}{\sigma}\left[\tau - I\left(\frac{y_i - \boldsymbol{x}_i' \boldsymbol{\beta}_\tau}{\sigma} < 0\right)\right] \\
&\equiv \arg\min_{\boldsymbol{\beta}_\tau \in \mathbb{R}^p} \sum_{i}^{n} (y_i - \boldsymbol{x}_i' \boldsymbol{\beta}_\tau)\left[\tau - I\left(y_i - \boldsymbol{x}_i' \boldsymbol{\beta}_\tau < 0\right)\right] \\
&= \arg\min_{\boldsymbol{\beta}_\tau \in \mathbb{R}^p} \sum_{i}^{n} \rho_\tau(y_i - \boldsymbol{x}_i' \boldsymbol{\beta}_\tau),
\end{aligned}
\tag{3.30}
$$

where the 4th and 5th equality are equivalent by the fact that $I((y_i - \boldsymbol{x}_i'\boldsymbol{\beta}_\tau)/\sigma < 0) = I((y_i - \boldsymbol{x}_i'\boldsymbol{\beta}_\tau) < 0)$ since σ is always positive.

3.3.3 Treating Additive Noise as Random Effect

We can treat additive noise as random effect. When additive noise is i.i.d., the QR estimators can be obtained from a minimization problem when integrating out the random noise U. Let $f_U(u|\boldsymbol{\theta})$ denote the distribution of U, the QR estimator

$$\hat{\boldsymbol{\beta}}_\tau \equiv \arg\min_{\boldsymbol{b}\in\mathbb{R}^p} \frac{1}{n}\int_{u\in\mathbb{R}} \sum_i^n \rho_\tau(y_i + u - \boldsymbol{x}_i\boldsymbol{b}) f_u(u|\boldsymbol{\theta})du, \tag{3.31}$$

where $\boldsymbol{\theta}$ is a vector of distribution parameters. The only implementation issue is that the parameters of Tweedie or Beta distribution are undetermined. Therefore, we aim to find the parameters of distribution of U that minimize the objective function as far as possible under given assumptions.

Definition 3.3.1. *Denote the optimal parameters as $\boldsymbol{\theta}^{Noise}$. Distribution parameters are said to be optimal if they satisfy condition* ***A2*** *and*

$$\begin{aligned}
&\min_{\boldsymbol{b}\in\mathbb{R}^p} \frac{1}{n}\int_{u\in\mathbb{R}} \sum_i^n \rho_\tau(y_i + u - \boldsymbol{x}_i\boldsymbol{b}) f_u(u|\boldsymbol{\theta}^{Noise})du \\
&\qquad\qquad \le \arg\min_{\boldsymbol{b}\in\mathbb{R}^p} \frac{1}{n}\int_{u\in\mathbb{R}} \sum_i^n \rho_\tau(y_i + u - \boldsymbol{x}_i\boldsymbol{b}) f_u(u|\boldsymbol{\theta})du,
\end{aligned} \tag{3.32}$$

where $\boldsymbol{\theta}$ are any other distribution parameters.

The noises generated from such a distribution smooth the discrete QR response as well as reduces the possible value of sum of absolute deviations, leading to less variation of QR estimators.

$\boldsymbol{\theta}^{\text{Noise}}$ can be determined by treating them as missing part of data. Let $(z_i, \boldsymbol{x}_i', u_i), i = 1, ..., n$, be the complete data-set, where $\boldsymbol{x}_i'$ are covariates, and u_i are the missing part. According to the connection between QR model and ALD, we assume that the conditional distribution of jittered response $z_i = y_i + u_i$ given u_i, follows the ALD with parameters (μ_i, σ, τ),

$$f_{Z_i|U}(z_i|\boldsymbol{\beta}_\tau, u_i, \sigma) = \frac{\tau(1-\tau)}{\sigma}\exp\left\{-\rho_\tau\left(\frac{z_i - \mu_i}{\sigma}\right)\right\}, \tag{3.33}$$

where $\mu_i = Q_\tau(U) + T^{-1}(\boldsymbol{x}_i'\boldsymbol{\beta}_\tau)$ is inverse transformation function of linear predictors of τ-th quantile. Also, we assume that the error term ϵ_i are independent, and u_i and ϵ_i are independent of each other. Therefore, the joint distribution of Z_i and U_i can be written as

$$f(z_i, u_i|\boldsymbol{\beta}_\tau, \sigma, \tau, \boldsymbol{\theta}) = f_{Z_i|U_i}(z_i|\boldsymbol{\beta}_\tau, u_i, \sigma) f_{U_i}(u_i|\boldsymbol{\theta}). \tag{3.34}$$

Let $\boldsymbol{z} = \{z_1, z_2, ..., z_n\}$ and $\boldsymbol{u} = \{u_1, u_2, ..., u_n\}$. The likelihood of complete dataset $\{\boldsymbol{z}, \boldsymbol{u}\}$ is obtained as

$$f(\boldsymbol{z}, \boldsymbol{u}|\boldsymbol{\beta}_\tau, \sigma, \tau, \boldsymbol{\theta}) = \prod_i^n f_{Z_i|U_i}(z_i|\boldsymbol{\beta}_\tau, u_i, \sigma) f_{U_i}(u_i|\boldsymbol{\theta}). \tag{3.35}$$

Based on the joint distribution, the likelihood of jittered QR response $\boldsymbol{z}$ can be obtained. It is equivalent to find the marginal pdf of $\boldsymbol{z}$ by simply integrating the missing part $\boldsymbol{u}$ out, which gives the marginal pdf of $\boldsymbol{z}$ as

$$\begin{aligned} L(\boldsymbol{\beta}_\tau, \sigma, \tau, \boldsymbol{\theta}; \boldsymbol{z}) &= f(\boldsymbol{z}|\boldsymbol{\beta}_\tau, \sigma, \tau, \boldsymbol{\theta}) \\ &= \int f(\boldsymbol{z}, \boldsymbol{u}|\boldsymbol{\beta}_\tau, \sigma, \tau, \boldsymbol{\theta}) d\boldsymbol{u} \\ &= \int \prod_i^n f_{Z_i|U_i}(z_i|\boldsymbol{\beta}, u_i, \sigma) f_U(u_i|\boldsymbol{\theta}) du_i, \end{aligned} \tag{3.36}$$

and the corresponding log-likelihood function is

$$l(\boldsymbol{\beta}_\tau, \sigma, \tau, \boldsymbol{\theta}; \boldsymbol{z}) = \int \sum_{i=1}^n \log f_{Z_i|U_i}(z_i|\boldsymbol{\beta}, u_i, \sigma) + \log f_{U_i}(u_i|\boldsymbol{\theta}) du. \tag{3.37}$$

We know that in order to estimate the QR parameters, we can alternatively maximize the likelihood of ALD. Therefore, the MLEs of distribution parameter of U that maximize $L(\beta_\tau, \sigma, \tau, \boldsymbol{\theta}; \boldsymbol{z})$ also minimize Equation 3.31. Denote such MLEs as $\hat{\boldsymbol{\theta}}^{\text{ALD}}$. $\boldsymbol{\theta}^{\text{Noise}} = \hat{\boldsymbol{\theta}}^{\text{ALD}}$ are determined to be the optimal parameters to generate additive noises for jittering the counts. We have to mention that even though maximizing the likelihood also returns the MLEs of $\boldsymbol{\beta}_\tau$, asymptotic theorems are not applicable to this estimate. To utilize the asymptotic properties and technique of averaging noises out, estimate $\hat{\boldsymbol{\beta}}_\tau$ can be only obtained by using methods in Section 3.2.

3.3.4 Computational Algorithms

To efficiently obtain the MLEs of parameters $\boldsymbol{\beta}_\tau, \sigma, \boldsymbol{\theta}$, denoted by $\hat{\boldsymbol{\beta}}_\tau, \hat{\sigma}, \hat{\boldsymbol{\theta}}$, respectively, we apply some computational algorithms. Theoretically, even though the response Z is not observed data, it is still possible to solve the MLEs, as Z has an ALD and $Z = Y + U$. We use $f(\cdot|\cdot)$ to denote a generic conditional PDF. The log-likelihood function

$$
\begin{aligned}
l(\Omega; z) = \log f(z|\Omega) &= \log \int f(z, u|\Omega) du \\
&= \log \int \frac{f(z, u|\Omega)}{f(u|z; \Omega^{old})} f(u|z; \Omega^{old}) du \\
&= \log \mathbb{E}\left[\frac{f(z, u|\Omega)}{f(u|z; \Omega^{old})} \middle| z; \Omega^{old} \right] \\
&\geq \mathbb{E}\left[\log \left(\frac{f(z, u|\Omega)}{f(u|z; \Omega^{old})} \right) \middle| z; \Omega^{old} \right] \quad \text{(Jensen's inequality)} \\
&= \mathbb{E}[\log f(z, u|\Omega) | z; \Omega^{old}] - \mathbb{E}[\log f(u|z; \Omega^{old}) | z; \Omega^{old}] \\
&= Q(\Omega|\Omega^{old}) - \mathbb{E}[\log f(u|z; \Omega^{old}) | z; \Omega^{old}],
\end{aligned}
\tag{3.38}
$$

where $\Theta = (\boldsymbol{\beta}_\tau, \sigma, \tau, \boldsymbol{\theta})$. Denote the last equality by $g(\Omega|\Omega^{old})$, we have

$$
l(\Omega; z) \geq g(\Omega|\Omega^{old}),
$$

for all Ω. The Jensen's inequality becomes equality when $\Omega = \Omega^{old}$ since the term inside expectation becomes constant. Therefore, any value of Θ that increases $g(\Theta|\Theta^{old})$ beyond $g(\Theta^{old}|\Theta^{old})$ must also increase $l(\Theta; z)$ beyond $l(\Theta^{old}; z)$. Thus, M-step in each iteration finds such a Θ by maximizing $Q(\Theta|\Theta^{old})$ over Θ is equivalent to maximizing $g(\Theta|\Theta^{old})$ over Θ. However, since the likelihood function includes Tweedie or Beta distribution, the E-step does not have a closed form, and the estimation of parameters is not straight-forward. To obtain the MLEs, we apply Monte Carlo Expectation Maximization (MCEM) algorithm (Levine and Casella, 2001; also see Geraci and Bottai, 2006). The MCEM algorithm is an extension of the classical Expectation Maximization (EM), which incorporates Monte Carlo methods. This algorithm is particularly useful in estimation problems where the expectation is difficult to compute. Unlike classical EM algorithm method, the E-step of MCEM algorithm avoid the intractable integrals, which also benefits the following M-step. This iterative algorithm can be realized by introducing a sampler and some computational steps.

The MCEM algorithm consists of three components: Monte Carlo simulation (MC-step), expectation computation (E-step), and maximization (M-step). In each iteration, the MC-step generates large number of random variables from the density of latent variables conditional on observed data and current parameters, and the E-step computes the expected value based on the simulated samples, instead of computing the expected value directly. The last M-step involves maximizing the expected likelihood found from the E-step. Similar to Geraci and Bottai (2006) who applied MCEM algorithm to estimate the random effect quantile regression model (REQR), we can modify the MC-step, E-step and M-step to fit our target. For the sake of convenience, let $\boldsymbol{\Omega} = \{\boldsymbol{\beta}_\tau, \sigma, \tau, \boldsymbol{\theta}\}$ be the set of all parameters involved. Firstly, we define the expectation function in E-step for the i^{th} subject at $(t+1)$ iteration as

$$\begin{aligned} Q_i(\boldsymbol{\Omega}|\boldsymbol{\Omega}^{(t)}) &= \mathbb{E}[l(\boldsymbol{\Omega}; z_i, u_i)|z_i; \boldsymbol{\Omega}^{(t)}] \\ &= \int \{\log f_{Z|U}(z_i|\boldsymbol{\beta}, u_i, \sigma) + \log f(u_i|\boldsymbol{\theta})\} \times f(u_i|z_i; \boldsymbol{\Omega}^{(t)}) du_i \\ &= \int \{\log f_{Z|U}(y_i + u_i|\boldsymbol{\beta}, u_i, \sigma) + \log f(u_i|\boldsymbol{\theta})\} \times f(u_i|z_i; \boldsymbol{\Omega}^{(t)}) du_i, \end{aligned} \tag{3.39}$$

where $l(\boldsymbol{\Omega}; z_i, u_i)$ is log-likelihood of the i^{th} jittered response. The expression of expectation is written with respect to missing data U and is conditional on $Z = z$. Then, a suitable Monte Carlo technique can evaluate the expected value of the above log-likelihood.

MC-step:

Since the additive noises U are unobserved, a sample $\boldsymbol{v_i} = \{v_{i1}, v_{i2}, ..., v_{ik}\}$ with sample size of k will be taken from the density function of U conditional on observed data Z so that the MC-step can be completed. Using the relationship of $Z = Y + U$, Bayesian's theorem gives that

$$\begin{aligned} f(u|z_i; \boldsymbol{\Omega}) &= \frac{f(z_i|u; \boldsymbol{\beta}_\tau, \sigma) f(u|\boldsymbol{\theta})}{f(z_i|\boldsymbol{\Omega})} \\ &= \frac{f(y_i + u|u; \boldsymbol{\beta}_\tau, \sigma) f(u|\boldsymbol{\theta})}{f(z_i|\boldsymbol{\Omega})} \end{aligned} \tag{3.40}$$

where $f(z_i|\boldsymbol{\Omega}) = \int_u f(z_i, u_i|\boldsymbol{\Omega}) du$ and $f(u|\boldsymbol{\theta})$ is the PDF of Tweedie or Beta. Hence, the density of latent variable U conditional on Z is obtained. In order to draw samples from the density $f(u|z_i; \boldsymbol{\Omega})$ efficiently, Metropolis-Hasting sampling algorithm can be applied (Hastings, 1970). Let $f(u|z_i; \boldsymbol{\Omega})$ be the target distribution, denote $\boldsymbol{v}$ be the previous draw from the density $f(u|z_i; \boldsymbol{\Omega})$, and generate new values $\boldsymbol{v}^*$ using the target distribution. Since we would like to restrict the generated value of $v_{ij} \in [0, 1)$, the Uniform distribution $U[0, 1)$ can be a good choice for proposal distribution. We denote $q(u|D)$ as the proposal distribution. Thus, we accept $\boldsymbol{u}^*$ as a new value with probability of $A_j(\boldsymbol{u}, \boldsymbol{u}^*)$, otherwise reject this value. Here, the probability $A_k(\boldsymbol{v}, \boldsymbol{v}^*)$ is given by

$$\begin{aligned} A_k(\boldsymbol{v}, \boldsymbol{v}^*) &= \min\left\{1, \frac{f(u^*|z_i; \boldsymbol{\Omega})q(u|D)/f(z_i|\boldsymbol{\Omega})}{f(u|z_i; \boldsymbol{\Omega})q(u^*|D)/f(z_i|\boldsymbol{\Omega})}\right\} \\ &= \min\left\{1, \frac{f(z_i|u^*; \boldsymbol{\beta}_\tau, \sigma)f(u^*|\boldsymbol{\theta})q(u|D)}{f(z_i|u; \boldsymbol{\beta}_\tau, \sigma)f(u|\boldsymbol{\theta})q(u^*|D)}\right\}. \end{aligned} \tag{3.41}$$

Since both $f(z_i)$ in the numerator and denominator are cancelled out, it is not necessary to know the density of Z. In order to ensure the generated samples are independent of one another, we can alleviate this by choosing every h^{th} value in the chain as our posterior sample, which is called 'thinning'. Therefore, the MH algorithm for generating m samples conditional on each $z_i, i = 1, ..., n$ is as follows:

Algorithm 1 Modified Metropolis-Hastings algorithm for MC-step

for iteration $i = 1, 2, \ldots, n$ **do**

  Initialize $v^{(0)} \sim q(u|D)$, where $q(u|D)$ is uniform $[0, 1)$

  for iteration $j = 1, 2, \ldots, kh$ **do**

    Propose $u^{\text{cand}} \sim q(u^{(j)}|u^{(j-1)})$

    Acceptance Probability:

$$A_j(v^{\text{cand}}|v^{(j-1)}) = \min\left\{1, \frac{q(u^{(j-1)}|u^{\text{cand}})p(u^{\text{cand}})}{q(u^{\text{cand}}|u^{(j-1)})p(u^{(j-1)})}\right\}$$
$$= \min\left\{1, \frac{f(z_i|u^{\text{cand}}; \boldsymbol{\beta}_\tau, \sigma)f(u^{\text{cand}}|\boldsymbol{\theta})q(u^{(j-1)}|D)}{f(z_i|u^{(j-1)}; \boldsymbol{\beta}_\tau, \sigma)f(u^{(j-1)}|\boldsymbol{\theta})q(u^{\text{cand}}|D)}\right\}$$

    Generate $\alpha \sim \text{Uniform}(0, 1)$

    if $\alpha < A_j$ **then**

      Accept the proposal: $v_{ij} \leftarrow v^{\text{cand}}$

    else

      Reject the proposal: $v_{ij} \leftarrow v^{(j-1)}$

    end if

  end for

  Choose every h^{th} element of $\boldsymbol{v}_i = (v_{i1}, ..., v_{i,kh})$, to form a vector consists of k independent samples, i.e., thinning.

end for

By applying the above algorithm, we can have independent variables $\boldsymbol{v}_i = (v_{i1}, ..., v_{im})$ for $i = 1, ..., n$.

E-step:

Based on the simulated samples, we let $\tilde{z}_{ij} = y_i + v_{ij}$, the approximate value of expectation for i^{th} subject in Equation 3.39 can be written as

$$\begin{aligned} Q_i^*(\mathbf{\Omega}|\mathbf{\Omega}^{(t)}) &= \frac{1}{k}\sum_{j=1}^{k} l(\mathbf{\Omega}; \tilde{z}_{ij}, v_{ij}) \\ &= \frac{1}{k}\sum_{j=1}^{k}\Bigg(\log\tau + \log(1-\tau) - \log\sigma - \rho_\tau\Big(\frac{\tilde{z}_{ij}-\mu_i}{\sigma}\Big) + \log f(u_i|\boldsymbol{\theta})\Bigg). \end{aligned} \tag{3.42}$$

Hence, consider all n subjects, the approximately expected value based on the log-likelihood is obtained as

$$\begin{aligned} Q^*(\mathbf{\Omega}|\mathbf{\Omega}^{(t)}) &= \sum_{i=1}^{n} Q_i^*(\mathbf{\Omega}|\mathbf{\Omega}^{(t)}) \\ &= \sum_{i=1}^{n}\frac{1}{k}\sum_{j=1}^{k}\Bigg(\log\tau + \log(1-\tau) - \log\sigma - \rho_\tau\Big(\frac{\tilde{z}_{ij}-\mu_i}{\sigma}\Big) + \log f(u_i|\boldsymbol{\theta})\Bigg). \end{aligned} \tag{3.43}$$

M-step:

In the M-step, we aim to maximize the approximately expected value of log-likelihood in Equation 3.42. Therefore, the expression or numerical solution of mle of parameters shall be found so that in each iteration the expectation can be maximized. Firstly, it is easy to notice that to maximize $E^*(\mathbf{\Omega}|\mathbf{\Omega}^{(t)})$, it is necessary to minimize the term of $\rho_\tau\big(\frac{\tilde{z}_{ij}-\mu_i}{\sigma}\big)$. Thus, the mle of β_τ can be obtained as

$$\begin{aligned} \hat{\beta}_\tau &\equiv \arg\min_{\boldsymbol{\beta}\in\mathbb{R}^p} \sum_{i}^{n}\frac{1}{k}\sum_{j=1}^{k}\rho\Big(\frac{\tilde{z}_{ij} - T^{-1}(\boldsymbol{x}_i'\boldsymbol{\beta})}{\sigma}\Big) \\ &\equiv \arg\min_{\boldsymbol{\beta}\in\mathbb{R}^p} \sum_{i}^{n}\frac{1}{k}\sum_{j=1}^{k}\rho(\tilde{z}_{ij} - T^{-1}(\boldsymbol{x}_i'\boldsymbol{\beta})), \end{aligned} \tag{3.44}$$

where the last equivalent holds since σ is always positive. After applying transformation, Equation 3.44 is equivalent to estimate the QR parameters, and thus, any software for QR estimation can be used. Secondly, to find the MLEs $\hat{\sigma}$, we set the derivative of $E^*(\mathbf{\Omega}|\mathbf{\Omega}^{(t)})$ with respect to σ be zero, that is,

$$\begin{aligned}\frac{\partial E^*(\mathbf{\Omega}|\mathbf{\Omega}^{(t)})}{\partial\sigma} &= \sum_{i=1}^{n}\frac{1}{k}\sum_{j=1}^{k}\Big[-\frac{1}{\sigma}+\frac{\tilde{z}_{ij}-T^{-1}(\boldsymbol{x}_i'\boldsymbol{\beta}))}{\sigma^2}\Big(\tau - I_{\{\frac{\tilde{z}_{ij}-T^{-1}(\boldsymbol{x}_i'\boldsymbol{\beta})}{\sigma}\}}\Big)\Big] \\ &= \sum_{i=1}^{n}\frac{1}{k}\sum_{j=1}^{k}\Big[-\sigma+(\tilde{z}_{ij}-T^{-1}(\boldsymbol{x}_i'\boldsymbol{\beta})))\Big(\tau - I_{\{\tilde{z}_{ij}-T^{-1}(\boldsymbol{x}_i'\boldsymbol{\beta}))\}}\Big)\Big] \\ &\overset{\text{set}}{=} 0\end{aligned} \tag{3.45}$$

which yields the mle of σ,

$$\hat{\sigma} = \frac{1}{nk}\sum_{i=1}^{n}\sum_{j=1}^{k}\rho(\tilde{z}_{ij}-T^{-1}(\boldsymbol{x}_i'\hat{\boldsymbol{\beta}}_\tau)). \tag{3.46}$$

Thirdly, the MLEs of $\boldsymbol{\theta}$ is the solution to the equation that the derivative of $Q^*(\mathbf{\Omega}|\mathbf{\Omega}^{(t)})$ with respect to $\boldsymbol{\theta}$ equals zero. In Equation 3.42, only the last term involves $\boldsymbol{\theta}$. Therefore, $\hat{\boldsymbol{\theta}}$ is just the MLE based on the simulated samples $\boldsymbol{v_i} = \{v_{i1}, ..., v_{ik}\}$ in MC-step. In R, efficient algorithms to obtain the MLEs of parameters are available. For MLEs of Tweedie distribution, Dunn and Smyth (2005) proposed R package

```
library(tweedie)
tweedie.profile(formula, p.vec, smooth=FALSE, do.plot=FALSE,
do.ci=smooth, eps=1/6, method="series", conf.level=0.95,
phi.method=ifelse(method == "saddlepoint", "saddlepoint", "mle"))
```

and for mle of Beta distribution Millard (2013) proposed R package

```
install.package('EnvStats')
library(EnvStats)
ebeta(x, method = "mle")
```

Regarding the above discussions, to obtain the maximum likelihood estimate of the parameter $\boldsymbol{\theta}$ for the τ^{th} quantile, we can apply the following iterative procedure.

Algorithm 2 MCEM Algorithm

1: Initialize the parameter $\mathbf{\Omega}^{(t)} = \{\boldsymbol{\beta}_\tau^{(t)}, \sigma^{(t)}, \boldsymbol{\theta}^{(t)}\}$. Set $t = 0$ and substitute $\mathbf{\Omega}^{(t)}$ to Equation 3.42;

2: Draw a sample $\boldsymbol{v}_i^{(t)} = \{v_{i1}^{(t)}, v_{i2}^{(t)}, ..., v_{ik}^{(t)}\}$ from the density $f(u_i|y_i, \mathbf{\Omega}^{(t)})$, for $i = 1, ..., n$, independently, by using any sampler, i.e., Metropolis-Hasting sampling algorithm;

3: Maximize the expected value of $Q^*(\mathbf{\Omega}|\mathbf{\Omega}^{(t)})$. Let $\tilde{z}_{ij}^{(t)} = y_i + v_{ij}^{(t)}$, solve a minimization problem and let

$$\hat{\beta}_\tau^{(t+1)} \equiv \underset{\boldsymbol{\beta} \in \mathbb{R}^{\boldsymbol{p}}}{\text{argmin}} \sum_i^n \frac{1}{k} \sum_{j=1}^k \rho(\tilde{z}_{ij}^{(t)} - T^{-1}(x_i'\beta)), \tag{3.47}$$

and

$$\hat{\sigma}^{(t+1)} = \frac{1}{nk} \sum_{i=1}^n \sum_{j=1}^k \rho(\tilde{z}_{ij}^{(t)} - T^{-1}(x_i'\hat{\beta}_\tau^{(t+1)})), \tag{3.48}$$

then, let $\boldsymbol{\theta}^{(t+1)}$ be the maximum likelihood estimate of density $f(v_i^{(t)}|\boldsymbol{\theta})$;

4: Set $t = t + 1$ and repeat step 1-3 until the parameter $\mathbf{\Omega}$ reaches the convergence.

Remark. *The sensitivity of MH sampler depends on the burn-in size and sample size. The MH sampler may generate variability of estimates, but a large burn-in size will address this problem as well as is required to ensure the convergence of generated samples. Take the computational cost into account, we use a burn-in sample size of 5000 and constant MC sample size of 5000. Various sample sizes k_i for different subjects, $i = 1, ..., n$, is also applicable. It is feasible to use other sampling techniques, i.e., Gibbs sampler.*

Now, the last issue is the exact upper bound of generated noises $u_i \sim f(u|\boldsymbol{\theta}^{\text{Noise}})$. Unfortunately, letting $\boldsymbol{\theta}^{\text{Noise}} = \hat{\boldsymbol{\theta}}^{\text{ALD}}$ does not automatically guarantee the generated values of noise are all less than 1 in practice. There are approximately 1% $\sim$ 5% of noises that equal or are greater than 1. In order to satisfy **A2**, we have to manually adjust them. Since the proportion of these 'unwanted' noises is very small, we shall re-generate a new value once the 'unwanted' value occurs, until it is less than 1. This approach ensures the generated noises approximately follow the Tweedie distribution. Finally, it is possible to guess the initial parameters. For $\boldsymbol{\beta}_\tau^{(0)}$, we can let it be the estimates obtained from QR model using existing uniform jittering method. For Tweedie distribution, the initial parameters can be chosen if most of random samples are less than 1 by some previous experiments, for example, previous experiments show $\mu^{(0)} = 0.5, \phi^{(0)} = 0.2, p^{(0)} = 1.5$ are good initial parameters. For beta distribution, we can simply set $a^{(0)} = 1, b^{(0)} = 1$.

3.4 Asymptotic Properties of Proposed Estimators

In this section, we derive asymptotic properties and distributions for estimates obtained from the proposed method. To ensure that the statistical inference is valid, the following regularity conditions are given with modifications (Machado & Silva, 2005; J. Powell, 1986).

(A3) The regressors have bounded second moment, i.e., $E[\|\boldsymbol{x}_i\|^2] < \infty$. The regressors $\boldsymbol{X}' = (\mathbf{1}, \ldots, \boldsymbol{x}_k)$ can be partitioned as $(\mathbf{1}, \boldsymbol{x}'^{(c)})$ where $\boldsymbol{x}'^{(c)} \in \mathbb{R}^{k_c}$, $1 \le k_b \le k-1$, satisfying $P(\boldsymbol{x}'^{(c)} \in \mathcal{B}) = 0$ for any countable subset $\mathcal{B}$ of $\mathbb{R}^{k_b}$;

(A4) Let $Z = Y + U$, where U is a Tweedie or Beta random variable, independent of $\boldsymbol{X}$ and Y. Denoting $T(\cdot;\tau)$ as a known monotone transformation, possibly depending on τ, the following restriction on the QR of Z given X holds:

$$Q_{T(Z;\tau)}(\tau|\boldsymbol{X}) = \boldsymbol{X}'\boldsymbol{\beta}_\tau \quad \text{for} \quad \tau \in (0,1),$$

and parameter β lies in a compact set $\mathcal{C}$ in $\mathbb{R}^p$. Furthermore, if $\beta_\tau^{(c)}$ denotes the components of β_τ corresponding to the continuous covariates $\boldsymbol{x}'^{(c)}$, then $\beta_\tau \neq 0$;

(A5) The error terms $\varepsilon_{\tau i} = z_i - \boldsymbol{x}_i'\boldsymbol{\beta}_\tau$ are continuously distributed given $\boldsymbol{x}_i$, with conditional density $f(\varepsilon|\boldsymbol{x}_i)$ satisfying the conditional quantile restriction, i.e.

$$\int_{-\infty}^{Q_\varepsilon(\tau|\boldsymbol{x}_i)} f(\lambda|\boldsymbol{x}_i)d\lambda = \tau,$$

where τ is the quantile level of interest, and $Q_\varepsilon(\cdot|\cdot)$ denotes the conditional quantile function of ε given X on the τ^{th} quantile of distribution;

(A6) The regressors and density of $T(Z;\tau)$ satisfy a "local identification" condition, that is, the matrix

$$\boldsymbol{D} = \mathbb{E}[f_{T(Z;\tau)}(\boldsymbol{X}'\boldsymbol{\beta}_\tau|\boldsymbol{X})\boldsymbol{X}\boldsymbol{X}']$$

is positive definite.

3.4.1 Consistency and Asymptotic Normality of Simple Jittering QR Estimator

Under jittering process, QR models are fit to Z but no longer Y, where Z is continuous. So, asymptotic properties of jittering QR estimator will simply follow the standard QR given by Koenker (2005). Based on the consistency of simple jittering estimator of median QR shown by (J. Powell, 1986), that of proposed QR model can be proved.

Theorem 3.4.1 (Consistency of Simple Jittering QR Estimator).
The $(y_i, \boldsymbol{x}_i, u_i)_{i=1}^n$ are a random sample of $(Y, \boldsymbol{X}, U)$. Let the QR estimators be the solution of a minimization problem

$$\hat{\boldsymbol{\beta}}_\tau \equiv \arg\min_{\boldsymbol{\beta}\in\mathbb{R}^p} \frac{1}{n}\sum_{i=1}^n \rho_\tau(z_i - \boldsymbol{x}_i'\boldsymbol{\beta}).$$

*Under conditions **A1** to **A6**, $\hat{\boldsymbol{\beta}}_\tau$ is consistent with the true QR coefficients, $\boldsymbol{\beta}_\tau$, i.e.,*

$$\hat{\boldsymbol{\beta}}_\tau \xrightarrow{P} \boldsymbol{\beta}_\tau.$$

Proof.

Also see J. Powell (1986) for median regression. We denote $\boldsymbol{\beta}_\tau$ as the vector of true QR coefficients. To prove the consistency of QR estimators clearly, it is more convenient to subtract the objective function of true QR coefficients, $S_n(\boldsymbol{\beta}_\tau)$, off the minimization problem because it will not affect the minimization solution. That is, we define

$$\begin{aligned}
\hat{\boldsymbol{\beta}}_\tau &\equiv \underset{\boldsymbol{\beta}\in\mathbb{R}^p}{\operatorname{argmin}} S_n(\boldsymbol{\beta}) - S_n(\boldsymbol{\beta}_\tau) \\
&= \underset{\boldsymbol{\beta}\in\mathbb{R}^p}{\operatorname{argmin}} \frac{1}{n}\sum_{i=1}^{n} \rho_\tau(T(z_i;\tau) - \boldsymbol{x}_i'\boldsymbol{\beta}) - \rho_\tau(T(z_i;\tau) - \boldsymbol{x}_i'\boldsymbol{\beta}_\tau) \\
&= \underset{\boldsymbol{\beta}\in\mathbb{R}^p}{\operatorname{argmin}} \frac{1}{n}\sum_{i=1}^{n} \rho_\tau(\varepsilon_{\tau i} - \boldsymbol{x}_i'\boldsymbol{\delta}_\tau) - \rho_\tau(\varepsilon_{\tau i})
\end{aligned}$$

where $\boldsymbol{\delta}_\tau \equiv \boldsymbol{\beta} - \boldsymbol{\beta}_\tau$. By the triangle and Cauchy-Schwarz inequalities, we have

$$-\|\boldsymbol{x}_i\| \cdot \|\boldsymbol{\delta}_\tau\| \leq |\varepsilon_{\tau i} - \boldsymbol{x}_i'\boldsymbol{\delta}_\tau| - |\varepsilon_{\tau i}| \leq \|\boldsymbol{x}_i\| \cdot \|\boldsymbol{\delta}_\tau\|.$$

Therefore, the normalized minimand is the average value of i.i.d. random variables with finite first and second moments under assumption **A2**. By Khintchine's Law of Large Numbers, which strengthens the traditional Weak Law of Large Numbers, we have

$$
\begin{aligned}
S_n(\boldsymbol{\beta}) - S_n(\boldsymbol{\beta}_\tau) \xrightarrow{p} \bar{S}_n(\boldsymbol{\delta}_\tau)& \\
&\equiv \mathbb{E}[S_n(\boldsymbol{\beta}) - S_n(\boldsymbol{\beta}_\tau)] \\
&= \mathbb{E}\left[\frac{1}{n}\sum_{i=1}^{n} \rho_\tau(\varepsilon_{\tau i} - \boldsymbol{x}_i'\boldsymbol{\delta}_\tau) - \rho_\tau(\varepsilon_{\tau i})\right] \\
&= \mathbb{E}[\rho_\tau(\varepsilon_{\tau i} - \boldsymbol{x}_i'\boldsymbol{\delta}_\tau) - \rho_\tau(\varepsilon_{\tau i})] \\
&= \mathbb{E}[(\varepsilon_{\tau i} - \boldsymbol{x}_i'\boldsymbol{\delta}_\tau)(\tau - I(\varepsilon_{\tau i} - \boldsymbol{x}_i'\boldsymbol{\delta}_\tau < 0)) - \varepsilon_{\tau i}(\tau - I(\varepsilon_{\tau i} < 0))] \\
&= \mathbb{E}[\mathbb{E}[(\varepsilon_{\tau i} - \boldsymbol{x}_i'\boldsymbol{\delta}_\tau)(\tau - I(\varepsilon_{\tau i} - \boldsymbol{x}_i'\boldsymbol{\delta}_\tau < 0)) - \varepsilon_{\tau i}(\tau - I(\varepsilon_{\tau i} < 0))|\boldsymbol{x}_i]] \\
&= \mathbb{E}[\mathbb{E}[(\varepsilon_{\tau i} - \boldsymbol{x}_i'\boldsymbol{\delta})(\tau - I(\varepsilon_{\tau i} - \boldsymbol{x}_i'\boldsymbol{\delta}_\tau < 0)) - (\varepsilon_{\tau i} - \boldsymbol{x}_i'\boldsymbol{\delta}_\tau)(\tau - I(\varepsilon_{\tau i} < 0))|\boldsymbol{x}_i]] \\
&= \mathbb{E}[\mathbb{E}[(\varepsilon_{\tau i} - \boldsymbol{x}_i'\boldsymbol{\delta}_\tau)((\tau - I(\varepsilon_{\tau i} - \boldsymbol{x}_i'\boldsymbol{\delta}_\tau < 0)) - (\tau - I(\varepsilon_{\tau i} < 0)))|\boldsymbol{x}_i]] \\
&= \mathbb{E}\left[2\int_{\boldsymbol{x}_i'\boldsymbol{\delta}_\tau}^{0} (\lambda - \boldsymbol{x}_i'\boldsymbol{\delta}_\tau) f(\lambda|\boldsymbol{x}_i) d\lambda\right],
\end{aligned}
$$

where the equation $\mathbb{E}[\mathbb{E}[\boldsymbol{x}_i'\boldsymbol{\delta}_\tau(\tau - I(\varepsilon_{\tau i} < 0))|\boldsymbol{x}_i]] = 0$ is added to the second-to-last equality. Also, the last equality is well defined for both positive value and negative value of $\boldsymbol{x}_i'\boldsymbol{\delta}_\tau$ (J. Powell, 1986). It is easy to check that the limit of $\bar{S}_n(\boldsymbol{\delta}_\tau) = 0$ at $\boldsymbol{\delta}_\tau = \boldsymbol{\beta} - \boldsymbol{\beta}_\tau = \mathbf{0}$ and is non-negative elsewhere. Since $\boldsymbol{\beta} - \boldsymbol{\beta}_\tau$ is convex for all n, the probability limit of $\bar{S}_n(\boldsymbol{\beta} - \boldsymbol{\beta}_\tau)$ is zero as well. Therefore, if $\boldsymbol{\beta} = \boldsymbol{\beta}_\tau$ is a unique local minimum, it is also a unique global minimum. Notice that since

$$
\begin{aligned}
\frac{\partial \bar{S}(\boldsymbol{\delta}_\tau)}{\partial \boldsymbol{\delta}_\tau} &= -2E\left[\boldsymbol{x}_i \int_{\boldsymbol{x}_i'\boldsymbol{\delta}_\tau}^{0} f(\lambda|\boldsymbol{x}_i) d\lambda\right], \\
\frac{\partial \bar{S}(\mathbf{0})}{\partial \boldsymbol{\delta}_\tau} &= \mathbf{0},
\end{aligned}
$$

and,

$$\frac{\partial^2 \bar{S}(\boldsymbol{\delta}_\tau)}{\partial \boldsymbol{\delta}_\tau \partial \boldsymbol{\delta}'_\tau} = 2E[\boldsymbol{x}_i \boldsymbol{x}'_i f(\boldsymbol{x}'_i \boldsymbol{\delta}_\tau | \boldsymbol{x}_i)],$$
$$\frac{\partial^2 \bar{S}(\mathbf{0})}{\partial \boldsymbol{\delta}_\tau \partial \boldsymbol{\delta}'_\tau} = 2E[\boldsymbol{x}_i \boldsymbol{x}'_i f(\mathbf{0} | \boldsymbol{x}_i)],$$

which is positive definite by condition **A4**. This implies that $\boldsymbol{\delta}_\tau = 0$ is a unique local minimum of $\bar{S}(\boldsymbol{\delta}_\tau)$, and yields that it is the unique global minimum of it. Therefore, $\boldsymbol{\beta}_\tau$ converges in probability to $\boldsymbol{\beta}_{\tau 0}$, i.e.,

$$\hat{\boldsymbol{\beta}}_\tau \xrightarrow{P} \boldsymbol{\beta}_\tau.$$

Hence, it shows consistency of $\hat{\boldsymbol{\beta}}_\tau$. □

Theorem 3.4.2 (Asymptotic Distribution of Simple Jittering QR Estimators)**.** *Let $(y_i, \boldsymbol{x_i}^T, u_i), i = 1, ..., n$, be a set of random samples from the population $(Y, \boldsymbol{X}, U)$. Let $Z = Y + U$ be the response and*

$$Q_{T(Z;\tau)}(\tau | \boldsymbol{X}) = \boldsymbol{X}' \boldsymbol{\beta}_\tau,$$

where $T(\cdot)$ is a suitable monotonic transformation function. If the QR estimator, $\hat{\beta}_\tau$, is defined by the minimization of check loss function,

$$\hat{\boldsymbol{\beta}}_\tau = \arg\min_{\boldsymbol{\beta} \in \mathbb{R}^p} S(\boldsymbol{\beta}) = \arg\min_{\boldsymbol{\beta} \in \mathbb{R}^p} \sum_{i=1}^{n} \rho_\tau(T(z_i; \tau) - \boldsymbol{x}'_i \boldsymbol{\beta})),$$

*where $z_i = y_i + u_i$ and $\rho_\tau(u) = u(\tau - I(u < 0))$, then, under conditions **A1** to **A6**, the asymptotic distribution of the estimator $\hat{\boldsymbol{\beta}}_\tau$ is*

$$\sqrt{n}(\hat{\boldsymbol{\beta}}_\tau - \boldsymbol{\beta}_\tau) \xrightarrow{D} \mathcal{N}(\mathbf{0}, \boldsymbol{D}^{-1} \boldsymbol{A} \boldsymbol{D}^{-1}). \quad (3.49)$$

with $\boldsymbol{A} = \tau(1-\tau)E[\boldsymbol{X}\boldsymbol{X}']$ *and* $\boldsymbol{D} = E[f_{T(Z;\tau)}(\boldsymbol{X}'\boldsymbol{\beta}_\tau|\boldsymbol{X})\boldsymbol{X}\boldsymbol{X}']$*, where* $f_T(Z;\tau)(\cdot|\cdot)$ *is the conditional density of* $T(Z;\tau)$ *given* $\boldsymbol{X}$*.*

Proof.

Let Z be continuous and the QR model $Q_{T(Z;\tau)}(\tau|\boldsymbol{X}) = \boldsymbol{X}'\boldsymbol{\beta}_\tau$. The estimator $\hat{\boldsymbol{\beta}}_\tau$ minimize the sample objective function $S(\boldsymbol{\beta}) = \sum_{i=1}^{n} \rho_\tau(T(z_i;\tau) - \boldsymbol{x}_i'\boldsymbol{\beta})$. Since the continuity assumption is satisfied in the jittering process, the asymptotic property of the QR estimator simply follows that of standard QR model, which is given and proven by Koenker (2005). □

Corollary 3.4.1.

Under conditions ***A1*** *to* ***A6****, the approximate covariance matrix of* $\hat{\boldsymbol{\beta}}_\tau$ *can be obtained from the asymptotic normal distribution, that is,*

$$Cov(\hat{\boldsymbol{\beta}}_\tau) = \frac{1}{n}\boldsymbol{D}^{-1}\boldsymbol{A}\boldsymbol{D}^{-1}. \tag{3.50}$$

3.4.2 Consistency and Asymptotic Normality of Average Jittering QR Estimator

In this section, we will discuss the asymptotic properties and distribution of QR estimators with the application of averaging-out noises technique.

Theorem 3.4.3 (Consistency of Average Jittering QR Estimators)**.**

The $(y_i, \boldsymbol{x}_i, u_i)_{i=1}^n$ *are a random sample of* $(Y, \boldsymbol{X}, U)$*. Let* $(y_i + u_i^{(l)}, \boldsymbol{x}_i)_{i=1}^n$ *be jittered sample and*

$$\hat{\boldsymbol{\beta}}_\tau^m = \frac{1}{m}\sum_{l=1}^{m}\hat{\boldsymbol{\beta}}_\tau^{(l)}, n \to \infty,$$

where $\hat{\boldsymbol{\beta}}_\tau^{(l)}$ *is the QR estimator based on* $(y_i + u_i^{(l)}, \boldsymbol{x}_i)$ *and* m *is fixed number of jittered samples. Then, under conditions* ***A1*** *to* ***A6****, the average-jittering QR estimator* $\hat{\boldsymbol{\beta}}_\tau^m$

is consistent with true QR coefficients , $\boldsymbol{\beta}_\tau$*, i.e.,*

$$\hat{\boldsymbol{\beta}}_\tau^A \xrightarrow{P} \boldsymbol{\beta}_\tau.$$

Proof.

For $l = 1, 2, ..., m$, the consistency of each QR estimator $\hat{\boldsymbol{\beta}}_\tau^{(l)}$ follows the result of Theorem 3.4.1 such that $\hat{\boldsymbol{\beta}}_\tau^{(l)} \xrightarrow{P} \boldsymbol{\beta}_\tau$. Therefore, for each $l \in (1, ..., m)$,

$$\lim_{n\to\infty} \Pr(|\hat{\boldsymbol{\beta}}_\tau^{(l)} - \boldsymbol{\beta}_\tau| < \delta) = 0,$$

hence there exists a N_l such that if $n > N_l$, then

$$\Pr(|\hat{\boldsymbol{\beta}}_\tau^{(l)} - \boldsymbol{\beta}_\tau| > \delta) < \frac{\epsilon}{m}$$

for all $\epsilon > 0$. By triangle inequality, we have

$$\begin{aligned} |\hat{\boldsymbol{\beta}}_\tau^A - \boldsymbol{\beta}_\tau| &= \left| \frac{1}{m} \sum_{l=1}^m \hat{\boldsymbol{\beta}}_\tau^{(l)} - \boldsymbol{\beta}_\tau \right| \\ &\leq \frac{1}{m} \sum_{l=1}^m |\hat{\boldsymbol{\beta}}_\tau^{(l)} - \boldsymbol{\beta}_\tau|. \end{aligned}$$

Then,

$$\begin{aligned} \Pr(|\hat{\boldsymbol{\beta}}_\tau^m - \boldsymbol{\beta}_\tau| > \delta) &= \Pr\left(\left| \frac{1}{m} \sum_{l=1}^m \hat{\boldsymbol{\beta}}_\tau^{(l)} - \boldsymbol{\beta}_\tau \right| > \epsilon \right) \\ &\leq \Pr\left(\frac{1}{m} \sum_{l=1}^m |\hat{\boldsymbol{\beta}}_\tau^{(l)} - \boldsymbol{\beta}_\tau| > \delta \right) \\ &\leq \sum_{l=1}^m \Pr\left(|\hat{\boldsymbol{\beta}}_\tau^{(l)} - \boldsymbol{\beta}_\tau| > \epsilon \right) < \epsilon, \end{aligned}$$

if $n \geq N = \max_{1 \leq l \leq m}\{N_l\}$. Thus, $\Pr(|\hat{\boldsymbol{\beta}}_\tau^A - \boldsymbol{\beta}_\tau| > \delta)$ converges to zero as $n \to \infty$. $\square$

Recall the procedure of average jittering, $m-1$ jittered samples are additionally generated for averaging the estimates. Therefore, the asymptotic covariance matrix of $(\hat{\boldsymbol{\beta}}_\tau^A - \boldsymbol{\beta}_\tau)$ can be written as the weighted sum of $\boldsymbol{D}^{-1}\boldsymbol{A}\boldsymbol{D}^{-1}$ and asymptotic covariance matrix of estimator for different jittered samples with a certain speed of convergence. Fixing the distribution parameters of U when averaging noise out, asymptotic normal distribution of normalized difference can be obtained as follows,

Theorem 3.4.4 (Asymptotic Normality of Average Jittering QR Estimator)**.** *The dataset $\{y_i, \boldsymbol{x}_i, u_i\}_{i=1}^n$ are a random sample taken from the population $\{Y, \boldsymbol{X}, U\}$ and are satisfying all the above conditions. Let $\boldsymbol{\beta}_\tau$ be the true QR coefficients. Then, the averaging jittered QR estimator follows normal distribution asymptotically such that*

$$\sqrt{n}(\hat{\boldsymbol{\beta}}_\tau^m - \boldsymbol{\beta}_\tau) \xrightarrow{D} \mathcal{N}(\mathbf{0}, \boldsymbol{V}^m), \text{ as } n \to \infty, \tag{3.51}$$

where $\hat{\boldsymbol{\beta}}_\tau^m$ is the averaging jittered QR estimator, and

$$\boldsymbol{V}^m = \frac{1}{m}\boldsymbol{D}^{-1}\boldsymbol{m}\boldsymbol{D}^{-1} + \Big(1 - \frac{1}{m}\Big)\boldsymbol{D}^{-1}\boldsymbol{B}\boldsymbol{D}^{-1}, \tag{3.52}$$

where $m = 1, 2, \ldots$ is the number of jittered samples, matrix $\boldsymbol{A}$ and $\boldsymbol{D}$ are given in Theorem 3.4.2, and

$$\boldsymbol{B} = \mathbb{E}\Big\{\Big[-\tau^2 + \sum_{y=0}^{Q_Y(\tau|\boldsymbol{X})} P(Y = y|\boldsymbol{X}) F_U(z_\tau - y)^2\Big]\boldsymbol{X}\boldsymbol{X}'\Big\} \tag{3.53}$$

where F_U is the CDF of U. Substituting CDFs of U and rearranging terms, we have

1. *For $U \sim Tw_p(\mu, \phi), 1 < p < 2$, and $U \in [0, \infty)$,*

$$\boldsymbol{B} = \mathbb{E}\Big\{\Big[-\tau^2 + \sum_{y=0}^{Q_Y(\tau|\boldsymbol{X})} P(Y = y|\boldsymbol{X})\Big[\Big(\sum_{k=0}^{\infty}\frac{1}{\Gamma(\alpha)}\gamma(k\alpha, \beta(z_\tau - y))\frac{e^{-\lambda}\lambda^k}{k!}\Big)^2\Big]\Big]\boldsymbol{X}\boldsymbol{X}'\Big\},$$

where $\lambda = \frac{\mu^{2-p}}{\phi(2-p)}$,$\alpha = \frac{2-p}{p-1}$ and $\beta = \phi(p-1)\mu^{p-1}$.

2. For $U \sim Beta(a,b)$, and $U \in (0,1)$

$$\boldsymbol{B} = \mathbb{E}\Big\{\Big[\tau(1-\tau) - f_{Y|\boldsymbol{X}}(Q_Y(\tau|\boldsymbol{X}))$$
$$\Big(Q_{Z|U}(\tau|\boldsymbol{X}) - Q_Y(\tau|\boldsymbol{X}) - \Big(\frac{B(Q_{Z|U}(\tau|\boldsymbol{X}) - Q_Y(\tau|\boldsymbol{X}); a, b)}{B(a,b)}\Big)^2\Big)\Big]\boldsymbol{X}\boldsymbol{X}'\Big\}.$$

Proof.

Machado and Silva (2005) shows that for fixed value of m, the asymptotic normality of averaging jittered QR estimator follows the result from Theorem 3.4.2 since the product of $\boldsymbol{D}^{-1}\boldsymbol{A}\boldsymbol{D}^{-1}$ is already the asymptotic covariance matrix of each $\sqrt{n}(\hat{\boldsymbol{\beta}}_\tau^{(l)} - \boldsymbol{\beta}_\tau)$, where $l = 1, 2, ..., m$. Then, it remains to evaluate the $m(m-1)$ asymptotic covariance matrix of the estimator for different sets of jittered sample.

Let $\omega_{ij} = \boldsymbol{\omega}(y_i, u_i^{(l)}, \boldsymbol{x}_i) \equiv \big[\tau - I_{\{y_i + u^{(l)} \le T^{-1}(\boldsymbol{x}_i'\boldsymbol{\beta}_\tau)\}}\big]\boldsymbol{x}_i$, so that

$$\sqrt{n}(\hat{\boldsymbol{\beta}}_\tau^{(l)} - \boldsymbol{\beta}\tau) = -\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\boldsymbol{D}^{-1}\boldsymbol{\omega}_{ij} + o_p(1).$$

Because $E(\boldsymbol{\omega}_{il}\boldsymbol{\omega}_{ik}) = 0$ for $l \neq k$,

$$\mathbb{E}\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\boldsymbol{D}^{-1}\boldsymbol{\omega}_{il}\right)\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\boldsymbol{D}^{-1}\boldsymbol{\omega}_{ik}\right)' = \boldsymbol{D}^{-1}\mathbb{E}(\boldsymbol{\omega}_{il}\boldsymbol{\omega}_{ik}')\boldsymbol{D}^{-1}.$$

It remains to evaluate $\mathbb{E}(\boldsymbol{\omega}_{il}\boldsymbol{\omega}_{ik})$ for $l \neq k$. To simplify the notation, put $z_{\tau_i} \equiv T^{-1}(\boldsymbol{x}_i'\boldsymbol{\beta}_\tau) \equiv Q_z(\tau|\boldsymbol{x}_i)$ and $y_{\tau_i} \equiv Q_{Y|\boldsymbol{x}_i}(\tau) \equiv \lceil Q_z(\tau|\boldsymbol{x}_i) - 1\rceil$. Apply the law of total expectation in propability theory, the relevant factor in the expectation $E(\boldsymbol{\omega}_{il}\boldsymbol{\omega}_{ik}')$

conditional on $\boldsymbol{X} = \boldsymbol{x}$ is

$$\begin{aligned}
&\mathbb{E}\left[(\tau - I\{y + u^{(l)} \leq z_\tau\})(\tau - I\{y + u^{(k)} \leq z_\tau\})|\boldsymbol{x}\right] \\
&= \tau^2 - \tau\mathbb{E}\left[I\{y + u^{(l)} \leq z_\tau\}|x\right] - \tau\mathbb{E}\left[I\{y + u^{(k)} \leq z_\tau\}|\boldsymbol{x}\right] \\
&\quad + \mathbb{E}\left[I\{y + u^{(l)} \leq z_\tau\}I\{y + u^{(k)} \leq z_\tau\}|\boldsymbol{x}\right] \\
&= \mathbb{E}\left[I\{y + u^{(l)} \leq z_\tau\}I\{y + u^{(k)} \leq z_\tau\}|\boldsymbol{x}\right] - \tau^2 \\
&= \mathbb{E}_{Y|\boldsymbol{x}}\left[\mathbb{E}\left[I\{u^{(l)} \leq z_\tau - y\}I\{u^{(k)} \leq z_\tau - y\}|\boldsymbol{x}, y\right]\right] - \tau^2 \\
&= \mathbb{E}_{Y|\boldsymbol{x}}\left[F_U(z_\tau - y)^2\right] - \tau^2. \qquad (*)
\end{aligned}$$

where $F_U(\cdot)$ is the distribution of U. Now, it remains to find the expression of CDF of U, $F_U(z_\tau - y)^2$. We derive for two cases where U taken from Tweedie distribution and Beta distribution one by one.

For U following Tweedie distribution, i.e., $U \sim \mathrm{Tw}_p(\mu, \phi)$, where distribution parameters μ, ϕ, p are known, by Proposition 3.2.1, the cdf of Tweedie variables can be written in compound Poisson-Gamma distribution. We substitute the value of $(z_\tau - y)$ to the cdf of Tweedie distribution shown in Equation 3.5, giving the support of Tweedie distribution $[0, +\infty)$, the cdf is written as

$$F_U(z_\tau - y)^2 = I_{\{y \leq z_\tau\}}\left(\sum_{k=0}^{\infty} \frac{1}{\Gamma(\alpha)}\gamma(k\alpha, \beta(z_\tau - y))\frac{e^{-\lambda}\lambda^k}{k!}\right)^2,$$

and the corresponding expectation is

$$\begin{aligned}
&E_{Y|\boldsymbol{x}}\left[F_U(z_\tau - y)^2\right] \\
&= \sum_{y=0}^{\lceil z_\tau - 1\rceil} P(Y = y|\boldsymbol{x})\left[I_{\{y \leq z_\tau\}}\left(\sum_{k=0}^{\infty} \frac{1}{\Gamma(\alpha)}\gamma(k\alpha, \beta(z_\tau - y))\frac{e^{-\lambda}\lambda^k}{k!}\right)^2\right] \\
&= \sum_{y=0}^{y_\tau} P(Y = y|\boldsymbol{x})\left[\left(\sum_{k=0}^{\infty} \frac{1}{\Gamma(\alpha)}\gamma(k\alpha, \beta(z_\tau - y))\frac{e^{-\lambda}\lambda^k}{k!}\right)^2\right].
\end{aligned}$$

Subtracting τ^2, we get the desired result that Tweedie distributed additive noise is used. The formulas to obtain parameters of compound Poisson-Gamma, i.e., λ, α, β, can be found in Proposition 3.2.1.

For U following the Beta distribution, i.e., $U \sim \text{Beta}(a, b)$, where the distribution parameters a, b are known, the cdf of Beta distribution can be expressed by using beta function and incomplete beta function. We substitute $(z_\tau - y)^2$ to Equation 3.8, the simplified cdf of Beta distribution. Since the support of Beta distributed random variable is $[0, 1]$, the CDF is written as

$$F_U(z_\tau - y)^2 = I_{\{y<z_\tau-1\}} + I_{\{z_\tau-1\leq y\leq z_\tau\}}\left(\frac{B(z_\tau - y; a, b)}{B(a, b)}\right)^2,$$

and the expectation

$$\begin{aligned}
&\mathbb{E}_{Y|\boldsymbol{x}}\left[F_U(z_\tau - y)^2\right] \\
&= P(Y < z_\tau - 1|\boldsymbol{x}) + P(Y = \lceil\, z_\tau - 1\rceil|\boldsymbol{x})\left(\frac{B(z_\tau - y_\tau; a, b)}{B(a, b)}\right)^2 \\
&= P(Y < y_\tau|\boldsymbol{x}) + P(Y = y_\tau|\boldsymbol{x})\left(\frac{B(z_\tau - y_\tau; a, b)}{B(a, b)}\right)^2 \\
&= \tau - P(Y = y_\tau|\boldsymbol{x})(z_\tau - y_\tau) + P(Y = y_\tau|\boldsymbol{x})\left(\frac{B(z_\tau - y_\tau; a, b)}{B(a, b)}\right)^2 \\
&= \tau - P(Y = y_\tau|\boldsymbol{x})\left[z_\tau - y_\tau - \left(\frac{B(z_\tau - y_\tau; a, b)}{B(a, b)}\right)^2\right],
\end{aligned}$$

where $P(Y < z_\tau|\boldsymbol{x}) = \tau - P(Y = y_\tau|\boldsymbol{x})(z_\tau - y_\tau)$ follows the Equation 2.53 with terms rearrangement. Subtracting τ^2, we get the desired result when Beta distributed additive noise is used. Note that Machado and Silva (2005), proved the uniform perturbation case. □

Remark. *Compare the asymptotic normal distributions of simple jittering estimator and average estimator in Theorem 3.4.2 and 3.4.4, respectively, changing noise distribution does not change the asymptotic covariance matrix of QR estimators using* *simple jittering method, but it does matter for that using average jittering method. Therefore, changing noise distribution requires modifications on Theorem 3.4.4.*

Corollary 3.4.2. *Under conditions **A1** to **A6**, covariance matrix of $\hat{\boldsymbol{\beta}}_\tau^A$ can be obtained from the asymptotic normal distribution, that is,*

$$Cov(\hat{\boldsymbol{\beta}}_\tau^A) = \frac{1}{n}\left[\frac{1}{m}\boldsymbol{D}^{-1}\boldsymbol{A}\boldsymbol{D}^{-1} + \left(1 - \frac{1}{m}\right)\boldsymbol{D}^{-1}\boldsymbol{B}\boldsymbol{D}^{-1}\right]. \tag{3.54}$$

3.4.3 Consistent Estimator of Asymptotic Covariance Matrices

The variance-covariance matrix can be consistently estimated, see Machado and Silva (2005). Assume that the following assumptions are satisfied.

(A7) $T^{-1}(\cdot)$ is twice continuously differentiable, that is, obtaining $T^{(1)} \equiv \frac{\partial T^{-1}(z;\tau)}{\partial z}$ and $T^{(2)} \equiv \frac{\partial^2 T^{-1}(z;\tau)}{\partial z^2}$ are valid;

(A8) Two expectations

$$(a)\mathbb{E}[|T^{-1}(\boldsymbol{x}_i'\boldsymbol{\beta}_\tau)|||\boldsymbol{x}_i||^2], \text{ and}$$
$$(b)\mathbb{E}\left[\sup_{||\beta-\beta_\tau||\le\delta}|T^{(1)}(\boldsymbol{x}_i'\boldsymbol{\beta}_\tau)|^l||\boldsymbol{x}_i||^2\right] \text{ for } l = 1, 2 \text{ and for some } \delta > 0,$$

exist;

(A9) There is a sequence c_n of real numbers in $(0, 0.5)$ such that $c_n = o_p(1)$ and

$$\frac{\sup_{1\le i\le n}|T^{-1}(\boldsymbol{x}_i'\hat{\boldsymbol{\beta}}_\tau) - T^{-1}(x_i'\boldsymbol{\beta}\tau)|}{c_n} = o_p(1).$$

Note that if $\sup_{1\le i\le n}\|\boldsymbol{x}_i\| = O_p(1)$, Theorem 3.4.2 actually implies that

$$\sup_{1\le i\le n} |\boldsymbol{x}_i'\hat{\boldsymbol{\beta}}_\tau - x_i'\boldsymbol{\beta}\tau| = O_p(\frac{1}{\sqrt{n}}),$$

Since the first derivative of transformation function $T^{(1)}(\cdot)$ is continuous and Γ is compact, we have

$$\sup_{1\le i\le n} |T^{-1}(\boldsymbol{x}_i'\hat{\boldsymbol{\beta}}_\tau) - T^{-1}(x_i'\boldsymbol{\beta}\tau)| = O_p(\frac{1}{\sqrt{n}}).$$

and c_n should converge to zero more slowly than $1/\sqrt{n}$, i.e., $c_n\sqrt{n}\to\infty$. We have the following theorems.

Theorem 3.4.5 (Consistent Estimator of Matrix A)**.**

Let the set $\{z_i, \boldsymbol{x}_i, u_i\}$ be a sample from the data set $\{Z, X, U\}$, where $i = 1, 2, ..., n$. Let

$$\hat{\boldsymbol{A}}_n = \frac{\tau(1-\tau)}{n}\sum_{i=1}^{n}\boldsymbol{x}_i\boldsymbol{x}_i'. \tag{3.55}$$

*Under conditions from **A1** to **A9**, $\hat{\boldsymbol{A}}_n$ converges in probability to $\boldsymbol{A}$, i.e.,*

$$\hat{\boldsymbol{A}}_n \xrightarrow{P} \boldsymbol{A}$$

as $n\to\infty$.

Proof.

Let $\boldsymbol{x}_i$ be a sample from the data set. Under the conditions, samples $\boldsymbol{x}_i$ are *i.i.d.*. It follows that the product of $\boldsymbol{x}_i\boldsymbol{x}_i'$ are also *i.i.d.*, since each $\boldsymbol{x}_i\boldsymbol{x}_i'$ is a function of $\boldsymbol{x}_i$. Then, the Law of Large Number (LLN) is applicable. Denote $a_{jk}^{(i)}$ as the $(j,k)^{th}$ element of product $\boldsymbol{x}_i\boldsymbol{x}_i'$. It indicates that for any element a_{jk} of product $\boldsymbol{x}_i\boldsymbol{x}_i'$, the sample mean $\frac{1}{n}\sum_{i=1}^{n} a_{jk}^{(i)}$ converges in probability to the expectation $E[a_{jk}]$ as $n\to\infty$.

Hence,

$$\frac{1}{n}\sum_{i=1}^{n}\boldsymbol{x}_i\boldsymbol{x}_i' \xrightarrow{P} \mathbb{E}[\boldsymbol{X}\boldsymbol{X}'],$$

as $n \to \infty$. Multiplying by the constant $\tau(1-\tau)$, where $\tau \in (0,1)$, it follows that

$$\frac{\tau(1-\tau)}{n}\sum_{i=1}^{n}\boldsymbol{x}_i\boldsymbol{x}_i' \xrightarrow{P} \tau(1-\tau)\mathbb{E}[\boldsymbol{X}\boldsymbol{X}'],$$

as $n \to \infty$. By letting $\hat{\boldsymbol{A}}_n = \frac{\tau(1-\tau)}{n}\sum_{i=1}^{n}\boldsymbol{x}_i\boldsymbol{x}_i'$, $\hat{\boldsymbol{A}}_n$ is a consistent estimator of matrix $\boldsymbol{A}$. $\square$

To construct consistent estimators of matrices B, two lemmas about inequality of indicator functions of non-negative real numbers are given.

Lemma 3.4.1.

Let $I_{\{\cdot\}}$ be an indicator function. For any non-negative $a, b, c \in \mathbb{R}$, it follows that

$$|I_{\{a\leq b\}} - I_{\{a\leq c\}}| \leq I_{\{|a-c|\leq|b-c|\}} \tag{3.56}$$

Proof.

We prove by identity.

Case 1: $a \leq b$ and $a \leq c$.

L.H.S. $I_{\{a\leq b\}} = 1$ and $I_{\{a\leq c\}} = 1$, which indicates that $|I_{\{a\leq b\}} - I_{\{a\leq c\}}| = 0 \leq$ R.H.S.;

Case 2: $a > b$ and $a > c$.

L.H.S. $I_{\{a\leq b\}} = 0$ and $I_{\{a\leq c\}} = 0$, which indicates that $|I_{\{a\leq b\}} - I_{\{a\leq c\}}| = 0 \leq$ R.H.S.;

Case 3: $c < a \leq b$.

L.H.S. $I_{\{a\leq b\}} = 1$ and $I_{\{a\leq c\}} = 0$, which indicates that $|I_{\{a\leq b\}} - I_{\{a\leq c\}}| = 1$;

R.H.S. The inequality $c < a \leq b$ is equivalent to $a - c \leq b - c$. Since both $a - c$ and $b - c$ are non-negative, it follows that $|a - c| \leq |b - c|$. Therefore, $I_{\{|a-c| \leq |b-c|\}} = 1$, hence, L.H.S. = R.H.S. in case 3;

Case 4: $b < a \leq c$.

L.H.S. $I_{\{a \leq b\}} = 0$ and $I_{\{a \leq c\}} = 1$, which indicates that $|I_{\{a \leq b\}} - I_{\{a \leq c\}}| = 1$;

R.H.S. The inequality $c < a \leq b$ is equivalent to $a - c > b - c$. Since $a - c$ is either zero or negative, and $b - c$ is negative, it follows that $|a - c| \leq |b - c|$. Thus, $I_{\{|a-c| \leq |b-c|\}} = 1$, hence, L.H.S. = R.H.S. in case 4;

Thus, consider all cases, the inequality of indicator function

$$|I_{\{a \leq b\}} - I_{\{a \leq c\}}| \leq I_{\{|a-c| \leq |b-c|\}}$$

holds. □

Lemma 3.4.2.

Let $I_{\{\cdot\}}$ be an indicator function. For any real number $a_1, a_2, b, c \in [0, +\infty)$, it follows that

$$|I_{\{a_1 \leq b\}} I_{\{a_2 \leq b\}} - I_{\{a_1 \leq c\}} I_{\{a_2 \leq c\}}| \leq I_{\{|a_1-c| \leq |b-c|\}} + I_{\{|a_2-c| \leq |b-c|\}} \tag{3.57}$$

Proof.

Notice that by applying the triangle inequality, we have

$$\begin{aligned}
\text{L.H.S} &= |I_{\{a_1 \leq b\}} I_{\{a_2 \leq b\}} - I_{\{a_1 \leq c\}} I_{\{a_2 \leq c\}}| \\
&= |I_{\{a_1 \leq b\}} I_{\{a_2 \leq b\}} - I_{\{a_1 \leq c\}} I_{\{a_2 \leq b\}} - I_{\{a_1 \leq c\}} I_{\{a_2 \leq c\}} + I_{\{a_1 \leq c\}} I_{\{a_2 \leq b\}}| \\
&\leq |I_{\{a_1 \leq b\}} I_{\{a_2 \leq b\}} - I_{\{a_1 \leq c\}} I_{\{a_2 \leq b\}}| + |I_{\{a_1 \leq c\}} I_{\{a_2 \leq c\}} - I_{\{a_1 \leq c\}} I_{\{a_2 \leq b\}}| \\
&\leq |I_{\{a_1 \leq b\}} \cdot 1 - I_{\{a_1 \leq c\}} \cdot 1| + |1 \cdot I_{\{a_2 \leq c\}} - 1 \cdot I_{\{a_2 \leq c\}}| \\
&\leq I_{\{|a_1-c| \leq |b-c|\}} + I_{\{|a_2-c| \leq |b-c|\}} = \text{R.H.S},
\end{aligned}$$

where the last two inequalities hold by Lemma 3.4.1. $\square$

Theorem 3.4.6 (Consistent Estimator of Matrix $\boldsymbol{B}$)**.**

Let the set $\{y_i, \boldsymbol{x}_i, u_i\}$ be a sample, where $i = 1, 2, ..., n$. Consider the asymptotic normality of $\sqrt{n}(\hat{\boldsymbol{\beta}}_\tau^m - \boldsymbol{\beta}_\tau)$ and the expression of matrix $\boldsymbol{B}$ is given in Theorem 3.4.4.

Let

$$\hat{\boldsymbol{\omega}}_i \equiv \left[\tau - I_{\{y_i + u_i^{(l)} \leq T^{-1}(\boldsymbol{x}_i' \hat{\beta}_\tau)\}}\right] \boldsymbol{x}_i$$

with elements

$$\hat{\omega}_{il} \equiv \left[\tau - I_{\{y_i + u_i^{(l)} \leq T^{-1}(\boldsymbol{x}_i' \hat{\beta}_\tau)\}}\right] x_{ij}$$

and

$$\hat{\boldsymbol{B}}_n \equiv \frac{1}{n} \sum_{i=1}^{n} \hat{\boldsymbol{\omega}}_i \hat{\boldsymbol{\omega}}_i' \tag{3.58}$$

with elements

$$\hat{B}_n(l,k) = \frac{1}{n} \sum_{i=1}^{n} \left[\tau^2 - 2\tau F_U(\hat{Q}_Z(\tau|\boldsymbol{x}_i) - y_i) + F_U(\hat{Q}_Z(\tau|\boldsymbol{x}_i) - y_i)^2\right] x_{il} x_{ik},$$

*where $F_U(\cdot)$ denotes the CDF of additive noise U. Under conditions **A1** to **A9**, $\hat{\boldsymbol{B}}_n$ converges in probability to the matrix $\boldsymbol{B}$, i.e.,*

$$\hat{\boldsymbol{B}}_n \xrightarrow{P} \boldsymbol{B}, \tag{3.59}$$

as $n \to \infty$. Furthermore, given the CDF of U, the specific expression of $\hat{\boldsymbol{B}}_n$ can be alternatively obtained as

1. *If $U \sim \mathcal{T}w_p(\mu, \phi)$, where parameters μ, ϕ, p are known then*

$$\begin{aligned}\hat{\boldsymbol{B}}_n \equiv \hat{\boldsymbol{B}}_n^{Tw} &= \frac{1}{n}\sum_{i=1}^{n} \tau^2 + I_{\{\hat{z}_{\tau_i} - y_i \geq 0\}} \\ &\times \Big[-2\tau \int_0^{\hat{Q}_Z(\tau|\boldsymbol{x}_i)-y_i} a(u;\phi) \exp\Big\{\frac{1}{\phi}\Big(u\frac{\mu^{1-p}}{1-p} - \frac{\mu^{2-p}}{2-p}\Big)\Big\} du \\ &\quad + \Big(\int_0^{\hat{Q}_Z(\tau|\boldsymbol{x}_i)-y_i} a(u;\phi) \exp\Big\{\frac{1}{\phi}\Big(u\frac{\mu^{1-p}}{1-p} - \frac{\mu^{2-p}}{2-p}\Big)\Big\} du \Big)^2 \Big];\end{aligned} \tag{3.60}$$

2. *If $U \sim \mathcal{B}eta(a, b)$, where parameters a, b are known, then*

$$\begin{aligned}\hat{\boldsymbol{B}}_n \equiv \hat{\boldsymbol{B}}_n^{Beta} &= \frac{1}{n}\sum_{i=1}^{n} \tau^2 + (1-2\tau) I_{\{y_i < \hat{z}_{\tau_i} - 1\}} \\ &\quad + I_{\{\hat{Q}_Z(\tau|\boldsymbol{x}_i) - 1 \leq Y_i \leq \hat{Q}_{Z_i}(\tau|x_i)\}} \Big[\Big(\int_0^{\hat{Q}_Z(\tau|\boldsymbol{x}_i)-y_i} \frac{u^{a-1}(1-u)^{b-1}}{B(a,b)} du \Big)^2 \\ &\quad - 2\tau \int_0^{\hat{Q}_Z(\tau|\boldsymbol{x}_i)-y_i} \frac{u^{a-1}(1-u)^{b-1}}{B(a,b)} du \Big].\end{aligned} \tag{3.61}$$

Remark. *Equation 3.60 and 3.61 involve integrating the PDF of Tweedie and Beta variables with respect to U when obtaining the consistent estimator $\hat{\boldsymbol{B}}_n$. The numerical solutions corresponding to integrals can be calculated by the R functions ptweedie and pbeta.*

Proof.

Recall Theorem 3.4.4, the expression of matrix $\boldsymbol{B}$ is given by

$$\boldsymbol{B} \equiv E(\boldsymbol{\omega}_i \boldsymbol{\omega}_i'),$$

where

$$\boldsymbol{\omega}_i = \boldsymbol{\omega}(y_i, u_i^{(l)}, \boldsymbol{x}_i) \equiv \big[\tau - I_{\{y_i + u_i^{(l)} \leq T^{-1}(\boldsymbol{x}_i'\boldsymbol{\beta}_\tau)\}}\big]\boldsymbol{x}_i.$$

Let

$$\widetilde{\boldsymbol{B}}_n \equiv \frac{1}{n}\sum_{i=1}^{n} \boldsymbol{\omega}_i \boldsymbol{\omega}_i'.$$

Since $\frac{1}{n}\sum_{i=1}^{n} \boldsymbol{\omega}_{il}\boldsymbol{\omega}_{ik}' \xrightarrow{P} \boldsymbol{\omega}_i\boldsymbol{\omega}_i'$, by the law of large numbers, we have

$$\widetilde{\boldsymbol{B}}_n \xrightarrow{P} \boldsymbol{B}.$$

Then, it remains to show that $\hat{\boldsymbol{B}}_n - \widetilde{\boldsymbol{B}}_n \xrightarrow{P} \mathbf{0}$. Let

$$\hat{\boldsymbol{\omega}}_{il} \equiv \left[\tau - I_{\{y_i + u_i^{(l)} \le T^{-1}(\boldsymbol{x}_i'\hat{\boldsymbol{\beta}}_\tau)\}}\right]\boldsymbol{x}_i.$$

and

$$\hat{\boldsymbol{B}}_n \equiv \frac{1}{n}\sum_{i=1}^{n} \hat{\boldsymbol{\omega}}_{il}\hat{\boldsymbol{\omega}}_{ik}'.$$

For the sake of simplicity, in the following context, we will denote $T^{-1}(\boldsymbol{x}_i'\boldsymbol{\beta}_\tau)$ and $T^{-1}(\boldsymbol{x}_i'\hat{\boldsymbol{\beta}}_\tau)$ as z_{τ_i} and $\hat{z}_{\tau_i}$, respectively.

Let us turn to the upper bound of $||\hat{\boldsymbol{B}}_n - \tilde{\boldsymbol{B}}_n||$. We firstly rewrite this term as

$$
\begin{aligned}
&\frac{1}{n}\sum_{i=1}^{n} |\hat{\boldsymbol{\omega}}_i\hat{\boldsymbol{\omega}}_i' - \boldsymbol{\omega}_i\boldsymbol{\omega}_i'| \\
&= \frac{1}{n}\sum_{i=1}^{n} ||\boldsymbol{x}_i||^2 \bigg| \tau^2 - \tau I_{\{y_i+u_i^{(l)} \le \hat{z}_{\tau_i}\}} - \tau I_{\{y_i+u_i^{(k)} \le \hat{z}_{\tau_i}\}} + I_{\{y_i+u_i^{(l)} \le \hat{z}_{\tau_i}\}} I_{\{y_i+u_i^{(k)} \le \hat{z}_{\tau_i}\}} \\
&\qquad\qquad - \tau^2 + \tau I_{\{y_i+u_i^{(l)} \le z_{\tau_i}\}} + \tau I_{\{y_i+u_i^{(k)} \le z_{\tau_i}\}} - I_{\{y_i+u_i^{(l)} \le z_{\tau_i}\}} I_{\{y_i+u_i^{(k)} \le z_{\tau_i}\}} \bigg| \\
&= \frac{1}{n}\sum_{i=1}^{n} ||\boldsymbol{x}_i||^2 \bigg| \tau I_{\{y_i+u_i^{(l)} \le z_{\tau_i}\}} - \tau I_{\{y_i+u_i^{(l)} \le \hat{z}_{\tau_i}\}} + \tau I_{\{y_i+u_i^{(k)} \le z_{\tau_i}\}} - \tau I_{\{y_i+u_i^{(k)} \le \hat{z}_{\tau_i}\}} \\
&\qquad + I_{\{y_i+u_i^{(l)} \le \hat{z}_{\tau_i}\}} I_{\{y_i+u_i^{(k)} \le \hat{z}_{\tau_i}\}} - I_{\{y_i+u_i^{(l)} \le z_{\tau_i}\}} I_{\{y_i+u_i^{(k)} \le z_{\tau_i}\}} \bigg| \\
&\le \frac{1}{n}\sum_{i=1}^{n} ||\boldsymbol{x}_i||^2 \\
&\qquad\qquad \times \bigg(\bigg| \tau I_{\{y_i+u_i^{(l)} \le z_{\tau_i}\}} - \tau I_{\{y_i+u_i^{(l)} \le \hat{z}_{\tau_i}\}} \bigg| && (*) \\
&\qquad\qquad\quad + \bigg| \tau I_{\{y_i+u_i^{(k)} \le z_{\tau_i}\}} - \tau I_{\{y_i+u_i^{(k)} \le \hat{z}_{\tau_i}\}} \bigg| && (**) \\
&\qquad\qquad\qquad + |I_{\{y_i+u_i^{(l)} \le \hat{z}_{\tau_i}\}} I_{\{y_i+u_i^{(k)} \le \hat{z}_{\tau_i}\}} - I_{\{y_i+u_i^{(l)} \le z_{\tau_i}\}} I_{\{y_i+u_i^{(k)} \le z_{\tau_i}\}} \bigg| \bigg). && (***)
\end{aligned}
$$

Now, focusing on the last inequality, by Lemma 3.4.1, the first term $(*)$ is bounded above by

$$I_{\{|y_i+u_i^{(l)} - z_{\tau_i}| \le |\hat{z}_{\tau_i} - z_{\tau_i}|\}}.$$

We only have to show the first of these terms are $o_p(1)$, since the proof of second term and third term are analogous. To simplify notation, let $\Delta_i \equiv \hat{z}_{\tau_i} - z_{\tau_i}$. For any $\eta > 0$

and any ε,

$$
\begin{aligned}
&P\Big[\frac{1}{n}\sum_{i=1}^{n}||x_i||^2 I_{\{|y_i+u_i^{(l)}-z_{\tau_i}|\le|\Delta_i|\}} > \eta\Big] \\
&\quad \le P\Big[\frac{1}{n}\sum_{i=1}^{n}||x_i||^2 I_{\{|y_i+u_i^{(l)}-z_{\tau_i}|\le|\varepsilon|\}} > \eta\Big] \\
&\quad\quad + P\Big[\sup_{1\le i\le n}|\Delta_i| > \varepsilon\Big].
\end{aligned}
$$

By the Markov's inequality, and the fact that conditional on $\boldsymbol{X} = \boldsymbol{x}$,

$$
\mathbb{E}\big[I_{\{|y_i+u_i^{(l)}-z_{\tau_i}|\le|\epsilon|\}}\big] \le 2\epsilon,
$$

we can show that $P\big(\frac{1}{n}\sum_{i=1}^{n}||\boldsymbol{x}_i||^2 I_{\{|y_i+u_i^{(l)}-z_{\tau_i}|\le|\epsilon|\}} > \eta\big)$ goes to 0 with ε.

Then, we have to show that $\Delta_i = o_p(1)$ uniformly in i. Note that condition **A8** states that $\mathbb{E}\big[\sup_{||\boldsymbol{\beta}-\boldsymbol{\beta}_\tau||\le\boldsymbol{\delta}}|T^{(1)}(\boldsymbol{x}_i'\boldsymbol{\beta})|^l||\boldsymbol{x}_i||^2\big]$ for $l = 1, 2$ and for some $\delta > 0$, which implies that

$$
\sup_{1\le i\le n}\sup_{||\boldsymbol{\beta}-\boldsymbol{\beta}_\tau||\le\boldsymbol{\delta}}|T^{(1)}(\boldsymbol{x}_i'\boldsymbol{\beta})| \times ||\boldsymbol{x}_i|| = o_p(1).
$$

Thus, by mean-value expansion,

$$
\sup_{1\le i\le n}|\hat{z}_{\tau_i} - z_{\tau_i}| = o_p(1).
$$

Under condition **A9**, this convergence must be faster than c_n; that is, for some sufficiently large n,

$$
P\Big[\sup_{1\le i\le n}|\hat{z}_{\tau_i} - z_{\tau_i}| \le \epsilon c_n\Big] \ge 1-\delta
$$

for any $\delta > 0$ and $\varepsilon > 0$. Hence, $\Delta_i = o_p(1)$. The proof of the second term $(**)$ is analogous to that of $(*)$. Now we turn to the third term $(***)$. Note that by

Lemma 3.4.2, it is bounded above by

$$I_{\{|y_i+u_i^{(l)}-z_{\tau_i}|\le|\hat{z}_{\tau_i}-z_{\tau_i}|\}} + I_{\{|y_i+u_i^{(k)}-z_{\tau_i}|\le|\hat{z}_{\tau_i}-z_{\tau_i}|\}}.$$

The remaining proof to show the sum is $o_p(1)$ is then analogous to that of the first term $(*)$ and the second term $(**)$. Use the same rationale, $I_{\{|y_i+u_i^{(l)}-z_{\tau_i}|\le|\hat{z}_{\tau_i}-z_{\tau_i}|\}} = o_p(1)$ and $I_{\{|y_i+u_i^{(k)}-z_{\tau_i}|\le|\hat{z}_{\tau_i}-z_{\tau_i}|\}} = o_p(1)$ and it immediately follows that the sum of them is $o_p(1)$ as well. Therefore, all the starred terms are $o_p(1)$, and for n sufficiently large,

$$\hat{\boldsymbol{B}}_n - \widetilde{\boldsymbol{B}}_n \xrightarrow{P} \boldsymbol{0}.$$

Recall that $\widetilde{\boldsymbol{B}}_n \xrightarrow{P} \boldsymbol{B}$, by Slutsky's Theorem,

$$\hat{\boldsymbol{B}}_n \xrightarrow{P} \boldsymbol{B},$$

as $n \to \infty$. Thus, $\hat{\boldsymbol{B}}_n$ is a consistent estimator of matrix $\boldsymbol{B}$. The proof of general expression of $\hat{\boldsymbol{B}}_n$ is complete.

The remainder of proof shows the specific expression of consistent estimator $\hat{\boldsymbol{B}}_n$, which is based on the given statistical distributions of noise U. A more specific expression helps compute numerical solution of covariance estimate when simulating the study. By the fact that $\hat{\boldsymbol{B}}_n \equiv \frac{1}{n}\sum_{i=1}^{n} \hat{\boldsymbol{\omega}}_{il}\hat{\boldsymbol{\omega}}'_{ik}$, to simplify the notation, put

$\hat{z}_{\tau_i} \equiv T^{-1}(\boldsymbol{x}_i'\hat{\boldsymbol{\beta}}_\tau) \equiv \hat{Q}_z(\tau|\boldsymbol{x}_i)$. We expand $\frac{1}{n}\sum_{i=1}^n \hat{\boldsymbol{\omega}}_{il}\hat{\boldsymbol{\omega}}_{ik}'$ as

$$
\begin{aligned}
&\frac{1}{n}\sum_{i=1}^n(\tau - I\{y_i + u_i^{(l)} \leq \hat{z}_{\tau_i}\})(\tau - I\{y_i + u_i^{(k)} \leq \hat{z}_{\tau_i}\}) \\
&= \frac{1}{n}\sum_{i=1}^n \big[\tau^2 - \tau I\{y_i + u_i^{(l)} \leq \hat{z}_{\tau_i}\} - \tau I\{y_i + u^{(k)} \leq \hat{z}_{\tau_i}\} \\
&\qquad\qquad + I\{y_i + u_i^{(l)} \leq \hat{z}_{\tau_i}\} I\{y_i + u_i^{(k)} \leq \hat{z}_{\tau_i}\}\big] \\
&= \frac{1}{n}\sum_{i=1}^n \big[\tau^2 - \tau I\{u_i^{(l)} \leq \hat{z}_{\tau_i} - y_i\} - \tau I\{u^{(k)} \leq \hat{z}_{\tau_i} - y_i\} \\
&\qquad\qquad + I\{u_i^{(l)} \leq \hat{z}_{\tau_i} - y_i\} I\{u_i^{(k)} \leq \hat{z}_{\tau_i} - y_i\}\big] \\
&= \frac{1}{n}\sum_{i=1}^n \big[\tau^2 - 2\tau F_U(\hat{z}_{\tau_i} - y_i) + \tau F_U(\hat{z}_{\tau_i} - y_i)^2\big].
\end{aligned}
$$

The last two equalities hold because $\frac{1}{n}\sum_{i=1}^n I\{u_i^{(l)} \leq \hat{z}_{\tau_i} - y_i\}$ converges in probability to $E[I\{u_i^{(l)} \leq \hat{z}_{\tau_i} - y_i\}]$ as $n \to \infty$ by the law of large numbers. Since U follows some specific distributions, $E[I\{u_i^{(l)} \leq \hat{z}_{\tau_i} - y_i\}] \equiv P(U \leq \hat{z}_{\tau_i} - y_i) \equiv F_U(\hat{z}_{\tau_i} - y_i)$. The convergence of the term $\frac{1}{n}\sum_{i=1}^n I\{u_i^{(k)} \leq \hat{z}_{\tau_i} - y_i\}$ is analogous. Substituting the cdf that integrates the pdf of U, we get desired results. □

Finally, we obtain the consistent estimator of matrix $\boldsymbol{D}$. Although the consistent estimator $\hat{\boldsymbol{D}}_n$ is complicated to obtain, J. L. Powell (1984) proposed an approach to estimate $\hat{\boldsymbol{D}}_n$. In particular, the probability density at $Z = z$ equals to the probability that Z is greater or equal to $\lfloor z \rfloor$ and smaller than $\lfloor z+1 \rfloor$, where $\lfloor c \rfloor$ is a floor function that returns the largest integer smaller than or equals to c. Thus, the conditional density function of Z at $Q_Z(\tau|\boldsymbol{x}_i)$ can be written as the conditional expectation of an indicator function

$$
I\{\lfloor Q_Z(\tau|\boldsymbol{x}_i)\rfloor \leq Z_i < \lfloor Q_Z(\tau|\boldsymbol{x}_i) + 1\rfloor. \tag{3.62}
$$

For the monotonic transformation, the density of $T(Z;\tau)$ at $\boldsymbol{x}_i'\boldsymbol{\beta}_\tau$ can be obtained by multiplying it by some Jacobian matrix of the transformation function to the density function of Z at $Q_Z(\tau|\boldsymbol{x}_i)$. Since the floor function is not continuous, Machado and Silva (2005) defined the alternative continuous function to approximate the floor function as follows,

$$F_n(x) = \begin{cases} \lfloor x \rfloor - \frac{1}{2} + \frac{x - \lfloor x \rfloor}{2c_n} \text{ if } x - \lfloor x \rfloor < c_n \text{ and } x \geq 1 \\ \lfloor x \rfloor \text{ if } c_n \leq x - \lfloor x \rfloor < 1 - c_n \text{ or } x < 1 \\ \lfloor x \rfloor + \frac{1}{2} + \frac{x - \lfloor x \rfloor - 1}{2c_n} \text{ if } x - \lfloor x \rfloor \geq 1 - c_n. \end{cases}$$

By Theorem 3.4.2, we can see that the distribution of U is orthogonal to the matrix $\boldsymbol{D}$, which actually leads to an identical estimator $\hat{\boldsymbol{D}}_n$ given by Machado and Silva (2005).

Theorem 3.4.7 (Consistent Estimator of Matrix $\boldsymbol{D}$, Machado and Silva, 2005)**.** *Let*

$$\hat{\boldsymbol{D}}_n \equiv \frac{1}{n} \sum_{i=1}^{n} \hat{w}_i \boldsymbol{x_i} \boldsymbol{x_i'},$$

with

$$\hat{w}_i \equiv T^{(1)}(\boldsymbol{x_i'}\hat{\boldsymbol{\beta}}_{\boldsymbol{\tau}} I(F_n(\hat{Q}_{Z_i}(\tau|\boldsymbol{x})) \leq Z_i < F_n(\hat{Q}_{Z_i}(\tau|\boldsymbol{x}) + 1)),$$

*where $\hat{Q}_{Z_i}(\tau|\boldsymbol{x}) \equiv T^{-1}(\boldsymbol{x_i'}\boldsymbol{\beta_\tau})$ is the estimated τ^{th} conditional quantile of Z, and continuous function $F_n(\cdot)$ is given above. Under all conditions **A1** to **A9**,*

$$\hat{\boldsymbol{D}}_n \xrightarrow{P} \boldsymbol{D},$$

as $n \to \infty$. Hence, $\hat{\boldsymbol{D}}_n$ is a consistent estimator of matrix $\boldsymbol{D}$.

Proof.

Let

$$\widetilde{\boldsymbol{D}}_n = \frac{1}{n} \sum_{i=1}^{n} \boldsymbol{\omega}_i \boldsymbol{x}_i \boldsymbol{x}_i',$$

with

$$\boldsymbol{\omega}_i = T^{(1)}(\boldsymbol{x}_i'\boldsymbol{\beta}_\tau)\left\{F_n(z_\tau) \le Z_i < F_n(z_\tau+1)\right\},$$

where z_τ is used as shorthand for $T^{-1}(\boldsymbol{x}_i'\boldsymbol{\beta}_\tau) = Q_Z(\tau|x_i)$. Notice that as n passes to ∞, $F_n(x) \to \lfloor x \rfloor$. Thus, by dominated convergence, as $n \to \infty$,

$$\mathbb{E}[\boldsymbol{\omega}_i\boldsymbol{x}_i\boldsymbol{x}_i'] \to E\big[T^{-1}(\boldsymbol{x}_i'\boldsymbol{\beta}_\tau))\left\{\lfloor z_\tau \rfloor \le Z_i < \lfloor z_\tau+1 \rfloor\right\}\boldsymbol{x}_i\boldsymbol{x}_i'\big],$$

with a simple change of variable, it equals $\mathbb{E}\big[f_T(\boldsymbol{x}_i'\boldsymbol{\beta}_\tau)\boldsymbol{x}_i\boldsymbol{x}_i'\big]$. The law of large numbers then yields

$$\widetilde{\boldsymbol{D}}_n \xrightarrow{P} \boldsymbol{D}.$$

The remainder of proof shows $\hat{\boldsymbol{D}}_n - \widetilde{\boldsymbol{D}}_n \xrightarrow{P} \mathbf{0}$ so that $\hat{\boldsymbol{D}}_n \xrightarrow{P} \boldsymbol{D}$, see Machado and Silva (2005). □

Before we end this chapter, the sandwich estimator of variance-covariance matrices of simple jittering estimators and average jittering estimators are given. They are practically useful in statistical inference on jittering QR model with finite sample, where variation of sample estimate can be estimated more easily.

Corollary 3.4.3.

By the Slutsky's theorem, the sandwich estimator of variance-covariance matrix of $\widehat{Cov}(\sqrt{n}(\hat{\boldsymbol{\beta}}_\tau - \boldsymbol{\beta}_\tau))$ *is*

$$\widehat{Cov}(\sqrt{n}(\hat{\boldsymbol{\beta}}_\tau - \boldsymbol{\beta}_\tau)) = \hat{\boldsymbol{D}}^{-1}\hat{\boldsymbol{A}}\hat{\boldsymbol{D}}^{-1}, \tag{3.63}$$

and the nature sandwich form estimator of variance-covariance matrix of $\widehat{Cov}(\sqrt{n}(\hat{\boldsymbol{\beta}}_\tau^m - \boldsymbol{\beta}_\tau))$ *is*

$$\widehat{Cov}(\sqrt{n}(\hat{\boldsymbol{\beta}}_\tau^m - \boldsymbol{\beta}_\tau)) = \frac{1}{m}\hat{\boldsymbol{D}}^{-1}\hat{\boldsymbol{A}}\hat{\boldsymbol{D}}^{-1} + \Big(1 - \frac{1}{m}\Big)\hat{\boldsymbol{D}}^{-1}\hat{\boldsymbol{B}}\hat{\boldsymbol{D}}^{-1}, \tag{3.64}$$

where m is the number of jittered samples.

Corollary 3.4.4.

The sandwich estimator of variance-covariance matrices for $\hat{\boldsymbol{\beta}}_\tau$ and $\hat{\boldsymbol{\beta}}_\tau^m$ are

$$\widehat{Cov}(\hat{\boldsymbol{\beta}}_\tau) = \frac{1}{n}\hat{\boldsymbol{D}}^{-1}\hat{\boldsymbol{A}}\hat{\boldsymbol{D}}^{-1}, \tag{3.65}$$

and

$$\widehat{Cov}(\hat{\boldsymbol{\beta}}_\tau^m) = \frac{1}{n}\left(\frac{1}{m}\hat{\boldsymbol{D}}^{-1}\hat{\boldsymbol{A}}\hat{\boldsymbol{D}}^{-1} + \left(1 - \frac{1}{m}\right)\hat{\boldsymbol{D}}^{-1}\hat{\boldsymbol{B}}\hat{\boldsymbol{D}}^{-1}\right), \tag{3.66}$$

respectively.

Chapter 4

Numerical Study

In order to evaluate the small sample performance of our QR method, we conduct two simulation studies: counts from discrete distributions, and that from truncated continuous distributions.

4.1 True QR Coefficients and Estimates

Due to non-linearity of the quantiles, the true coefficients of QR model may vary with quantiles, and hence may not be the value we set. In practice, true value can be obtained from a sufficiently large sample with size N where we call it pseudo population, denoted by $\boldsymbol{\beta}_\tau$, which is asymptotically equivalent to true parameter according to Theorem 3.4.2. Therefore, let $(\boldsymbol{Y}^{\text{pop}}_{[N]}, \boldsymbol{X}^{\text{pop}}_{[N\times p]})$ be a generated population under the conditions and consist of QR response and matrix of predicts. The mechanism of defining true QR coefficients and obtaining QR estimates is summarized as follows,

1. True QR coefficients $\boldsymbol{\beta}_\tau$: Generate population with large size N. True values

$\boldsymbol{\beta}_\tau$ are obtained using corresponding QR method,

$$\boldsymbol{Y}^{\text{pop}} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{pmatrix}, \quad \boldsymbol{X}^{\text{pop}} = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ x_{N1} & x_{N2} & \cdots & x_{Np} \end{pmatrix} \xRightarrow{\text{QR method}} \boldsymbol{\beta}_\tau. \tag{4.1}$$

2. QR estimates $\hat{\boldsymbol{\beta}}_\tau^{(k)}$ in the k-th simulation run: Let $(y_i^{(k)}, \boldsymbol{x}_i^{(k)})_{i=1}^n$ be a size n random sample selected from the generated population in the $k = 1, ..., S$-th simulation run. The estimates $\hat{\boldsymbol{\beta}}_\tau^{(k)}$ are obtained from that sample using corresponding QR method,

$$\boldsymbol{y} = \begin{pmatrix} y_1^{(k)} \\ y_2^{(k)} \\ \vdots \\ y_n^{(k)} \end{pmatrix}, \quad \boldsymbol{x} = \begin{pmatrix} x_{11}^{(k)} & x_{12}^{(k)} & \cdots & x_{1p}^{(k)} \\ x_{21}^{(k)} & x_{22}^{(k)} & \cdots & x_{2p}^{(k)} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1}^{(k)} & x_{n2}^{(k)} & \cdots & x_{np}^{(k)} \end{pmatrix} \xRightarrow{\text{QR method}} \hat{\boldsymbol{\beta}}_\tau^{(k)}. \tag{4.2}$$

To carry out simulations, we set the number of replications used in both simulations as $S = 1000$.

Assessments of proposed QR method are based on bias, mean squared error (MSE), asymptotic standard error (SE) and sample standard deviation (SD). The rejection rate (RR) and power $(1 - \beta)$ are for assessment of hypothesis tests. Notice that the relative efficiencies (REF) of estimators can be calculated with MSEs directly, which are not included in the tables.

4.2 Simulation Study I: QR with Discrete Distributed Data

In the 1st simulation we assess the performance of QR method when modeling the counts from four discrete distributions: Poisson, negative binomial (NB), zero-inflated Poisson (ZIP), and zero-inflated negative binomial (ZINB) distributions. We simulated the proposed QR method on Poisson and negative binomial data to account for different variance structures.

4.2.1 Data and Simulation Setup

Pseudo population $(y_i, \boldsymbol{x}_i)_{i=1}^N$ is generated from the aforementioned discrete distributions conditional on predictors. Sample with size n is randomly selected from the population. Data generation for Poisson model and negative binomial model are straightforward. We discuss a little details about sampling from zero-inflated Poisson model and zero-inflated negative binomial model. Erdman and Sinko (2008) introduced that for each observation i, two data generation processes run simultaneously, where the result of a Bernoulli trial determines the process. Process 1 is selected with probability $\varphi_i \in [0, 1]$ and process 2 with probability $1-\varphi_i$, where φ_i is called proportion of zero inflation. Thus, process 1 generates only zeros, and process 2 generates discrete variables from $g(y_i|\boldsymbol{x}_i)$, where $g(y_i|\boldsymbol{x}_i), i = 1, 2, ...,$ denotes either Poisson or negative binomial model. Then, we generate the i^{th} observation

$$y_i \sim \begin{cases} 0 & \text{with probability } \varphi_i \\ g(y_i|\boldsymbol{x}_i) & \text{with probability } 1 - \varphi_i. \end{cases} \tag{4.3}$$

For fixed proportion of zero inflation, let $\varphi_i = \varphi$. Also, by letting $\varphi = 0$, the zero-inflated distribution becomes its regular version (i.e., non zero-inflation).

To compare the simulation results clearly, we generate count variable Y according to the same non-linear conditional mean

$$\mu_i = \exp(\beta_0 + x_{i1}\beta_1 + x_{i2}\beta_2), \tag{4.4}$$

where $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2)'$ is vector of regression coefficients. The first predictor is initially sampled from beta distribution,

$$x_{i1}^* \sim \text{Beta}(5/3, 5/3),$$

and then centered and scaled, i.e., $x_{i1} = (x_{i1}^* - \bar{x}_{i1}^*)/s.d.(x_1^*)$. The second predictor is a dummy variable,

$$x_{i2} = \begin{cases} 1 & \text{with probability } 0.5 \\ 0 & \text{with probability } 0.5. \end{cases}$$

Four experiments for QR response Y following distributions below are simulated:

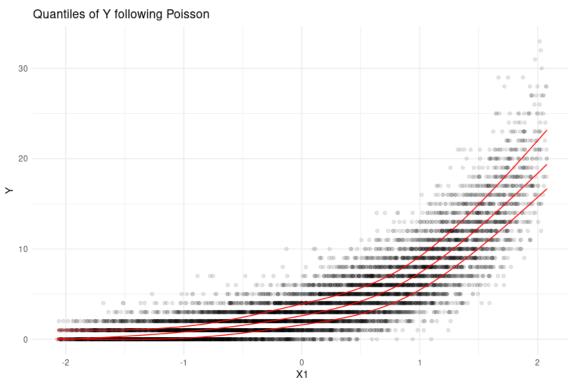
***Case* 1.** Poisson distribution with rate $\lambda_i = \mu_i$;

***Case* 2.** Zero-inflated Poisson distribution with rate $\lambda_i = \mu_i$, and the proportion of zero inflation of 0.2;

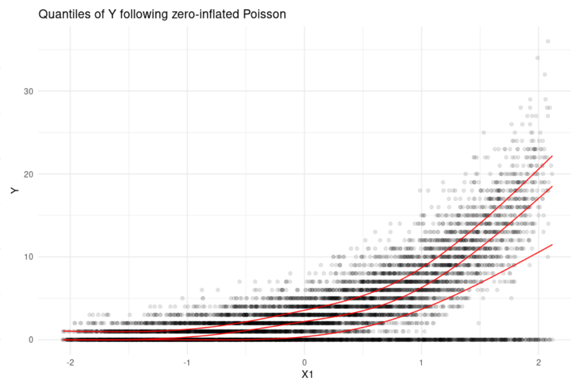
***Case* 3.** Negative binomial distribution with mean μ_i and variance $\sigma_i^2 = \mu_i + 0.5\mu_i^2$.

***Case* 4.** Zero-inflated negative binomial distribution with mean μ_i, variance $\sigma_i^2 = \mu_i + 0.5\mu_i^2$, and the proportion of zero inflation of 0.2.

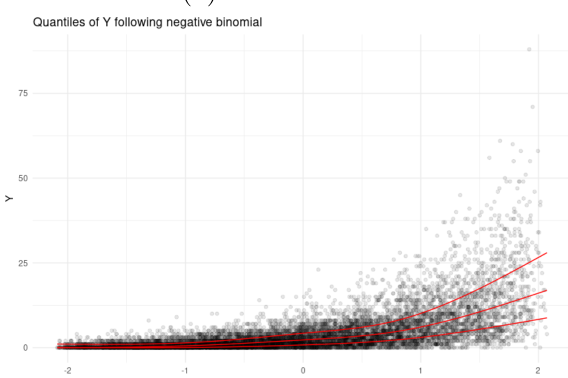
Figure 4.1 provides an overview of count data following four different discrete distributions. The red curves indicate quantiles of distribution on $\tau = (0.25, 0.5, 0.75)$. As we can see, all counts are exponentially increasing with $\boldsymbol{x}_1$, especially for those following negative binomial and zero-inflated negative binomial. Over-dispersion is displayed as well (see Fig. 4.1c and Fig. 4.1d). Also, a number of zero values are



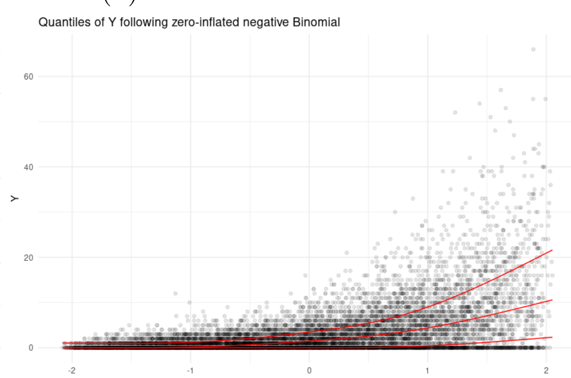
(a) *Case* 1. Pois.



(b) *Case* 2. Zero-inflated Pois.



(c) *Case* 3. N.binomial



(d) *Case* 4. Zero-inflated n.binomial

Figure 4.1: Count Y with respects to $\boldsymbol{x}_1$ generated from four discrete distributions. Red lines indicate quantiles of distributions on $\tau = (0.25, 0.5, 0.75)$

clearly shown at the bottom of graphs of zero-inflated distributions (see Fig. 4.1b and Fig. 4.1d), which affect the quantiles of distribution on lower levels, i.e., $\tau = 0.25$.

In the simulation, we set $\beta_0 = \beta_1 = 1$ and $\beta_2 = 0$. The zero effect allows us to investigate the finite-sample behaviour of hypothesis tests based on simple jittering estimate $\hat{\boldsymbol{\beta}}_\tau$ and average jittering estimate $\hat{\boldsymbol{\beta}}_\tau^A$. Also, their asymptotic covariance matrices can be estimated. In this simulation, we set a large population with size of $N = 50000$ and sample with size of $n = (200, 500)$, and $\tau = (0.25, 0.5, 0.75)$.

To estimate the asymptotic covariance matrix, a sequence c_n should be decided. It can be shown that predictors $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ are $O_p(1)$. We choose the sequence $c_n = 0.5\ln(\ln(n))/\sqrt{n}$ to satisfy the condition **A9** since its limit

$$\begin{aligned}
&\lim_{n\to\infty} \frac{0.5\ln(\ln(n))/\sqrt{n}}{1/\sqrt{n}} \\
&= \lim_{n\to\infty} 0.5\ln(\ln(n))
\end{aligned} \tag{4.5}$$

increases without bound as $n \to \infty$ but very slowly due to double logarithmic functions. It is also worth-noting that the choice of sequence c_n is not unique as long as the condition is satisfied, but restricted by the behaviours of predictors and corresponding transformation if any.

For our optimized jittering methods, distribution parameters are obtained by using the MCEM algorithm introduced in Section 3.3. Constant MC sample size of 5000 and burn-in size of 2500 are used. To ensure the independence of samples from MC-step in each iteration, thinning with $h = 5$ is preferred.

4.2.2 QR Methods and Transformations

The performance of jittering approaches and ordinary QR method are compared. Also, The performance of simple jittering methods and average jittering methods are compared. To see the performance of average jittering technique, jittered sample size of $m = (50, 150)$ are used. Therefore, QR methods and their acronyms used in simulation are as follows:

1. Ordinary quantile regression method (ORD);

2. Simple uniform-jittering quantile regression method (UJ);

3. Average uniform-jittering quantile regression method with jittered samples $m = 50$ (AUJ50);

4. Simple Tweedie-jittering quantile regression method (TJ);

5. Average Tweedie-jittering quantile regression method with jittered samples $m = 50$ (ATJ50);

6. Average Tweedie-jittering quantile regression method with jittered samples $m = 150$ (ATJ150);

7. Simple Beta-jittering quantile regression method (BJ);

8. Average beta-jittering quantile regression method with jittered samples $m = 50$ (ABJ50); and

9. Average beta-jittering quantile regression method with jittered samples $m = 150$ (ABJ150).

Since the data are exponential, different transformations of response are applied for all regression methods. For ordinary quantile regression method, we simply apply the logarithmic transformation of counts. Since no additional noises are added, the transformation is unconditional to quantile level τ.

$$T(Y;\tau) = \begin{cases} \log(Y) & \text{if} \quad Y > 0 \\ \log(\varsigma) & \text{if} \quad Y \leq 0, \end{cases} \tag{4.6}$$

where ς is small positive number and τ is the quantile. The QR model is specified as

$$Q_{T(Y;\tau)}(\tau|\boldsymbol{x}) = \boldsymbol{x}'\boldsymbol{\beta}_{\tau}, \tag{4.7}$$

where $\boldsymbol{\beta}_{\tau}$ is the vector of QR coefficients. For jittering and average-jittering methods, logarithmic transformation is applied to the jittered response, $Z = Y + U$, that is,

$$T(Z;\tau) = \begin{cases} \log(Z - Q_U(\tau)) & \text{if} \quad Z > Q_U(\tau) \\ \log(\varsigma) & \text{if} \quad Z \leq Q_U(\tau), \end{cases} \tag{4.8}$$

and the corresponding QR model is

$$Q_{T(Z;\tau)}(\tau|\boldsymbol{x}) = \boldsymbol{x}'\boldsymbol{\beta}_{\tau}, \tag{4.9}$$

where U is additive noise following unif$[0, 1)$ for uniform jittering method, $\text{Tw}_p(\mu, \phi)$ for Tweedie jittering method, and $\text{Beta}(a, b)$ for beta jittering methods.

4.2.3 Comparison of Different QR Methods

In this section $S = 1000$ simulation runs of QR for counts using different methods will be analyzed. We report average bias ($bias$), mean squared error (MSE) and relative efficiency (REF) of the estimates of $\beta_{\tau 0}$, $\beta_{\tau 1}$ and $\beta_{\tau 2}$ of QR model by using seven methods mentioned in Section 4.2.2. Also, results from the simulations of four discrete distributions (Poisson, negative binomial, zero-inflated Poisson and zero-inflated negative binomial) are reported in following Tables.

Table 4.1 and Table 4.1A report the biases and MSEs of the estimators using different methods when counts follow Poisson, negative binomial and corresponding zero-inflated distributions, respectively with samples size $n = (200, 500)$. Ordinary

QR method gives large biases and MSEs and low relative efficiencies of estimators especially on low quantiles, implying a serious systematic bias due to discreteness. Also, ordinary method is unable to estimate the QR parameters for both zero-inflated Poisson and zero-inflated negative binomial distribution on $\tau = 0.25$ due to massive zeros.

Estimators from both proposed jittering methods are approximately $10\% \sim 50\%$ more efficient than that of uniform-jittering methods depending on the distributions and quantiles. This is as expected because the proposed method aims to optimize the model fit so that the performance of the estimator obtained by the proposed jittering methods is improved. Also, it is interesting to see that the biases and MSEs of estimators using our optimized jittering method are small at low quantile of zero-inflation.

Now, let us compare Tweedie and Beta jittering methods and their average jittering versions. The biases of estimators between regular jittering methods and average jittering methods are slightly different, whereas the average jittering methods can provide extra $10\% \sim 80\%$ higher efficiency of estimator across four combinations of distribution and quantiles. This outcome is similar to Machado's study (2005) whose efficiency of estimator is higher when averaging out noises. However, increasing the value of m does not always lead to a significant improvement of efficiency. For example, when we increase m from 50 to 150, the biases and MSEs of estimators are only slightly improved by $1\% \sim 5\%$ roughly. Therefore, recommended number of jittered sample is $m = 50$. In short, average jittering methods are suggested to use in order to obtain more efficient estimates. The Tweedie and Beta jittering methods, as well as their average jittering version, exhibit similar performance. Beta jittering method provides slightly higher estimator efficiency on average. But importantly, the algorithm runtime of Beta jittering method are always shorter than that of Tweedie jittering method. Therefore, Beta jittering method is much more flexible to use in practice.

Table 4.1A in Appendix shows the biases, MSEs and REFs of estimators when Y follows Poisson, negative binomial and corresponding zero-inflated versions, respectively, with sample size $n = 500$. The performance of different QR methods are similar to those observed from Table 4.1 but with some differences. The increase in sample size significantly reduces the biases and mean squared errors of estimators observed from all jittering methods but maintains the relative efficiencies.

In a nutshell, Tweedie and Beta jittering methods grant QR the ability to model counts, with better performance than uniform jittering methods. Applying the technique of averaging-noise out and increasing sample size improve the efficiency and reduce bias of estimator.

Table 4.1: Bias ($bias$) and mean squared errors (MSE) of the estimators of β_0, β_1, and β_2 using different methods (OQR, TJQR, ATJQR, BJQR, ABJQR, UJQR, AUJQR) at three quantiles (0.25, 0.5, 0.75). Sample size $n = 200$.

		$\beta_{\tau 0}$		$\beta_{\tau 1}$		$\beta_{\tau 2}$	
τ	Method	Bias	MSE	Bias	MSE	Bias	MSE
Case 1. *Poisson*							
0.25	ORD	0.6291	2.8557	-0.4424	1.4117	-0.0062	0.4466
	UJ	0.0293	0.0177	-0.0048	0.0098	-0.0306	0.0174
	AUJ	0.0226	0.0119	-0.0052	0.0055	-0.0334	0.0161
	TJ	0.0213	0.0191	-0.0026	0.0099	-0.0369	0.0210
	ATJ50	0.0082	0.0113	0.0010	0.0052	-0.0256	0.0152
	ATJ150	-0.0091	0.0078	0.0031	0.0030	-0.0011	0.0109
	BJ	0.0185	0.0185	-0.0033	0.0101	-0.0267	0.0198
	ABJ50	0.0112	0.0107	-0.0083	0.0048	-0.0237	0.0146
	ABJ150	-0.0067	0.0079	0.0031	0.0031	0.0033	0.0121

Continued on next page

		$\beta_{\tau 0}$		$\beta_{\tau 1}$		$\beta_{\tau 2}$	
τ	Method	Bias	MSE	Bias	MSE	Bias	MSE
0.5	ORD	-0.0262	0.0193	0.0182	0.0121	0.0166	0.0176
	UJ	0.0127	0.0088	0.0027	0.0058	-0.0152	0.0105
	AUJ	-0.0137	0.0071	0.0097	0.0042	0.0092	0.0083
	TJ	0.0018	0.0082	0.0022	0.0054	-0.0130	0.0098
	ATJ50	-0.0072	0.0061	0.0001	0.0029	0.0028	0.0081
	ATJ150	-0.0036	0.0058	0.0142	0.0029	-0.0071	0.0069
	BJ	0.0024	0.0084	0.0031	0.0049	-0.0147	0.0094
	ABJ50	0.0027	0.0065	0.0088	0.0039	-0.0210	0.0087
	ABJ150	-0.0051	0.0063	0.0109	0.0039	-0.0032	0.0081
0.75	ORD	-0.0050	0.0075	0.0050	0.0038	-0.0039	0.0120
	UJ	-0.0196	0.0103	0.0057	0.0061	0.0165	0.0131
	AUJ	-0.0194	0.0089	0.0093	0.0051	0.0145	0.0116
	TJ	-0.0168	0.0069	0.0064	0.0038	0.0074	0.0085
	ATJ50	-0.0209	0.0055	0.0055	0.0028	0.0175	0.0069
	ATJ150	-0.0066	0.0054	0.0042	0.0027	0.0000	0.0068
	BJ	-0.0237	0.0064	0.0055	0.0039	0.0177	0.0085
	ABJ50	-0.0181	0.0054	0.0052	0.0051	0.0147	0.0118
	ABJ150	-0.0158	0.0052	0.0072	0.0030	0.0062	0.0071
Case 2. *Zero-inflated Poisson*							
0.25	ORD	-	-	-	-	-	-
	UJ	-0.1541	0.1699	-0.1316	0.1492	-0.0400	0.2872
	AUJ	-0.0876	0.1047	-0.1211	0.0949	-0.0549	0.1598
	TJ	-0.1386	0.1641	-0.1143	0.1285	-0.0334	0.2301

Continued on next page

Table 4.1 – continued from previous page

τ	Method	$\beta_{\tau 0}$ Bias	$\beta_{\tau 0}$ MSE	$\beta_{\tau 1}$ Bias	$\beta_{\tau 1}$ MSE	$\beta_{\tau 2}$ Bias	$\beta_{\tau 2}$ MSE
	ATJ50	-0.0519	0.1141	-0.1206	0.0894	-0.0562	0.1549
	ATJ150	-0.0643	0.1120	-0.1261	0.0893	-0.1156	0.1539
	BJ	-0.0956	0.1592	-0.0900	0.1247	-0.0834	0.2551
	ABJ50	-0.0544	0.1053	-0.1199	0.0921	-0.0694	0.1580
	ABJ150	-0.0424	0.1020	-0.1040	0.0904	-0.0306	0.1541
0.5	ORD	-0.1346	0.3638	0.1346	0.2808	-0.0137	0.0742
	UJ	0.0027	0.0210	0.0013	0.0105	-0.0080	0.0254
	AUJ	-0.0047	0.0163	-0.0026	0.0086	0.0089	0.0205
	TJ	-0.0039	0.0197	-0.0037	0.0094	0.0149	0.0246
	ATJ50	-0.0080	0.0151	-0.0007	0.0074	-0.0076	0.0189
	ATJ150	-0.0198	0.0150	0.0007	0.0063	0.0056	0.0179
	BJ	-0.0078	0.0191	-0.0045	0.0099	0.0070	0.0204
	ABJ50	0.0113	0.0147	0.0078	0.0074	-0.0429	0.0203
	ABJ150	-0.0065	0.0136	-0.0016	0.0072	0.0303	0.0183
0.75	ORD	0.0178	0.0195	-0.0214	0.0118	0.0178	0.0240
	UJ	-0.0262	0.0110	0.0140	0.0061	0.0179	0.0134
	AUJ	-0.0226	0.0093	0.0101	0.0052	0.0206	0.0123
	TJ	-0.0191	0.0093	0.0206	0.0054	0.0033	0.0132
	ATJ50	-0.0061	0.0078	0.0121	0.0042	-0.0066	0.0117
	ATJ150	-0.0060	0.0075	-0.0099	0.0035	-0.0051	0.0111
	BJ	-0.0411	0.0087	0.0302	0.0049	0.0213	0.0114
	ABJ50	-0.0274	0.0094	0.0183	0.0054	0.0045	0.0111
	ABJ150	-0.0274	0.0079	0.0111	0.0040	0.0166	0.0103

Continued on next page

Table 4.1 – continued from previous page

		$\beta_{\tau 0}$		$\beta_{\tau 1}$		$\beta_{\tau 2}$	
τ	Method	Bias	MSE	Bias	MSE	Bias	MSE
Case 3. *Negative binomial*							
0.25	ORD	0.0556	0.4045	-0.0444	0.0455	-0.0141	0.6098
	UJ	-0.0236	0.0542	0.0215	0.0335	-0.0477	0.0782
	AUJ	0.0289	0.0254	0.0183	0.0128	-0.0801	0.0545
	TJ	0.0281	0.0351	-0.0126	0.0199	-0.0516	0.0650
	ATJ50	0.0275	0.0245	-0.0052	0.0101	-0.0523	0.0473
	ATJ150	0.0116	0.0332	0.0011	0.0131	-0.0053	0.0599
	BJ	0.0327	0.0331	-0.0061	0.0191	-0.0527	0.0604
	ABJ50	-0.0057	0.0215	0.0036	0.0105	-0.0260	0.0436
	ABJ150	0.0206	0.0213	-0.0091	0.0103	-0.0387	0.0324
0.5	ORD	-0.0928	0.1605	0.1094	0.1604	-0.0418	0.0800
	UJ	-0.0061	0.0233	-0.0071	0.0170	0.0019	0.0367
	AUJ	-0.0060	0.0191	0.0126	0.0118	0.0016	0.0350
	TJ	0.0061	0.0204	0.0145	0.0109	-0.0404	0.0347
	ATJ50	0.0014	0.0152	0.0124	0.0092	-0.0348	0.0309
	ATJ150	0.0012	0.0145	0.0119	0.0087	0.0255	0.0283
	BJ	0.0036	0.0201	0.0118	0.0121	-0.0314	0.0398
	ABJ50	0.0098	0.0164	0.0112	0.0096	-0.0252	0.0288
	ABJ150	0.0054	0.0160	0.0115	0.0086	-0.0151	0.0256
0.75	ORD	-0.0345	0.0276	-0.0246	0.0175	0.0559	0.0527
	UJ	-0.0465	0.0188	0.0084	0.0110	-0.0018	0.0299
	AUJ	-0.0539	0.0162	0.0046	0.0082	0.0272	0.0291
	TJ	-0.0197	0.0150	0.0089	0.0097	-0.0033	0.0255

Continued on next page

		$\beta_{\tau 0}$		$\beta_{\tau 1}$		$\beta_{\tau 2}$	
τ	Method	Bias	MSE	Bias	MSE	Bias	MSE
	ATJ50	-0.0143	0.0147	0.0072	0.0087	-0.0029	0.0246
	ATJ150	-0.0084	0.0124	-0.0068	0.0067	-0.0126	0.0213
	BJ	-0.0089	0.0138	0.0031	0.0090	-0.0135	0.0267
	ABJ50	-0.0175	0.0151	0.0141	0.0071	-0.0190	0.0269
	ABJ150	-0.0271	0.0141	0.0144	0.0070	-0.0068	0.0257
Case 4. *Zero-inflated negative binomial*							
0.25	ORD	-	-	-	-	-	-
	UJ	-0.0503	0.1441	0.0010	0.1157	0.0049	0.2509
	AUJ	-0.0403	0.0841	0.0078	0.0822	-0.031	0.1539
	TJ	0.0122	0.0939	0.0410	0.0955	-0.0615	0.1758
	ATJ50	0.0259	0.0617	0.0645	0.0643	-0.0367	0.1064
	ATJ150	-0.0211	0.0585	0.0435	0.0591	-0.0299	0.0923
	BJ	0.0278	0.0914	0.0393	0.0750	-0.0591	0.1846
	ABJ50	0.0331	0.0626	0.0296	0.0635	-0.0382	0.1243
	ABJ150	-0.0121	0.0610	0.0156	0.0617	0.0000	0.1121
0.5	ORD	-0.5316	2.5768	0.3500	1.3772	0.0144	0.8754
	UJ	0.0223	0.0559	0.0061	0.0314	-0.0531	0.0966
	AUJ	-0.0093	0.0381	0.0047	0.0175	-0.0211	0.0765
	TJ	0.0315	0.0366	-0.0096	0.0179	-0.0178	0.0686
	ATJ50	0.0213	0.0270	-0.0190	0.0126	-0.0030	0.0535
	ATJ150	-0.0129	0.0268	0.0122	0.0115	-0.0022	0.0516
	BJ	0.0179	0.0343	-0.0050	0.0181	-0.0256	0.0694
	ABJ50	0.0119	0.0321	-0.0020	0.0116	-0.0284	0.0592

Continued on next page

Table 4.1 – continued from previous page

τ	Method	$\beta_{\tau 0}$ Bias	$\beta_{\tau 0}$ MSE	$\beta_{\tau 1}$ Bias	$\beta_{\tau 1}$ MSE	$\beta_{\tau 2}$ Bias	$\beta_{\tau 2}$ MSE
	ABJ150	0.0111	0.0310	-0.0018	0.0111	0.0258	0.0551
0.75	ORD	-0.0472	0.0289	0.0202	0.0237	0.0208	0.0396
	UJ	-0.0700	0.0302	0.0264	0.0180	0.0268	0.0460
	AUJ	-0.0495	0.0272	0.0318	0.0189	-0.0117	0.0439
	TJ	-0.0128	0.0236	-0.0048	0.0196	0.0156	0.0444
	ATJ50	-0.0124	0.0223	0.0050	0.0119	0.0045	0.0380
	ATJ150	-0.0217	0.0195	0.0138	0.0165	0.0099	0.0419
	BJ	-0.0146	0.0221	0.0047	0.0134	0.0045	0.0403
	ABJ50	-0.0217	0.0209	-0.0062	0.0130	0.0020	0.0363
	ABJ150	-0.0156	0.0203	0.0066	0.0113	-0.0097	0.0361

4.2.4 The Estimation of Variances

In this section, we evaluate the variance estimation. The sample standard deviations of optimized jittering estimator $\hat{\beta}_\tau$ and average-jittering estimator $\hat{\beta}_\tau^{(50)}$ and corresponding asymptotic standard errors are reported in Table 4.2 and 4.2A in Appendix, respectively for sample sizes of $n = 500$ and $n = 200$. We focus on the results observed from the Table where $n = 500$ is used, since the asymptotic properties are much more clearly evaluated with larger sample size. The coverage probabilities, the percentages of simulation runs when the true QR coefficients falls into $(1-\alpha) \times 100\%$ Wald confidence intervals constructed based on sandwich estimator of asymptotic covariance matrix using formulas in Corollary 3.65 and 3.66, are recorded. Table 4.2 reports the comparison of coverage probabilities of estimators, denoted by $P_{1-\alpha}$, to nominal level of confidence $1-\alpha$, respectively according to the sample sizes $n = 500$ and $n = 200$. We set levels of confidence with $\alpha = (0.01, 0.05, 0.1)$.

Table 4.2 and Table 4.2A report the comparison of standard deviations of 1000 estimates of the proposed quantile regression models and their corresponding estimated standard errors when counts follow different distributions, respectively based on sample sizes $n = (500, 200)$. Across four discrete distributions, the difference between most standard deviations and estimated SEs are small and acceptable. This shows that the standard errors obtained by estimated variance-covariance matrix are reliable to construct confidence intervals and conduct parametric hypothesis tests. However, on lower quantile of counts when following zero-inflated negative binomial, we observe that the estimated standard errors are moderately larger than corresponding sample standard deviations of estimator. Consider the percentage of excessive zeros, the result is not surprising. On the other hand, when smaller sample size $n = 200$ is used, the estimated standard errors of estimator using all methods are much larger than corresponding sample standard deviations. See Table 4.3, the coverage probabilities of $\beta_{\tau 2}$ are much higher than their due nominal levels at lower quantile of zero-inflated negative binomial (case 4). For other distributions, most values of $P_{0.99}$, $P_{0.95}$ and $P_{0.90}$ are close to the corresponding nominal levels.

In conclusion, the proposed estimators are asymptotically normally distributed and inferences are reliable. Compare to results observed in Table 4.3A, the coverage probability calculated based on larger sample size $n = 500$ is much more accurate than sample size $n = 200$ is used.

Table 4.2: Standard deviations (s.d.) of 1000 estimates of QR parameters $\beta_{\tau 0}$, $\beta_{\tau 1}$ and $\beta_{\tau 2}$ and 1000 sample averages of estimated asymptotic standard errors (s.e.) when counts follow different Poisson, negative binomial and zero-inflated distributions are reported by using different QR methods (TJ, ATJ50, BJ, ABJ50) based on sample size $n = 500$.

		$\beta_{\tau 0}$		$\beta_{\tau 1}$		$\beta_{\tau 2}$	
τ	Method	s.d.	s.e.	s.d.	s.e.	s.d.	s.e.
Case 1. *Poisson*							
0.25	TJ	0.0609	0.0609	0.0450	0.0466	0.0655	0.0682
	ATJ50	0.0691	0.0634	0.0433	0.0418	0.0600	0.0639
	BJ	0.0791	0.0810	0.0570	0.0594	0.0855	0.0850
	ABJ50	0.0641	0.0656	0.0427	0.0454	0.0731	0.0757
0.50	TJ	0.0499	0.0512	0.0452	0.0455	0.0598	0.0635
	ATJ50	0.0470	0.0462	0.0382	0.0418	0.0589	0.0635
	BJ	0.0572	0.0590	0.0443	0.0445	0.0632	0.0662
	ABJ50	0.0528	0.0535	0.0396	0.0400	0.0588	0.0612
0.75	TJ	0.0478	0.0499	0.0371	0.0381	0.0601	0.0599
	ATJ50	0.0452	0.0456	0.0332	0.0331	0.0584	0.0562
	BJ	0.0494	0.0509	0.0376	0.0384	0.0571	0.0608
	ABJ50	0.0469	0.0459	0.0335	0.0335	0.0554	0.0555
Case 2. *Zero-inflated Poisson*							
0.25	TJ	0.2343	0.2348	0.2128	0.2072	0.2247	0.2515
	ATJ50	0.1945	0.2091	0.1824	0.1934	0.1853	0.2319
	BJ	0.2154	0.2180	0.2098	0.1965	0.2208	0.2542
	ABJ50	0.1925	0.2108	0.1876	0.1846	0.1850	0.2453
0.50	TJ	0.0795	0.0810	0.0578	0.0598	0.0907	0.0958
	ATJ50	0.0759	0.0790	0.0532	0.0540	0.0898	0.0927

Continued on next page

Table 4.2 – continued from previous page

		$\beta_{\tau 0}$		$\beta_{\tau 1}$		$\beta_{\tau 2}$	
τ	Method	s.d.	s.e.	s.d.	s.e.	s.d.	s.e.
	BJ	0.0765	0.0851	0.0534	0.0573	0.0904	0.0964
	ABJ50	0.0675	0.0714	0.0426	0.0491	0.0828	0.0867
0.75	TJ	0.0679	0.0658	0.0481	0.0498	0.0691	0.0721
	ATJ50	0.0639	0.0617	0.0452	0.0478	0.0637	0.0722
	BJ	0.0566	0.0617	0.0431	0.0466	0.0688	0.0723
	ABJ50	0.0538	0.0570	0.0392	0.0432	0.0627	0.0685
Case 3. *Negative binomial*							
0.25	TJ	0.1089	0.1101	0.0834	0.0856	0.1678	0.1711
	ATJ50	0.1059	0.1062	0.0620	0.0670	0.1481	0.1509
	BJ	0.1052	0.1142	0.0845	0.0847	0.1557	0.1588
	ABJ50	0.0925	0.0977	0.0603	0.0678	0.1311	0.1398
0.50	TJ	0.1013	0.1024	0.0691	0.0724	0.1197	0.1221
	ATJ50	0.0948	0.0101	0.0675	0.0712	0.1114	0.1149
	BJ	0.0891	0.0947	0.0669	0.0718	0.1197	0.1275
	ABJ50	0.0866	0.0945	0.0637	0.0713	0.1163	0.1265
0.75	TJ	0.0811	0.0825	0.0715	0.0789	0.0968	0.9930
	ATJ50	0.0759	0.0762	0.0620	0.0670	0.0881	0.0899
	BJ	0.0788	0.0816	0.0621	0.0641	0.1069	0.1126
	ABJ50	0.0710	0.0784	0.0539	0.0607	0.0997	0.1088
Case 4. *Zero-inflated negative binomial*							
0.25	TJ	0.1847	0.1989	0.1624	0.1821	0.2089	0.2187
	ATJ50	0.1558	0.1723	0.1511	0.1678	0.1756	0.1949
	BJ	0.1754	0.1917	0.1672	0.1804	0.2162	0.2311

Continued on next page

Table 4.2 – continued from previous page

τ	Method	$\beta_{\tau 0}$ s.d.	$\beta_{\tau 0}$ s.e.	$\beta_{\tau 1}$ s.d.	$\beta_{\tau 1}$ s.e.	$\beta_{\tau 2}$ s.d.	$\beta_{\tau 2}$ s.e.
	ABJ50	0.1495	0.1621	0.1552	0.1699	0.1781	0.1964
0.50	TJ	0.0905	0.0943	0.0714	0.0689	0.1321	0.1398
	ATJ50	0.0890	0.0868	0.0607	0.0597	0.1209	0.1222
	BJ	0.1057	0.1100	0.0744	0.0707	0.1535	0.1601
	ABJ50	0.1048	0.1097	0.0619	0.0613	0.1477	0.1512
0.75	TJ	0.1021	0.1068	0.0821	0.0867	0.1432	0.1479
	ATJ50	0.0949	0.1015	0.0803	0.0861	0.1240	0.1398
	BJ	0.1034	0.1050	0.0767	0.0794	0.1301	0.1311
	ABJ50	0.0934	0.0997	0.0757	0.0793	0.1143	0.1162

Table 4.3: Converage probabilities of 1000 estimates of QR parameters $\beta_{\tau 0}$, $\beta_{\tau 1}$ and $\beta_{\tau 2}$ when counts follow Poisson, negative binomial and zero-inflated distributions are reported by using different QR methods including simple Tweedie-jittering method, average Tweedie-jittering method, simple Beta-jittering method, and average Beta-jittering method. Nominal levels at $\alpha = (0.01, 0.05, 0.1)$, sample size $n = 500$.

τ	Method	$\beta_{\tau 0}$ $P_{0.99}$	$\beta_{\tau 0}$ $P_{0.95}$	$\beta_{\tau 0}$ $P_{0.90}$	$\beta_{\tau 1}$ $P_{0.99}$	$\beta_{\tau 1}$ $P_{0.95}$	$\beta_{\tau 1}$ $P_{0.90}$	$\beta_{\tau 2}$ $P_{0.99}$	$\beta_{\tau 2}$ $P_{0.95}$	$\beta_{\tau 2}$ $P_{0.90}$
Case 1. *Poisson*										
0.25	TJ	0.985	0.931	0.887	0.995	0.935	0.889	0.989	0.947	0.891
	ATJ50	0.970	0.963	0.906	0.988	0.942	0.886	0.990	0.951	0.912
	BJ	0.991	0.956	0.897	0.988	0.945	0.892	0.985	0.950	0.893
	ABJ50	0.989	0.949	0.897	0.992	0.952	0.988	0.982	0.955	0.903
0.50	TJ	0.982	0.943	0.905	0.883	0.955	0.905	0.992	0.947	0.893
	ATJ50	0.990	0.950	0.917	0.990	0.950	0.913	0.991	0.952	0.908

Continued on next page

Table 4.3 – continued from previous page

τ	Method	$\beta_{\tau 0}$			$\beta_{\tau 1}$			$\beta_{\tau 2}$		
		$P_{0.99}$	$P_{0.95}$	$P_{0.90}$	$P_{0.99}$	$P_{0.95}$	$P_{0.90}$	$P_{0.99}$	$P_{0.95}$	$P_{0.90}$
	BJ	0.988	0.949	0.894	0.985	0.943	0.892	0.990	0.954	0.902
	ABJ50	0.987	0.947	0.897	0.986	0.944	0.902	0.992	0.948	0.906
0.75	TJ	0.985	0.958	0.895	0.986	0.953	0.889	0.992	0.953	0.901
	ATJ50	0.988	0.952	0.891	0.983	0.943	0.905	0.987	0.953	0.906
	BJ	0.986	0.932	0.890	0.988	0.938	0.878	0.992	0.948	0.904
	ABJ50	0.987	0.936	0.897	0.986	0.958	0.911	0.987	0.945	0.893
Case 2. *Zero-inflated Poisson*										
0.25	TJ	0.992	0.965	0.910	0.990	0.953	0.904	0.986	0.942	0.895
	ATJ50	0.992	0.957	0.905	0.986	0.945	0.896	0.988	0.946	0.905
	BJ	0.995	0.957	0.911	0.994	0.964	0.912	0.988	0.947	0.895
	ABJ50	0.994	0.955	0.907	0.991	0.961	0.914	0.988	0.949	0.900
0.50	TJ	0.992	0.958	0.909	0.993	0.958	0.906	0.991	0.953	0.914
	ATJ50	0.993	0.960	0.906	0.985	0.943	0.897	0.991	0.949	0.908
	BJ	0.998	0.966	0.915	0.990	0.974	0.915	0.984	0.954	0.914
	ABJ50	0.994	0.966	0.912	0.996	0.962	0.916	0.989	0.939	0.904
0.75	TJ	0.988	0.945	0.891	0.981	0.943	0.892	0.986	0.947	0.913
	ATJ50	0.987	0.954	0.893	0.983	0.958	0.906	0.988	0.947	0.899
	BJ	0.994	0.952	0.910	0.994	0.960	0.912	0.988	0.948	0.897
	ABJ50	0.986	0.956	0.910	0.994	0.962	0.912	0.986	0.946	0.896
Case 3. *Negative binomial*										
0.25	TJ	0.995	0.957	0.908	0.987	0.941	0.890	0.987	0.947	0.897
	ATJ50	0.993	0.943	0.885	0.995	0.955	0.906	0.991	0.952	0.902
	BJ	0.998	0.960	0.910	0.994	0.960	0.884	0.990	0.953	0.906

Continued on next page

Table 4.3 – continued from previous page

		$\beta_{\tau 0}$			$\beta_{\tau 1}$			$\beta_{\tau 2}$		
τ	Method	$P_{0.99}$	$P_{0.95}$	$P_{0.90}$	$P_{0.99}$	$P_{0.95}$	$P_{0.90}$	$P_{0.99}$	$P_{0.95}$	$P_{0.90}$
	ABJ50	0.998	0.968	0.916	0.994	0.958	0.906	0.987	0.942	0.899
0.50	TJ	0.993	0.960	0.911	0.986	0.940	0.899	0.990	0.948	0.910
	ATJ50	0.995	0.965	0.915	0.985	0.930	0.885	0.990	0.952	0.899
	BJ	0.999	0.962	0.912	0.984	0.940	0.898	0.984	0.949	0.901
	ABJ50	0.988	0.962	0.910	0.996	0.960	0.902	0.984	0.946	0.905
0.75	TJ	0.990	0.951	0.902	0.991	0.952	0.908	0.987	0.950	0.892
	ATJ50	0.988	0.946	0.890	0.994	0.966	0.916	0.900	0.951	0.902
	BJ	0.994	0.942	0.889	0.980	0.939	0.898	0.988	0.948	0.906
	ABJ50	0.983	0.948	0.900	0.984	0.965	0.901	0.991	0.949	0.902
Case 4. *Zero-inflated negative binomial*										
0.25	TJ	0.996	0.959	0.910	0.993	0.956	0.907	0.992	0.954	0.906
	ATJ50	0.996	0.954	0.905	0.991	0.953	0.895	0.995	0.956	0.921
	BJ	0.988	0.940	0.900	0.976	0.939	0.886	0.990	0.962	0.907
	ABJ50	0.990	0.938	0.900	0.986	0.966	0.888	0.994	0.952	0.902
0.50	TJ	0.987	0.941	0.893	0.987	0.951	0.898	0.987	0.952	0.902
	ATJ50	0.985	0.939	0.888	0.980	0.950	0.885	0.987	0.947	0.899
	BJ	0.996	0.966	0.994	0.996	0.946	0.902	0.990	0.946	0.896
	ABJ50	0.992	0.960	0.906	0.996	0.958	0.903	0.989	0.954	0.903
0.75	TJ	0.991	0.952	0.907	0.982	0.942	0.891	0.992	0.944	0.903
	ATJ50	0.980	0.945	0.905	0.985	0.945	0.905	0.988	0.946	0.903
	BJ	0.986	0.958	0.916	0.984	0.946	0.904	0.986	0.944	0.895
	ABJ50	0.986	0.954	0.906	0.986	0.944	0.898	0.987	0.946	0.898

4.2.5 Performance of Statistical Testing

Calculating the rejection rate and power of a test is essential to evaluate the effectiveness and reliability of different statistical tests.

Table 4.4: Rejection rates of $S = 1000$ null hypotheses based on the quantile regression parameters $\beta_{\tau 2}$ are reported when using different quantile regression methods TJ, ATJ50, BJ, and ABJ50. Nominal levels $\alpha = (0.01, 0.05, 0.1)$ and sample sizes $n = (200, 500)$.

		$n = 200$			$n = 500$		
τ	Method	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$
Case 1. *Poisson*							
0.25	TJ	0.010	0.047	0.083	0.012	0.063	0.109
	ATJ50	0.016	0.039	0.087	0.010	0.049	0.098
	BJ	0.015	0.054	0.111	0.015	0.050	0.107
	ABJ50	0.010	0.065	0.090	0.012	0.045	0.097
0.50	TJ	0.013	0.049	0.091	0.008	0.053	0.097
	ATJ50	0.006	0.045	0.088	0.009	0.048	0.092
	BJ	0.018	0.059	0.098	0.010	0.046	0.098
	ABJ50	0.015	0.064	0.096	0.008	0.052	0.094
0.75	TJ	0.014	0.047	0.109	0.008	0.047	0.099
	ATJ50	0.014	0.045	0.089	0.013	0.047	0.104
	BJ	0.019	0.063	0.102	0.008	0.046	0.096
	ABJ50	0.008	0.053	0.100	0.013	0.055	0.107
Case 2. *Zero-inflated Poisson*							
0.25	TJ	0.008	0.039	0.073	0.014	0.058	0.105
	ATJ50	0.015	0.048	0.078	0.012	0.054	0.095

Continued on next page

Table 4.4 – continued from previous page

τ	Method	$n = 200$			$n = 500$		
		$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$
	BJ	0.011	0.042	0.072	0.012	0.053	0.105
	ABJ50	0.010	0.041	0.072	0.012	0.051	0.100
0.50	TJ	0.010	0.049	0.101	0.009	0.047	0.086
	ATJ50	0.006	0.049	0.095	0.009	0.051	0.092
	BJ	0.010	0.043	0.085	0.015	0.044	0.089
	ABJ50	0.005	0.035	0.085	0.011	0.061	0.095
0.75	TJ	0.016	0.060	0.096	0.015	0.053	0.097
	ATJ50	0.010	0.044	0.088	0.012	0.053	0.101
	BJ	0.019	0.058	0.104	0.012	0.053	0.104
	ABJ50	0.018	0.058	0.103	0.015	0.054	0.104
Case 3. *Negative binomial*							
0.25	TJ	0.015	0.049	0.088	0.013	0.053	0.103
	ATJ50	0.008	0.043	0.083	0.009	0.048	0.098
	BJ	0.008	0.045	0.084	0.010	0.047	0.094
	ABJ50	0.010	0.040	0.081	0.013	0.058	0.105
0.50	TJ	0.018	0.056	0.097	0.010	0.052	0.090
	ATJ50	0.008	0.044	0.089	0.010	0.048	0.101
	BJ	0.014	0.062	0.106	0.014	0.051	0.097
	ABJ50	0.011	0.040	0.089	0.016	0.054	0.095
0.75	TJ	0.010	0.044	0.105	0.013	0.050	0.102
	ATJ50	0.021	0.068	0.107	0.010	0.049	0.098
	BJ	0.018	0.050	0.098	0.011	0.052	0.094
	ABJ50	0.016	0.046	0.100	0.009	0.051	0.098

Continued on next page

Table 4.4 – continued from previous page

τ	Method	$n = 200$			$n = 500$		
		$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$
Case 4. *Zero-inflated Negative binomial*							
0.25	TJ	0.004	0.037	0.078	0.008	0.046	0.094
	ATJ50	0.005	0.041	0.079	0.005	0.044	0.083
	BJ	0.007	0.035	0.078	0.010	0.038	0.093
	ABJ50	0.007	0.028	0.054	0.006	0.041	0.094
0.50	TJ	0.014	0.050	0.095	0.013	0.048	0.098
	ATJ50	0.010	0.053	0.092	0.013	0.053	0.101
	BJ	0.019	0.047	0.091	0.010	0.054	0.104
	ABJ50	0.015	0.066	0.110	0.011	0.046	0.097
0.75	TJ	0.023	0.064	0.104	0.008	0.056	0.097
	ATJ50	0.010	0.047	0.096	0.012	0.056	0.097
	BJ	0.018	0.061	0.106	0.015	0.055	0.104
	ABJ50	0.015	0.056	0.098	0.012	0.053	0.102

Table 4.4 reports the comprehensive overview of the percentage of rejecting the null hypothesis $H_0 : \beta_{\tau 2} = 0$ vs. $H_0 : \beta_{\tau 2} \neq 0$ using different jittering methods across four discrete distributions at nominal levels of $\alpha = (0.01, 0.05, 0.10)$, respectively, where the sample sizes are $n = (200, 500)$. The rejection rates of H_0, when $n = 200$, are found to be mildly further away from the nominal levels, especially for the lower quantile of zero-inflated negative binomial distribution (Case 4). When the sample size is increased to $n = 500$, the rejection rates of H_0 using all methods on lower quantile of zero-inflated distributions are also close to the corresponding nominal levels. Therefore, the statistical tests associated with our QR methods are both effective and reliable.

Next, we conduct a series of power analyses to assess the performance of the statistical test when testing null hypothesis $H_0 : \beta_{\tau 1} = 0$ vs. $H_0 : \beta_{\tau 1} \neq 0$. The five true values represent the effect size. The objective is to check the capacity of a statistical test to correctly reject a false null hypothesis, with the given size of effect. Firstly. Table 4.5 and 4.5A report the rejection rates of null hypothesis $H_0 : \beta_{\tau 1} = 0$ against alternative regarding QR parameter $\beta_{\tau 1}$ with varying effect size of $(0.2, 0.4, 0.6, 0.8, 1.0)$ and with sample sizes $n = (200, 500)$. The results find all QR methods (TJ, TJ50, BJ, BJ50) give an increasing power of the test with the effect size, which implies an enhanced probability of correctly identifying an effect. Many of them even reach the power of 100% when effect size is 1.0. The power of the test associated with average jittering (ATJ50, ABJ50) are much higher than simple jittering (TJ, BJ), especially on lower quantile of distribution. More specifically, for Poisson and negative binomial distributions, the powers of the test using all proposed methods are all greater than 80% and continue to increase across three quantiles. This solidifies the capacity of the test to identify from small to large effects when the observations follow Poisson and negative binomial distributions. When the observations follow zero-inflated Poisson (Case 3) and zero-inflated negative binomial (Case 4), however, the powers of test using all proposed methods become lower especially on low quantile $\tau = 0.25$. The possible reason regarding the lower power is that the excessive zeros make the standard errors significantly larger, making effect harder to detect. This evidence can be found in Table 4.2. Also, the presence of massive zeros dilutes the observable effect size. On the low quantile of zero-inflated distribution, the dependent variable Y often equals zero regardless of the value of predictor $\boldsymbol{X}$, making identifying an effect becomes difficult. On the other hand, insufficient sample size also contributes to the worse statistical test's capacity of identifying an effect. When we increase the sample size to $n = 500$, the variability of regression estimator decreases, and the powers of the test associated with proposed methods are significantly increased for all distributions. In most of the cases the tests are able to 100% correctly reject a false null hypothesis, detecting small to large effects. It is worth noting that after the sample size increases, the threshold of effect size at which the test demonstrates high efficacy in detecting effects (i.e., $1-\beta \geq 80\%$) becomes 0.6, which is a medium effect size (Cohen, 1988, p. 25), the power of the test associated with proposed method can detect a medium effect size with $n = 500$ even though excessive zeros are present.

In conclusion, the hypothesis test associated with proposed methods is capable of fairly controlling the type I error. The performance of hypothesis test that detects an effect is generally good. By increasing sample sizes n and jittered samples size m, hypothesis test associated with proposed methods can detect much smaller effect size. Therefore, the parametric hypothesis tests underpinning the proposed methods demonstrate reliable efficacy for hypothesis test.

Table 4.5: Power of $S = 1000$ hypotheses test (i.e., $1-\beta$) as related to the quantile regression parameters $\beta_{\tau 1}$ are reported by using different quantile regression methods TJ, ATJ50, BJ, and ABJ50. Nominal level $\alpha = 0.05$ and sample size $n = 200$.

τ	Method	$\beta_{\tau 1} = 0.2$	$\beta_{\tau 1} = 0.4$	$\beta_{\tau 1} = 0.6$	$\beta_{\tau 1} = 0.8$	$\beta_{\tau 1} = 1.0$
Case 1. *Poisson*						
0.25	TJ	0.862	1.000	1.000	1.000	1.000
	ATJ50	0.947	1.000	1.000	1.000	1.000
	BJ	0.842	1.000	1.000	1.000	1.000
	ABJ50	0.942	1.000	1.000	1.000	1.000
0.50	TJ	0.922	1.000	1.000	1.000	1.000
	ATJ50	0.977	1.000	1.000	1.000	1.000
	BJ	0.933	1.000	1.000	1.000	1.000
	ABJ50	0.976	1.000	1.000	1.000	1.000

Continued on next page

Table 4.5 – continued from previous page

τ	Method	$\beta_{\tau 1} = 0.2$	$\beta_{\tau 1} = 0.4$	$\beta_{\tau 1} = 0.6$	$\beta_{\tau 1} = 0.8$	$\beta_{\tau 1} = 1.0$
0.75	TJ	0.932	1.000	1.000	1.000	1.000
	ATJ50	0.989	1.000	1.000	1.000	1.000
	BJ	0.921	1.000	1.000	1.000	1.000
	ABJ50	0.981	1.000	1.000	1.000	1.000
Case 2. *Zero-inflated Poisson*						
0.25	TJ	0.232	0.467	0.620	0.713	0.755
	ATJ50	0.356	0.512	0.688	0.720	0.788
	BJ	0.212	0.464	0.645	0.711	0.748
	ABJ50	0.323	0.492	0.683	0.735	0.794
0.50	TJ	0.729	0.997	0.999	1.000	1.000
	ATJ50	0.764	1.000	1.000	1.000	1.000
	BJ	0.703	0.990	0.999	1.000	1.000
	ABJ50	0.779	1.000	1.000	1.000	1.000
0.75	TJ	0.847	0.999	1.000	1.000	1.000
	ATJ50	0.875	1.000	1.000	1.000	1.000
	BJ	0.853	0.998	1.000	1.000	1.000
	ABJ50	0.873	1.000	1.000	1.000	1.000
Case 3. *Negative binomial*						
0.25	TJ	0.370	0.898	0.991	0.997	1.000
	ATJ50	0.388	0.916	1.000	1.000	1.000
	BJ	0.387	0.896	0.993	0.998	1.000
	ABJ50	0.415	0.919	1.000	1.000	1.000
0.50	TJ	0.599	0.978	1.000	1.000	1.000

Continued on next page

Table 4.5 – continued from previous page

τ	Method	$\beta_{\tau 1} = 0.2$	$\beta_{\tau 1} = 0.4$	$\beta_{\tau 1} = 0.6$	$\beta_{\tau 1} = 0.8$	$\beta_{\tau 1} = 1.0$
	ATJ50	0.629	1.000	1.000	1.000	1.000
	BJ	0.614	0.980	1.000	1.000	1.000
	ABJ50	0.634	1.000	1.000	1.000	1.000
0.75	TJ	0.657	0.995	1.000	0.997	1.000
	ATJ50	0.639	1.000	1.000	0.999	1.000
	BJ	0.622	0.993	1.000	0.997	1.000
	ABJ50	0.659	1.000	1.000	0.997	1.000
Case 4. *Zero-inflated Negative binomial*						
0.25	TJ	0.095	0.204	0.361	0.494	0.589
	ATJ50	0.125	0.315	0.552	0.732	0.823
	BJ	0.073	0.219	0.357	0.489	0.580
	ABJ50	0.123	0.309	0.545	0.718	0.810
0.50	TJ	0.309	0.816	0.976	0.992	0.997
	ATJ50	0.375	0.845	0.975	0.995	0.999
	BJ	0.339	0.783	0.968	0.994	0.998
	ABJ50	0.376	0.843	0.974	0.994	0.999
0.75	TJ	0.506	0.941	0.989	0.998	1.000
	ATJ50	0.555	0.955	1.000	1.000	1.000
	BJ	0.558	0.945	1.000	1.000	1.000
	ABJ50	0.491	0.957	0.991	0.996	1.000

4.3 Simulation Study II: QR with Truncated Data

In the 2^{nd} simulation we investigate the performance of the QR methods when fitting truncated data whose errors follow normal, student's t and Chi-squared distributions, respectively. Truncation is a fundamental method to bound a number at a particular value and round it down to a given degree of accuracy. In this simulation, truncation is used as a specific data processing and keeps integer part as the degree of accuracy. Notice that a truncation function operates differently from a floor function, denoted by $\lfloor \cdot \rfloor$, when rounding negative numbers. For example, for the truncation function denoted by $\text{trunc}(\cdot)$, we have $\text{trunc}(3.8) = 3 = \lfloor 3.8 \rfloor$ but $\text{trunc}(-4.9) = -4 \neq -5 = \lfloor -4.9 \rfloor$. Truncation is a popular data processing method in real world. By eliminating decimal points and the digits that follow, the use of storage space and processing speed can be optimized. Floating-point arithmetic is an example that represents subsets of real numbers using an integer with a fixed precision (Muller et al., 2010). Therefore, understanding the performance of QR methods on truncated data is highly in demand.

In this section, the application of linear transformation is demonstrated. The comparison of bias, mean squared error and relative efficiency of QR estimator using different methods are reported.

4.3.1 Data and Simulation Setup

Counts data are generated by truncating continuous data with errors following standard normal, student's t and Chi-squared distributions. We first generate continuous data, for $i = 1, ..., n$,

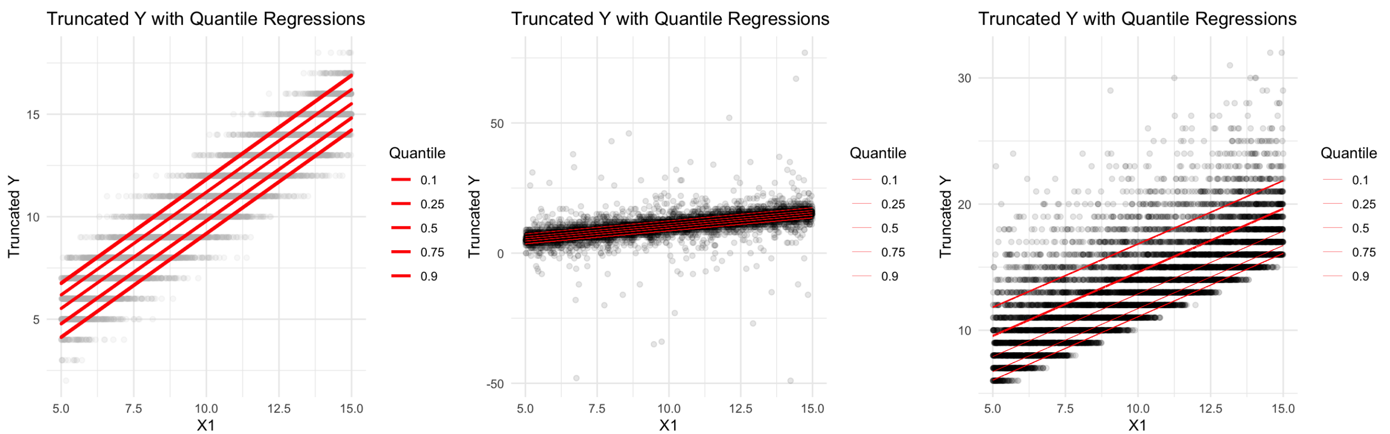
$$y_i^c = \beta_{\tau 0} + \beta_{\tau 1} x_{i1} + \beta_{\tau 2} x_{i2} + \varepsilon_i, \tag{4.10}$$

where y_i^c denotes the continuous observation, $\boldsymbol{\beta}' = (\beta_0, \beta_1, \beta_2)$ is the vector of linear regression coefficients and

$$x_{i1} \sim \text{Uniform } (5, 15),$$

$$x_{i2} = \begin{cases} 1 & \text{with probability } 0.2 \\ 0 & \text{with probability } 0.8. \end{cases}$$

Then, apply the truncation function and let $y_i = \text{trunc}(y_i^c)$ be the count. The sample $(y_i, x_{i1}, x_{i2})_{i=1}^n$ is selected from the population in each of the simulation runs. The error term follows the three different distributions:



(a) *Case* 1. Standard Normal (b) *Case* 2. Student's t with $df = 2$ (c) *Case* 3. Chi-squared with $df = 3$

Figure 4.2: Truncated Y with respect to $\boldsymbol{x}_1$ generated from linear models whose error term follows normal, student's t and Chi-squared distributions, respectively. Red lines indicate quantiles of distributions on $\tau = (0.10, 0.25, 0.5, 0.75, 0.90)$

***Case* 1.** Normal distribution. Assume the mean and variance are 0 and 1, respectively, $\varepsilon \sim \mathcal{N}(0, 1)$;

***Case* 2.** Student's t distribution. Assume the degree of freedom is 2, $\varepsilon \sim$ Student's t_2;

***Case* 3.** Chi-squared distribution. Assume the degree of freedom is 3, $\varepsilon \sim \mathcal{X}_3^2$.

Figure 4.2 plots the relationship of truncated Y and first independent variable x_1 where the error term follows standard normal, student's t $df = 2$, and Chi-squared $df = 3$ distributions, respectively. The five red lines in each graph indicate the $(0.1, 0.25, 0.5, 0.75, 0.9)$-th quantiles of corresponding distribution. As we can see in Fig 4.2a, the error term following standard normal distribution is symmetric and reveals the homoscedasticity. In Fig 4.2b, the error term follows student's t distribution is also symmetric, but has a significantly heavier tail than normal distribution. In Fig 4.2c, Chi-squared distributed error term shows right skewness, with the distribution of error concentrated on the left side and exhibiting a long tail to the right side. We will fit the QR model to these data sets and show the performance of proposed QR model against outliers, extreme values and skewness.

In this simulation, population size $N = 50000$ is used to obtain the true QR coefficients and $\beta_0 = 1, \beta_1 = 1, \beta_2 = 0$. Note that same as simulation I, since $\beta_2 = 0$ does not contribute any effect in the resulting QR model, $\beta_{\tau 2} = 0$. $S = 1000$ iterations are run to assess the performance of proposed QR models on the $\tau = (0.1, 0.25, 0.5, 0.75, 0.9) - th$ quantile of data set.

4.3.2 QR Methods and Transformations

The following QR methods are used in the simulation for comparison:

1. Ordinary quantile regression method (ORD);

2. Mid-quantile regression method (MID);

3. Simple uniform jittering method (UJ);

4. Average uniform jittering method with $m = 50$ jittered samples (AUJ50);

5. Simple Tweedie jittering method (TJ);

6. Average Tweedie jittering method with $m = 50$ jittered samples (ATJ50);

7. Simple Beta jittering method (BJ); and

8. Average Beta jittering method with $m = 50$ jittered samples (ABJ50).

Firstly the performance of optimized jittering method is compared with other methods. We report bias, mean squared error and relative efficiency of estimators using all the above QR models. Then, average value of approximate standard error of $S = 1000$ QR estimates and their standard deviations are compared.

Let $z_i = y_i + u_i$ be the i-th QR response, for $i = 1, ..., n$,

$$T(z_i;\tau) = \begin{cases} z_i - Q_U(\tau) & \text{if } z_i \geq Q_U(\tau), \\ \varsigma & \text{if } z_i < Q_U(\tau). \end{cases} \tag{4.11}$$

where ς is a small positive number and $Q_U(\tau)$ is the τ^{th} quantile of U following either Tweedie or Beta distribution. On the sample level, the QR model is specified as

$$Q_{T(z;\tau)} = \boldsymbol{x}'\boldsymbol{\beta}_\tau,$$

where $\boldsymbol{\beta}_\tau$ is a vector of QR coefficients.

4.3.3 Comparison of Different QR Methods

The results of $S = 1000$ simulation iterations of QR models with different error distributions are analyzed.

Table 4.6: Biases ($bias$), mean squared errors (MSE) and relative efficiencies (REF) of estimators $\beta_{\tau 0}$, $\beta_{\tau 1}$, and $\beta_{\tau 2}$ on five quantiles $\tau = (0.10, 0.25, 0.5, 0.75, 0.90)$ of data when error follows normal, student's t and Chi-squared are reported for different methods (ORD, MID, TJ, ATJ50, ATJ150, BJ, ABJ50, ABJ150). Sample size $n = 200$.

		$\beta_{\tau 0}$				$\beta_{\tau 1}$				$\beta_{\tau 2}$			
τ	Method	Bias	Bias*	MSE	REF	Bias	Bias*	MSE	REF	Bias	Bias*	MSE	REF
Case 1. $\varepsilon \sim \mathcal{N}(0,1)$													
0.10	ORD	0.1609	-0.0661	0.2487	1.000	-0.0036	-0.0018	0.0032	1.000	0.0507	0.0301	0.1108	1.000
	MID	-	-		-	-	-	-	-	-	-	-	-
	UJ	0.1935	-0.0661	0.2375	1.047	-0.0165	-0.0018	0.0027	1.185	-0.2966	0.0301	0.0880	1.259
	AUJ50	0.0252	-0.0661	0.1743	1.426	-0.0009	-0.0018	0.0016	2.000	0.0496	0.0301	0.0879	1.261
	TJ	-0.0046	-0.0661	0.2388	1.041	0.0035	-0.0018	0.0025	1.28	0.0094	0.0301	0.1054	1.051
	ATJ50	0.0093	-0.0661	0.1593	1.561	-0.0044	-0.0018	0.0014	2.286	0.0241	0.0301	0.0911	1.216
	BJ	-0.0036	-0.0661	0.2277	1.092	0.0020	-0.0018	0.0020	1.600	-0.0080	0.0301	0.1032	1.074
	ABJ50	0.0089	-0.0661	0.1497	1.661	-0.0049	-0.0018	0.0010	3.200	0.0229	0.0301	0.0805	1.376
0.25	ORD	-0.0166	-0.0047	0.1572	1.000	0.0034	0.0009	0.0015	1.000	0.0049	0.0057	0.0709	1.000
	MID	-0.1858	-0.0047	0.1389	1.132	0.0054	0.0009	0.0010	1.500	-0.0020	0.0057	0.0646	1.098

Continued on next page

Table 4.6 – continued from previous page

		$\beta_{\tau 0}$				$\beta_{\tau 1}$				$\beta_{\tau 2}$			
τ	Method	Bias	Bias*	MSE	REF	Bias	Bias*	MSE	REF	Bias	Bias*	MSE	REF
	UJ	-0.0513	-0.0047	0.1381	1.138	0.0035	0.0009	0.0012	1.250	0.0392	0.0057	0.0707	1.003
	AUJ50	-0.0221	-0.0047	0.1289	1.220	0.0023	0.0009	0.0012	1.250	0.0033	0.0057	0.0503	1.410
	TJ	0.0029	-0.0047	0.1535	1.024	0.0013	0.0009	0.0017	0.882	0.0128	0.0057	0.0675	1.050
	ATJ50	0.0027	-0.0047	0.1055	1.490	0.0024	0.0009	0.0010	1.500	-0.0250	0.0057	0.0491	1.444
	BJ	-0.0017	-0.0047	0.1438	1.093	-0.0000	0.0009	0.0012	1.250	-0.0137	0.0057	0.0637	1.113
	ABJ50	-0.0068	-0.0047	0.1042	1.509	0.0002	0.0009	0.0009	1.667	-0.0143	0.0057	0.0471	1.505
0.50	ORD	-0.0475	-0.0331	0.1396	1.000	0.0065	-0.0004	0.0012	1.000	-0.0283	0.0079	0.0652	1.000
	MID	0.1362	-0.0331	0.1351	1.033	0.0009	-0.0004	0.0010	1.200	0.0036	0.0079	0.0501	1.301
	UJ	0.0137	-0.0331	0.1323	1.055	-0.0006	-0.0004	0.0011	1.091	0.0281	0.0079	0.0629	1.037
	AUJ50	-0.0181	-0.0331	0.1002	1.393	0.0024	-0.0004	0.0008	1.500	-0.0263	0.0079	0.0471	1.384
	TJ	-0.0299	-0.0331	0.1245	1.121	-0.0001	-0.0004	0.0011	1.091	0.0034	0.0079	0.0599	1.088
	ATJ50	-0.0388	-0.0331	0.1077	1.296	0.0033	-0.0004	0.0010	1.200	0.0055	0.0079	0.0444	1.468
	BJ	-0.0000	-0.0331	0.1233	1.132	-0.0000	-0.0004	0.0010	1.200	-0.0066	0.0079	0.0559	1.166
	ABJ50	-0.0370	-0.0331	0.1041	1.341	0.0043	-0.0004	0.0009	1.333	-0.0092	0.0079	0.0403	1.618

Continued on next page

Table 4.6 – continued from previous page

		$\beta_{\tau 0}$				$\beta_{\tau 1}$				$\beta_{\tau 2}$			
τ	Method	Bias	Bias*	MSE	REF	Bias	Bias*	MSE	REF	Bias	Bias*	MSE	REF
0.75	ORD	-0.0371	-0.0564	0.1777	1.000	-0.0054	-0.0011	0.0017	1.000	0.0175	0.0100	0.0746	1.000
	MID	0.0728	-0.0564	0.1455	1.221	-0.0086	-0.0011	0.0012	1.417	0.0224	0.0100	0.0757	0.985
	UJ	0.0722	-0.0564	0.1706	1.042	-0.0067	-0.0011	0.0015	1.133	-0.0361	0.0100	0.0609	1.225
	AUJ50	0.0116	-0.0564	0.1206	1.473	-0.0023	-0.0011	0.0010	1.700	0.0059	0.0100	0.0503	1.483
	TJ	-0.0336	-0.0564	0.1590	1.118	0.0013	-0.0011	0.0013	1.308	0.0388	0.0100	0.0662	1.127
	ATJ50	-0.0354	-0.0564	0.1025	1.734	0.0004	-0.0011	0.0010	1.700	0.0234	0.0100	0.0598	1.247
	BJ	-0.0093	-0.0564	0.1588	1.119	0.0015	-0.0011	0.0014	1.214	-0.0447	0.0100	0.0665	1.122
	ABJ50	-0.0242	-0.0564	0.1013	1.754	0.0003	-0.0011	0.0009	1.889	0.0211	0.0100	0.0572	1.304
0.90	ORD	0.0429	-0.0117	0.2254	1.000	-0.0041	0.0002	0.0028	1.000	-0.0172	0.0029	0.1133	1.000
	MID	-	-	-	-	-	-	-	-	-	-	-	-
	UJ	-0.0390	-0.0117	0.2167	1.040	0.0038	0.0003	0.0020	1.400	-0.0091	0.0029	0.1055	1.074
	AUJ50	0.0166	-0.0117	0.1544	1.460	-0.0015	0.0003	0.0014	2.000	-0.0270	0.0029	0.0725	1.563
	TJ	0.0087	-0.0117	0.2165	1.041	-0.0015	0.0003	0.0019	1.474	-0.0099	0.0029	0.0756	1.499
	ATJ50	-0.0114	-0.0117	0.1648	1.368	0.0003	0.0003	0.0013	2.154	0.0017	0.0029	0.0783	1.447

Continued on next page

Table 4.6 – continued from previous page

τ	Method	$\beta_{\tau 0}$				$\beta_{\tau 1}$				$\beta_{\tau 2}$			
		Bias	Bias*	MSE	REF	Bias	Bias*	MSE	REF	Bias	Bias*	MSE	REF
	BJ	0.0072	-0.0117	0.2155	1.046	-0.0013	0.0003	0.0019	1.474	-0.0098	0.0029	0.1064	1.065
	ABJ50	-0.0113	-0.0117	0.1629	1.384	0.0003	0.0003	0.0014	2.000	0.0016	0.0029	0.0742	1.527
Case 2. *$\varepsilon \sim$ student's t_2*													
0.10	ORD	-0.0634	0.0422	1.0538	1.000	-0.0008	0.0009	0.0097	1.000	-0.0707	0.0065	0.5574	1.000
	MID	0.0977	0.0422	1.0306	1.023	-0.0094	0.0009	0.0088	1.102	-0.1409	0.0065	0.4195	1.329
	UJ	0.0706	0.0422	1.1157	0.945	-0.0027	0.0009	0.0107	0.907	-0.0712	0.0065	0.5564	1.002
	AUJ50	0.0547	0.0422	0.9339	1.128	-0.0067	0.0009	0.0086	1.128	-0.0594	0.0065	0.4296	1.297
	TJ	0.0657	0.0422	0.9987	1.055	-0.0095	0.0009	0.0089	1.090	-0.0465	0.0065	0.4451	1.252
	ATJ50	0.0565	0.0422	0.8646	1.219	0.0007	0.0009	0.0079	1.228	-0.0642	0.0065	0.3800	1.467
	BJ	0.0648	0.0422	0.9965	1.058	-0.0090	0.0009	0.0087	1.115	-0.0482	0.0065	0.4441	1.255
	ABJ50	-0.0544	0.0422	0.8634	1.221	0.0006	0.0009	0.0077	1.260	-0.0628	0.0065	0.3790	1.471
0.25	ORD	-0.0624	0.0353	0.2771	1.000	0.0030	-0.0002	0.0037	1.000	-0.0888	0.0133	0.1372	1.000
	MID	0.1757	0.0353	0.2640	1.050	-0.0263	-0.0002	0.0023	1.609	0.0321	0.0133	0.1306	1.051
	UJ	0.0637	0.0353	0.2723	1.018	-0.0032	-0.0002	0.0037	1.000	-0.0426	0.0133	0.1387	0.989

Continued on next page

Table 4.6 – continued from previous page

τ	Method	$\beta_{\tau 0}$				$\beta_{\tau 1}$				$\beta_{\tau 2}$			
		Bias	Bias*	MSE	REF	Bias	Bias*	MSE	REF	Bias	Bias*	MSE	REF
	AUJ50	0.0038	0.0353	0.2015	1.375	0.0009	-0.0002	0.0018	2.056	-0.0745	0.0133	0.1117	1.228
	TJ	0.0609	0.0353	0.2627	1.055	-0.0034	-0.0002	0.0031	1.194	-0.0370	0.0133	0.1173	1.170
	ATJ50	0.0358	0.0353	0.2264	1.224	-0.0027	-0.0002	0.0024	1.542	0.0599	0.0133	0.0843	1.628
	BJ	-0.0460	0.0353	0.2708	1.023	0.0031	-0.0002	0.0025	1.480	0.0297	0.0133	0.1108	1.238
	ABJ50	0.0352	0.0353	0.2259	1.227	-0.0025	-0.0002	0.0021	1.762	0.0594	0.0133	0.0802	1.711
0.50	ORD	-0.1324	0.0165	0.1785	1.000	0.0071	0.0018	0.0017	1.000	-0.0060	0.0205	0.0931	1.000
	MID	-0.0350	0.0165	0.1749	1.021	0.0042	0.0018	0.0014	1.214	0.03160	0.0205	0.0787	1.183
	UJ	0.0145	0.0165	0.1679	1.063	-0.0015	0.0018	0.0015	1.133	-0.0172	0.0205	0.0730	1.137
	AUJ50	-0.0177	0.0165	0.1386	1.288	0.0032	0.0018	0.0012	1.417	-0.0187	0.0205	0.0699	1.187
	TJ	-0.0124	0.0165	0.1502	1.188	0.0027	-0.0018	0.0016	1.062	0.0355	0.0205	0.0923	0.899
	ATJ50	-0.0532	0.0165	0.1205	1.481	0.0027	0.0018	0.0013	1.308	0.0026	0.0205	0.0598	1.388
	BJ	-0.0120	0.0165	0.1499	1.191	0.0026	0.0018	0.0015	1.133	0.0343	0.0205	0.0917	0.905
	ABJ50	-0.0140	0.0165	0.1201	1.486	0.0020	0.0018	0.0012	1.417	0.0025	0.0205	0.0589	1.409
0.75	ORD	-0.1281	0.0186	0.3044	1.000	0.0132	0.0004	0.0026	1.000	-0.0511	0.0231	0.1275	1.000

Continued on next page

Table 4.6 – continued from previous page

		$\beta_{\tau 0}$				$\beta_{\tau 1}$				$\beta_{\tau 2}$			
τ	Method	Bias	Bias*	MSE	REF	Bias	Bias*	MSE	REF	Bias	Bias*	MSE	REF
	MID	-0.0852	0.0186	0.2484	1.225	0.0077	0.0004	0.0020	1.000	0.0530	0.0231	0.1038	1.228
	UJ	-0.0820	0.0186	0.3017	1.009	0.0079	0.0004	0.0026	1.300	0.0173	0.0231	0.1299	0.982
	AUJ50	-0.0827	0.0186	0.2091	1.456	0.0089	0.0004	0.0019	1.368	-0.0202	0.0231	0.1214	1.050
	TJ	0.0364	0.0186	0.2897	1.051	-0.0025	0.0004	0.0025	1.040	0.0135	0.0231	0.1214	1.050
	ATJ50	-0.0110	0.0186	0.2059	1.478	0.0001	0.0004	0.0019	1.368	0.0256	0.0231	0.1210	1.054
	BJ	0.0342	0.0186	0.2875	1.059	-0.0022	0.0004	0.0025	1.040	0.0124	0.0231	0.1272	1.312
	ABJ50	-0.0104	0.0186	0.2051	1.484	0.0001	0.0004	0.0018	1.444	0.0248	0.0231	0.1194	1.068
0.90	ORD	-0.0193	0.0890	1.0314	1.000	0.0025	-0.0028	0.0093	1.000	0.1427	0.0440	0.5735	1.000
	MID	-	-	-	-	-	-	-	-	-	-	-	-
	UJ	-0.0135	0.0890	1.0797	0.955	-0.0040	-0.0028	0.0100	0.930	-0.0192	0.0440	0.5562	1.031
	AUJ50	-0.0564	0.0890	0.8667	1.190	0.0031	-0.0028	0.0077	1.208	0.1387	0.0440	0.4893	1.172
	TJ	-0.0734	0.0890	0.9721	1.061	-0.0072	-0.0028	0.0088	1.057	-0.0262	0.0440	0.5412	1.060
	ATJ50	-0.0878	0.0890	0.7798	1.323	0.0145	-0.0028	0.0073	1.274	0.0315	0.0440	0.4884	1.174
	BJ	-0.0726	0.0890	0.9715	1.062	-0.0069	-0.0028	0.0086	1.081	-0.0251	0.0440	0.5406	1.061

Continued on next page

Table 4.6 – continued from previous page

τ	Method	$\beta_{\tau 0}$				$\beta_{\tau 1}$				$\beta_{\tau 2}$			
		Bias	Bias*	MSE	REF	Bias	Bias*	MSE	REF	Bias	Bias*	MSE	REF
	ABJ50	-0.0894	0.0890	0.7786	1.325	0.0044	-0.0028	0.0073	1.274	0.0311	0.0440	0.4814	1.191
Case 3. $\varepsilon \sim \mathcal{X}_2^2$													
0.10	ORD	0.0580	0.0295	0.2202	1.000	-0.0010	0.0020	0.0020	1.000	0.0314	0.0016	0.0893	1.000
	MID	-	-	-	-	-	-	-	-	-	-	-	-
	UJ	-0.0441	0.0295	0.2180	1.010	0.0066	0.0020	0.0016	1.250	0.0288	0.0016	0.0883	1.011
	UJ50	0.0072	0.0295	0.1348	1.634	0.0018	0.0020	0.0012	1.667	0.0232	0.0016	0.0551	1.621
	TJ	0.0321	0.0295	0.2145	1.027	-0.0014	0.0020	0.0018	1.111	0.0012	0.0016	0.0884	1.01
	ATJ50	-0.0386	0.0295	0.1310	1.681	-0.0018	0.0020	0.0010	2.000	-0.0011	0.0016	0.0533	1.675
	BJ	0.0342	0.0295	0.2173	1.013	-0.0018	0.0020	0.0019	1.053	-0.0015	0.0016	0.0897	0.996
	ABJ50	-0.0377	0.0295	0.1316	1.673	0.0046	0.0020	0.0011	1.818	-0.0016	0.0016	0.0544	1.642
0.25	ORD	-0.0746	-0.0424	0.2468	1.000	0.0052	-0.0023	0.0022	1.000	0.0470	0.0218	0.1113	1.000
	MID	-0.1237	-0.0424	0.2025	1.219	0.0093	-0.0023	0.0016	1.375	0.1285	0.0218	0.0804	1.384
	UJ	-0.0705	-0.0424	0.2589	0.953	0.0070	-0.0023	0.0022	1.000	0.0168	0.0218	0.1005	1.150
	AUJ50	-0.0569	-0.0424	0.1806	1.367	0.0059	-0.0023	0.0017	1.294	0.0253	0.0218	0.0840	1.325

Continued on next page

		$\beta_{\tau 0}$				$\beta_{\tau 1}$				$\beta_{\tau 2}$			
τ	Method	Bias	Bias*	MSE	REF	Bias	Bias*	MSE	REF	Bias	Bias*	MSE	REF
	TJ	0.0744	-0.0424	0.2410	1.024	-0.0051	-0.0023	0.0019	1.158	-0.0361	0.0218	0.1100	1.012
	ATJ50	0.0345	-0.0424	0.1732	1.425	-0.0029	-0.0023	0.0014	1.571	-0.0287	0.0218	0.0778	1.431
	BJ	0.0753	-0.0424	0.2414	1.022	-0.0053	-0.0023	0.0020	1.100	0.0371	0.0218	0.1106	1.006
	ABJ50	0.0360	-0.0424	0.1753	1.408	-0.0037	-0.0023	0.0015	1.467	0.0293	0.0218	0.0787	1.414
0.50	ORD	0.0805	-0.0216	0.4956	1.162	-0.0038	-0.0018	0.0069	1.000	-0.0450	0.0181	0.2298	1.000
	MID	0.1827	-0.0216	0.3170	1.563	-0.0177	-0.0018	0.0028	2.464	-0.0177	0.0181	0.1566	1.467
	UJ	0.0262	-0.0216	0.4887	1.014	-0.0001	-0.0018	0.0048	1.438	0.0287	0.0181	0.1216	1.037
	AUJ50	0.0369	-0.0216	0.4027	1.231	0.0008	-0.0018	0.0040	1.725	-0.0318	0.0181	0.2081	1.104
	TJ	0.0673	-0.0216	0.4908	1.010	-0.0021	-0.0018	0.0044	1.568	0.0229	0.0181	0.2175	1.057
	ATJ50	-0.0109	-0.0216	0.4266	1.162	0.0052	-0.0018	0.0038	1.816	-0.0347	0.0181	0.1948	1.180
	BJ	0.0001	-0.0216	0.4842	1.024	-0.0015	-0.0018	0.0043	1.605	0.0218	0.0181	0.2136	1.076
	ABJ50	0.0056	-0.0216	0.4309	1.15	0.0006	-0.0018	0.0039	1.769	0.0123	0.0181	0.1755	1.309
0.75	ORD	0.0621	0.0471	1.1255	1.000	-0.0043	0.0055	0.0110	1.000	0.1243	0.0042	0.5406	1.000
	MID	-0.4472	0.0471	1.1112	1.013	0.0398	0.0055	0.0102	1.078	-0.0119	0.0042	0.5020	1.077

Continued on next page

Table 4.6 – continued from previous page

τ	Method	$\beta_{\tau 0}$ Bias	$\beta_{\tau 0}$ Bias*	$\beta_{\tau 0}$ MSE	$\beta_{\tau 0}$ REF	$\beta_{\tau 1}$ Bias	$\beta_{\tau 1}$ Bias*	$\beta_{\tau 1}$ MSE	$\beta_{\tau 1}$ REF	$\beta_{\tau 2}$ Bias	$\beta_{\tau 2}$ Bias*	$\beta_{\tau 2}$ MSE	$\beta_{\tau 2}$ REF
	UJ	-0.0339	0.0471	1.0230	1.100	-0.0009	0.0055	0.0104	1.058	0.0204	0.0042	0.5384	1.004
	AUJ50	0.0725	0.0471	1.1119	1.012	-0.0093	0.0055	0.0107	1.028	0.0148	0.0042	0.5115	1.057
	TJ	0.0399	0.0471	1.1225	1.003	-0.0022	0.0055	0.0106	1.038	0.0146	0.0042	0.5314	1.017
	ATJ50	0.1380	0.0471	1.1112	1.013	-0.0098	0.0055	0.0097	1.134	0.0125	0.0042	0.5055	1.069
	BJ	0.0825	0.0471	1.0980	1.025	-0.0066	0.0055	0.0099	1.111	-0.0273	0.0042	0.5335	1.013
	ABJ50	-0.0683	0.0471	1.0389	1.083	0.0058	0.0055	0.0099	1.111	0.0093	0.0042	0.5087	1.063
0.90	ORD	-0.0843	-0.0403	3.3425	1.000	-0.0143	0.0085	0.0324	1.000	0.0260	0.0239	1.5954	1.000
	MID	0.0915	-0.0403	3.0554	1.094	0.0304	0.0085	0.0267	1.213	0.0173	0.0239	1.4367	1.110
	UJ	-0.0636	-0.0403	3.5619	0.938	-0.0049	0.0085	0.0319	1.016	0.0925	0.0239	1.4421	1.106
	AUJ50	-0.0770	-0.0403	3.1195	1.071	-0.0082	0.0085	0.0273	1.187	0.0777	0.0239	1.4403	1.108
	TJ	0.0668	-0.0403	3.2948	1.014	-0.0112	0.0085	0.0297	1.091	0.0178	0.0239	1.3594	1.174
	ATJ50	0.0531	-0.0403	2.6741	1.250	-0.0102	0.0085	0.0242	1.339	0.0819	0.0239	1.5419	1.035
	BJ	0.0535	-0.0403	3.1760	1.052	-0.0099	0.0085	0.0288	1.125	-0.0740	0.0239	1.4209	1.123
	ABJ50	-0.0490	-0.0403	2.7923	1.197	-0.0028	0.0085	0.0256	1.266	0.0278	0.0239	1.4336	1.113

Table 4.6 and Table 4.6A report the biases, mean squared errors and relative efficiencies of estimators using all QR methods where the error term follows normal, student's t and Chi-squared distributions, respectively with sample size $n = (200, 500)$. The table finds that the biases of estimator using different methods are comparable, and small for $\beta_{\tau 1}$ and $\beta_{\tau 2}$. In most of the cases, the jittering methods (UJ, AUJ50, BJ, ABJ50, TJ, ATJ50) provides less biased QR estimators than other methods, and the presence of bias of estimator can be slightly improved by using technique of averaging noises out. Comparing the mean squared error and relative efficiency of estimator using different methods, the proposed methods outperform ordinary method and uniform jittering methods. Meanwhile, mid-QR method sometimes gives more efficient estimator than our simple jittering methods and performs very stably across quantiles of data with different error, and hence the strength of mid-QR is emphasized. However, it sometimes fails to estimate QR parameters on extreme quantiles (e.g., $\tau = (0.1, 0.9)$) due to quantile level out of estimable range. Also, check both mean squared error and relative efficiency of the estimator, the proposed simple jittering methods (TJ, BJ) and average jittering methods (ABJ50, ATJ50) are still able to provide more efficient estimators for QR coefficients. We also found that the Beta jittering gives an efficient estimator more frequently than Tweedie jittering method. Considering its better performance and lighter computational burden, Beta distribution seems to be a better choice for jittering the counts. On the other hand, estimator obtained by using proposed methods (TJ, BJ, ATJ50, ABJ50) are found to be more efficient on the lower quantile ($\tau = 0.1$) and upper quantile ($\tau = 0.9$) of the student's t distribution and upper quantile ($\tau = 0.9$) of Chi-squared distribution. The proposed jittering methods are robust against heavy tails and skewness. From Table 4.6A we find that the biases and mean squared errors of estimators obtained by using different methods are significantly reduced for $\beta_{\tau 0}$, $\beta_{\tau 1}$ and $\beta_{\tau 2}$ by increasing the sample size.

In short, the proposed methods provide consistent estimators for data with small and moderate sample size. By applying the technique of averaging-out noise and increasing the sample size, the performance of the methods can be improved. The proposed jittering methods are robust to varying patterns of error term.

4.3.4 The Estimation of Variances

The estimated standard errors and corresponding standard deviations of 1000 estimators obtained using proposed methods (BJ, TJ, ABJ50, ATJ50) are compared with those obtained from

$$Q_{y^c}(\tau|\boldsymbol{x}) = \boldsymbol{x}'\boldsymbol{\gamma}_\tau, \tag{4.12}$$

where $\boldsymbol{\gamma}_\tau$ denotes the QR coefficient based on continuous response y^c before the truncation.

Table 4.7: Standard deviations ($s.d.$) of 1000 estimates of quantile regression parameters $\beta_{\tau 0}$, $\beta_{\tau 1}$ and $\beta_{\tau 2}$ and 1000 sample averages of estimated asymptotic standard errors ($s.e.$) based on truncated data by using different quantile regression methods and Standard deviations ($s.d.^*$) of 1000 estimates of quantile regression parameters $\gamma_{\tau 0}$, $\gamma_{\tau 1}$ and $\gamma_{\tau 2}$ based on original continuous data are reported. Quantile levels $\tau = (0.1, 0.25, 0.5, 0.75, 0.9)$, sample size $n = 200$.

		$\beta_{\tau 0}$			$\beta_{\tau 1}$			$\beta_{\tau 2}$		
τ	Method	$s.d.$	$s.e.$	$s.d.^*$	$s.d.$	$s.e.$	$s.d.^*$	$s.d.$	$s.e.$	$s.d.^*$
Case 1. $\varepsilon \sim \mathcal{N}(0,1)$										
0.1	TJ	0.4787	0.4995	0.4674	0.0464	0.0489	0.0465	0.3315	0.3645	0.3145
	ATJ50	0.3984	0.4245	0.4255	0.0365	0.0393	0.0398	0.2985	0.2985	0.2998
	BJ	0.4776	0.4994	0.4564	0.0451	0.0473	0.0440	0.3214	0.3524	0.2834
	ABJ50	0.3842	0.4137	0.4175	0.0356	0.0392	0.0395	0.2832	0.2832	0.2979
0.25	TJ	0.3831	0.3898	0.3621	0.0367	0.0381	0.0356	0.2567	0.2791	0.2655
	ATJ50	0.3326	0.3376	0.3468	0.0311	0.0324	0.0314	0.2168	0.2354	0.2355
	BJ	0.3803	0.3833	0.3529	0.0359	0.0370	0.0326	0.2528	0.2761	0.2387
	ABJ50	0.3225	0.3280	0.3344	0.0306	0.0311	0.0324	0.2153	0.2222	0.2285
0.50	TJ	0.3622	0.3611	0.3674	0.0334	0.0311	0.0318	0.2321	0.2398	0.2202

Continued on next page

Table 4.7 – continued from previous page

τ	Method	$\beta_{\tau 0}$			$\beta_{\tau 1}$			$\beta_{\tau 2}$		
		s.d.	*s.e.*	*s.d.**	*s.d.*	*s.e.*	*s.d.**	*s.d.*	*s.e.*	*s.d.**
	ATJ50	0.3311	0.3122	0.3311	0.0314	0.0299	0.0314	0.2014	0.2099	0.2100
	BJ	0.3521	0.3529	0.3225	0.0322	0.0336	0.0308	0.2369	0.2469	0.2329
	ABJ50	0.3213	0.3036	0.3317	0.0307	0.0289	0.0308	0.2011	0.2133	0.2168
0.75	TJ	0.3975	0.3934	0.3516	0.0367	0.0365	0.0325	0.2115	0.2109	0.2280
	ATJ50	0.3198	0.3256	0.3357	0.0311	0.0321	0.0333	0.2401	0.2398	0.2410
	BJ	0.3987	0.3831	0.3639	0.0375	0.0365	0.0345	0.2542	0.2612	0.2612
	ABJ50	0.3182	0.3267	0.3462	0.0302	0.0311	0.0329	0.2390	0.2260	0.2365
0.9	TJ	0.4067	0.4645	0.4166	0.0446	0.0456	0.0416	0.3289	0.3298	0.2825
	ATJ50	0.4067	0.4044	0.4156	0.0398	0.0399	0.4001	0.2784	0.2784	0.2924
	BJ	0.4654	0.4630	0.4167	0.0441	0.0444	0.0409	0.3269	0.3282	0.2817
	ABJ50	0.4045	0.4011	0.4198	0.0388	0.0389	0.4000	0.2732	0.2737	0.2916
Case 2. $\varepsilon \sim$ *Student's* t_2										
0.10	TJ	0.9994	1.0221	1.0147	0.0946	0.0967	0.1001	0.6376	0.6451	0.6589
	ATJ50	0.9302	0.9265	0.9545	0.0895	0.0902	0.0889	0.6187	0.6199	0.6275

Continued on next page

Table 4.7 – continued from previous page

		$\beta_{\tau 0}$			$\beta_{\tau 1}$			$\beta_{\tau 2}$		
τ	Method	$s.d.$	$s.e.$	$s.d.^*$	$s.d.$	$s.e.$	$s.d.^*$	$s.d.$	$s.e.$	$s.d.^*$
	BJ	0.9987	1.0230	1.0167	0.0934	0.0958	0.1005	0.6364	0.6295	0.6590
	ABJ50	0.9299	0.9158	0.9676	0.0880	0.0894	0.0873	0.6140	0.6136	0.6298
0.25	TJ	0.5201	0.5258	0.5179	0.0482	0.0478	0.0493	0.3357	0.3532	0.3231
	ATJ50	0.4778	0.4797	0.4698	0.0467	0.0459	0.0478	0.2786	0.2998	0.2898
	BJ	0.5196	0.5236	0.5145	0.0504	0.0518	0.0497	0.3324	0.3569	0.3198
	ABJ50	0.4765	0.4784	0.4675	0.0460	0.0458	0.0469	0.2776	0.2963	0.3133
0.50	TJ	0.3891	0.4202	0.3667	0.0378	0.0365	0.0376	0.2465	0.2586	0.2534
	ATJ50	0.3465	0.3589	0.3564	0.0354	0.0358	0.0365	0.2456	0.2587	0.2515
	BJ	0.3881	0.4162	0.3656	0.0388	0.0398	0.0362	0.3016	0.2932	0.2857
	ABJ50	0.3432	0.3690	0.3636	0.0348	0.0350	0.0363	0.2434	0.2589	0.2511
0.75	TJ	0.5376	0.5257	0.5198	0.0510	0.0512	0.0489	0.3133	0.3323	0.3411
	ATJ50	0.4578	0.4768	0.4689	0.0445	0.0455	0.0445	0.3488	0.3341	0.3321
	BJ	0.5365	0.5230	0.5180	0.0508	0.0504	0.0494	0.3123	0.3317	0.3424
	ABJ50	0.4539	0.4726	0.4728	0.0434	0.0449	0.0453	0.3457	0.3337	0.3309

Continued on next page

Table 4.7 – continued from previous page

		$\beta_{\tau 0}$			$\beta_{\tau 1}$			$\beta_{\tau 2}$		
τ	Method	*s.d.*	*s.e.*	*s.d.**	*s.d.*	*s.e.*	*s.d.**	*s.d.*	*s.e.*	*s.d.**
0.90	TJ	0.9877	1.0099	1.0044	0.0954	0.0989	0.0995	0.7389	0.7088	0.6997
	ATJ50	0.9301	1.0577	1.0266	0.0943	0.0952	0.0997	0.6834	0.7056	0.7315
	BJ	0.9854	1.0103	0.9912	0.0943	0.0978	0.0924	0.7367	0.6931	0.6993
	ABJ50	0.9675	1.0280	1.0172	0.0932	0.0982	0.0982	0.6955	0.7004	0.7315
Case 3. $\varepsilon \sim \mathcal{X}_3^2$										
0.10	TJ	0.4653	0.4401	0.3735	0.0428	0.0408	0.0326	0.3000	0.3045	0.2478
	ATJ50	0.3611	0.3561	0.3434	0.0325	0.0326	0.0163	0.2336	0.2405	0.2396
	BJ	0.4661	0.4407	0.3756	0.0437	0.0418	0.0332	0.3004	0.3059	0.2499
	ABJ50	0.3617	0.3563	0.3446	0.0338	0.0335	0.0348	0.2340	0.2410	0.2381
0.25	TJ	0.4846	0.4869	0.4448	0.0451	0.0455	0.0422	0.3298	0.3436	0.3024
	ATJ50	0.4156	0.4410	0.4448	0.0378	0.0416	0.0422	0.2796	0.3105	0.3024
	BJ	0.4858	0.4874	0.4448	0.0455	0.0464	0.0422	0.3308	0.3452	0.3024
	ABJ50	0.4177	0.4416	0.4448	0.0389	0.0422	0.0422	0.2794	0.3117	0.3024
0.50	TJ	0.6977	0.7078	0.6729	0.0667	0.0678	0.0641	0.4661	0.4843	0.4640

Continued on next page

Table 4.7 – continued from previous page

τ	Method	$\beta_{\tau 0}$			$\beta_{\tau 1}$			$\beta_{\tau 2}$		
		s.d.	*s.e.*	*s.d.**	*s.d.*	*s.e.*	*s.d.**	*s.d.*	*s.e.*	*s.d.**
	ATJ50	0.6534	0.6661	0.6729	0.0622	0.0638	0.0641	0.4403	0.4689	0.4640
	BJ	0.6962	0.7062	0.6729	0.0657	0.0673	0.0641	0.4568	0.4915	0.4640
	ABJ50	0.6568	0.6614	0.6729	0.0628	0.0633	0.0641	0.4190	0.4614	0.4640
0.75	TJ	1.0827	1.1151	1.0739	0.1031	0.1064	0.1024	0.7083	0.7510	0.7550
	ATJ50	1.0456	1.0731	1.0952	0.0985	0.1026	0.1027	0.7319	0.7580	0.7557
	BJ	1.1452	1.1218	1.0639	0.0998	0.1074	0.1074	0.7303	0.7679	0.7399
	ABJ50	1.0175	1.0862	1.0733	0.0967	0.1033	0.1019	0.7135	0.7536	0.7400
0.90	TJ	1.8135	1.8885	1.8229	0.1721	0.1833	0.1725	1.1664	1.2163	1.1946
	ATJ50	1.6323	1.8818	1.7098	0.1654	0.1801	0.1614	1.2403	1.1657	1.2897
	BJ	1.7794	1.9720	1.7719	0.1695	0.1981	0.1704	1.1603	1.1852	1.1668
	ABJ50	1.7713	1.8964	1.7585	0.1702	0.1743	0.1691	1.1982	1.1438	1.2514

Table 4.7 and Table 5.7A report the estimated standard errors and sample standard deviations of 1000 estimates using proposed methods based on truncated data, and sample standard deviations of estimators based on original continuous data, denoted by $s.d.^*$, when errors follows normal, student's t, Chi-squared distributions, respectively with sample size $n = (200, 500)$. First of all, we compare the estimated standard errors and sample standard deviation of QR estimators. Overall, the table finds that the estimated standard errors of estimates using proposed methods (BJ, TJ, ABJ50, ATJ50) are very close to their corresponding sample standard deviations regardless of error distributions. The estimated standard error obtained from the sandwich estimator of covariance matrix is reliable across small sample size. It is also worth-noting that when error follows t distribution, the sample standard deviations of estimates using all proposed methods are much larger on extreme quantiles of data, say when $\tau = (0.1, 0.9)$. Such a U-shaped relationship of quantile level and variablity of estimates is caused by heavy tails of t distribution with small degree of freedom. Similarly, large sample standard deviations of estimates are also found on upper quantiles of data with Chi-squared distributed errors, as the Chi-squared distribution is notably right-skewed. Despite the challenges posed by heavy tailed and large error values from t and Chi-squared distribution, the proposed methods is still able to offer reliable estimated variability of estimator with small sample size. The estimated standard errors and sample deviations of estimates using average jittering methods (ABJ50, ATJ50) are slightly smaller than using simple jittering methods (BJ, TJ). The results also reveal that the the estimated standard error of estimates based on truncated data is close to that of estimates based on original continuous data in most of cases. Table 5.7A shows clearly that the sample standard deviation and estimated standard error of estimators using all methods are getting lower, particularly for those on the lower and upper quantiles of t distribution and upper quantiles of Chi-squared distribution.

Table 4.8: Converage probabilities of 1000 estimates of QR parameters $\beta_{\tau 0}$, $\beta_{\tau 1}$ and $\beta_{\tau 2}$ are reported for different QR methods TJ, ATJ50, BJ, and ABJ50 on the $\tau = (0.1, 0.25, 0.5, 0.75, 0.9)^{th}$ quantile of data. Nominal levels $\alpha = (0.01, 0.05, 0.1)$ and sample size $n = 200$ are used.

		$\beta_{\tau 0}$			$\beta_{\tau 1}$			$\beta_{\tau 2}$		
τ	Method	$P_{0.99}$	$P_{0.95}$	$P_{0.90}$	$P_{0.99}$	$P_{0.95}$	$P_{0.90}$	$P_{0.99}$	$P_{0.95}$	$P_{0.90}$
Case 1. $\varepsilon \sim \mathcal{N}(0,1)$										
0.10	TJ	0.983	0.945	0.912	0.981	0.934	0.889	0.993	0.964	0.889
	ATJ50	0.982	0.936	0.911	0.985	0.944	0.913	0.996	0.961	0.891
	BJ	0.982	0.940	0.888	0.990	0.950	0.902	0.988	0.942	0.906
	ABJ50	0.984	0.934	0.906	0.988	0.950	0.912	0.966	0.936	0.884
0.25	TJ	0.986	0.946	0.905	0.991	0.956	0.906	0.985	0.940	0.895
	ATJ50	0.994	0.958	0.910	0.987	0.944	0.901	0.995	0.954	0.906
	BJ	0.980	0.940	0.900	0.985	0.940	0.895	0.985	0.960	0.915
	ABJ50	0.990	0.955	0.905	0.985	0.945	0.885	0.990	0.945	0.915
0.50	TJ	0.987	0.946	0.909	0.991	0.956	0.911	0.990	0.950	0.902
	ATJ50	0.986	0.956	0.898	0.988	0.950	0.911	0.992	0.945	0.910
	BJ	0.985	0.940	0.910	0.990	0.955	0.910	0.985	0.940	0.910
	ABJ50	0.980	0.955	0.885	0.980	0.950	0.910	0.990	0.935	0.900
0.75	TJ	0.989	0.938	0.899	0.991	0.948	0.890	0.985	0.944	0.890
	ATJ50	0.989	0.953	0.909	0.990	0.952	0.908	0.983	0.944	0.894
	BJ	0.990	0.940	0.898	0.981	0.945	0.891	0.987	0.933	0.888
	ABJ50	0.985	0.950	0.905	0.990	0.945	0.910	0.985	0.940	0.886
0.90	TJ	0.990	0.957	0.906	0.995	0.956	0.910	0.985	0.935	0.890
	ATJ50	0.987	0.955	0.900	0.994	0.955	0.905	0.995	0.945	0.899
	BJ	0.980	0.945	0.903	0.990	0.945	0.885	0.985	0.945	0.895

Continued on next page

		$\beta_{\tau 0}$			$\beta_{\tau 1}$			$\beta_{\tau 2}$		
τ	Method	$P_{0.99}$	$P_{0.95}$	$P_{0.90}$	$P_{0.99}$	$P_{0.95}$	$P_{0.90}$	$P_{0.99}$	$P_{0.95}$	$P_{0.90}$
	ABJ50	0.985	0.955	0.895	0.990	0.935	0.900	0.980	0.940	0.885
Case 2. $\varepsilon \sim \mathcal{T}_2$										
0.10	TJ	0.982	0.940	0.890	0.990	0.939	0.893	0.984	0.941	0.889
	ATJ50	0.984	0.941	0.889	0.992	0.941	0.904	0.994	0.945	0.891
	BJ	0.979	0.939	0.889	0.990	0.945	0.892	0.982	0.943	0.892
	ABJ50	0.988	0.940	0.888	0.990	0.943	0.912	0.980	0.935	0.895
0.25	TJ	0.983	0.944	0.982	0.988	0.939	0.891	0.986	0.954	0.905
	ATJ50	0.983	0.948	0.911	0.982	0.953	0.883	0.992	0.953	0.907
	BJ	0.987	0.942	0.893	0.990	0.942	0.890	0.985	0.956	0.910
	ABJ50	0.986	0.950	0.912	0.983	0.950	0.885	0.993	0.956	0.911
0.50	TJ	0.986	0.954	0.903	0.984	0.947	0.886	0.986	0.941	0.902
	ATJ50	0.984	0.958	0.904	0.986	0.942	0.883	0.984	0.941	0.901
	BJ	0.991	0.960	0.905	0.995	0.960	0.894	0.987	0.935	0.900
	ABJ50	0.985	0.956	0.903	0.984	0.945	0.887	0.988	0.963	0.900
0.75	TJ	0.989	0.951	0.892	0.982	0.943	0.905	0.991	0.956	0.906
	ATJ50	0.986	0.942	0.904	0.988	0.958	0.913	0.987	0.956	0.893
	BJ	0.990	0.950	0.887	0.983	0.938	0.905	0.990	0.960	0.913
	ABJ50	0.984	0.935	0.906	0.985	0.963	0.916	0.985	0.960	0.888
0.90	TJ	0.992	0.954	0.892	0.986	0.944	0.912	0.982	0.939	0.895
	ATJ50	0.991	0.959	0.894	0.993	0.956	0.909	0.987	0.938	0.885
	BJ	0.995	0.960	0.907	0.993	0.961	0.915	0.984	0.939	0.900
	ABJ50	0.995	0.955	0.894	0.995	0.956	0.911	0.983	0.956	0.920
Case 3. $\varepsilon \sim \mathcal{T}_2$										

Continued on next page

Table 4.8 – continued from previous page

τ	Method	$\beta_{\tau 0}$			$\beta_{\tau 1}$			$\beta_{\tau 2}$		
		$P_{0.99}$	$P_{0.95}$	$P_{0.90}$	$P_{0.99}$	$P_{0.95}$	$P_{0.90}$	$P_{0.99}$	$P_{0.95}$	$P_{0.90}$
0.10	TJ	0.983	0.950	0.887	0.991	0.956	0.894	0.990	0.957	0.902
	ATJ50	0.991	0.938	0.892	0.989	0.948	0.895	0.991	0.954	0.896
	BJ	0.981	0.950	0.885	0.995	0.960	0.884	0.990	0.964	0.905
	ABJ50	0.990	0.936	0.888	0.990	0.955	0.896	0.993	0.955	0.895
0.25	TJ	0.987	0.935	0.890	0.985	0.940	0.889	0.986	0.947	0.905
	ATJ50	0.985	0.952	0.891	0.992	0.953	0.910	0.991	0.946	0.909
	BJ	0.983	0.931	0.886	0.983	0.936	0.885	0.982	0.948	0.903
	ABJ50	0.986	0.954	0.888	0.994	0.962	0.914	0.992	0.944	0.912
0.50	TJ	0.984	0.941	0.894	0.985	0.945	0.897	0.985	0.943	0.893
	ATJ50	0.984	0.935	0.892	0.983	0.938	0.893	0.987	0.943	0.891
	BJ	0.982	0.936	0.893	0.983	0.943	0.889	0.982	0.942	0.899
	ABJ50	0.988	0.935	0.884	0.988	0.938	0.885	0.988	0.949	0.907
0.75	TJ	0.992	0.957	0.911	0.991	0.954	0.909	0.995	0.956	0.911
	ATJ50	0.994	0.948	0.914	0.992	0.963	0.916	0.987	0.964	0.914
	BJ	0.993	0.968	0.919	0.995	0.965	0.918	0.995	0.962	0.914
	ABJ50	0.993	0.960	0.914	0.992	0.959	0.918	0.990	0.967	0.914
0.90	TJ	0.993	0.948	0.909	0.995	0.961	0.909	0.984	0.950	0.894
	ATJ50	0.996	0.964	0.924	0.998	0.966	0.928	0.988	0.952	0.890
	BJ	0.990	0.957	0.918	0.989	0.962	0.914	0.984	0.937	0.892
	ABJ50	0.998	0.980	0.948	0.996	0.954	0.916	0.988	0.956	0.904

Table 4.8 and Table 4.8A reports the coverage probabilities of Wald type confidence intervals using different QR methods (TJ, ATJ50, BJ, and ABJ50) on the $\tau = (0.1, 0.25, 0.5, 0.75, 0.9)^{th}$ quantiles with nominal levels $\alpha = (0.99, 0.95, 0.90)$, re-

spectively and sample size of $n = (200, 500)$. The observed coverage probabilities of estimates are close to their corresponding nominal levels $(0.99, 0.95, 0.90)$ when small sample size $n = 200$ is used and getting enhanced when sample size increased to $n = 500$.

4.3.5 Performance of Statistical Tests

This section evaluates the performance of the Wald-type test based on the proposed methods. Two null hypotheses $H_0 : \beta_{\tau 2} = 0$ and $H_0 : \beta_{\tau 1} = 0$ against complete alternatives are tested.

Table 4.9: Rejection rates of 1000 simulation runs on $H_0 : \beta_{\tau 2} = 0$ using different QR methods. Levels of significance $\alpha = (0.01, 0.05, 0.1)$ and sample size $n = (200, 500)$.

		$n = 200$			$n = 500$		
τ	Method	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$
Case 1. $\varepsilon \sim \mathcal{N}(0, 1)$							
0.1	TJ	0.008	0.043	0.109	0.014	0.048	0.094
	ATJ50	0.010	0.065	0.105	0.012	0.057	0.098
	BJ	0.012	0.056	0.096	0.015	0.055	0.095
	ABJ50	0.014	0.054	0.116	0.015	0.050	0.105
0.25	TJ	0.011	0.043	0.093	0.011	0.054	0.102
	ATJ50	0.010	0.058	0.095	0.012	0.054	0.106
	BJ	0.015	0.040	0.085	0.010	0.048	0.096
	ABJ50	0.010	0.055	0.085	0.015	0.055	0.100
0.50	TJ	0.012	0.057	0.101	0.011	0.051	0.105
	ATJ50	0.010	0.051	0.101	0.012	0.057	0.104
	BJ	0.015	0.065	0.105	0.015	0.045	0.110
	ABJ50	0.010	0.060	0.095	0.015	0.055	0.105

Continued on next page

Table 4.9 – continued from previous page

		$n = 200$			$n = 500$		
τ	Method	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$
0.75	TJ	0.015	0.048	0.105	0.010	0.049	0.099
	ATJ50	0.013	0.065	0.110	0.011	0.047	0.097
	BJ	0.017	0.057	0.109	0.010	0.050	0.100
	ABJ50	0.015	0.056	0.105	0.011	0.045	0.095
0.9	TJ	0.012	0.056	0.107	0.012	0.056	0.104
	ATJ50	0.010	0.053	0.105	0.012	0.052	0.103
	BJ	0.015	0.055	0.111	0.013	0.046	0.094
	ABJ50	0.010	0.065	0.110	0.015	0.053	0.107
Case 2. $\varepsilon \sim \mathcal{T}_2$							
0.10	TJ	0.015	0.058	0.108	0.013	0.047	0.095
	ATJ50	0.020	0.044	0.089	0.011	0.051	0.100
	BJ	0.017	0.061	0.109	0.011	0.052	0.102
	ABJ50	0.018	0.057	0.105	0.013	0.052	0.098
0.25	TJ	0.013	0.056	0.094	0.013	0.058	0.096
	ATJ50	0.007	0.044	0.093	0.010	0.046	0.095
	BJ	0.014	0.050	0.092	0.015	0.055	0.100
	ABJ50	0.005	0.043	0.091	0.009	0.047	0.093
0.50	TJ	0.020	0.059	0.107	0.011	0.052	0.101
	ATJ50	0.018	0.058	0.101	0.013	0.054	0.095
	BJ	0.020	0.057	0.110	0.015	0.055	0.100
	ABJ50	0.012	0.045	0.100	0.015	0.053	0.097
0.75	TJ	0.011	0.048	0.092	0.014	0.044	0.098
	ATJ50	0.013	0.042	0.117	0.008	0.047	0.100

Continued on next page

		$n = 200$			$n = 500$		
τ	Method	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$
	BJ	0.017	0.046	0.090	0.008	0.046	0.105
	ABJ50	0.018	0.040	0.112	0.011	0.048	0.097
0.90	TJ	0.014	0.056	0.115	0.010	0.050	0.096
	ATJ50	0.013	0.062	0.115	0.010	0.050	0.097
	BJ	0.016	0.061	0.101	0.008	0.050	0.096
	ABJ50	0.017	0.045	0.080	0.012	0.050	0.096
Case 3. $\varepsilon \sim \mathcal{X}_3^2$							
0.10	TJ	0.010	0.043	0.098	0.008	0.050	0.103
	ATJ50	0.009	0.046	0.106	0.014	0.048	0.097
	BJ	0.010	0.036	0.095	0.008	0.048	0.102
	ABJ50	0.007	0.045	0.105	0.013	0.054	0.104
0.25	TJ	0.014	0.053	0.092	0.013	0.053	0.104
	ATJ50	0.009	0.054	0.091	0.013	0.045	0.103
	BJ	0.015	0.052	0.097	0.010	0.050	0.099
	ABJ50	0.008	0.052	0.092	0.012	0.054	0.104
0.50	TJ	0.015	0.057	0.107	0.014	0.048	0.093
	ATJ50	0.012	0.062	0.111	0.012	0.052	0.102
	BJ	0.015	0.060	0.100	0.014	0.053	0.099
	ABJ50	0.012	0.054	0.091	0.010	0.049	0.103
0.75	TJ	0.006	0.046	0.090	0.009	0.047	0.096
	ATJ50	0.013	0.038	0.089	0.008	0.047	0.104
	BJ	0.007	0.039	0.089	0.009	0.046	0.102
	ABJ50	0.010	0.039	0.088	0.008	0.048	0.098

Continued on next page

Table 4.9 – continued from previous page

τ	Method	$n = 200$			$n = 500$		
		$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$
0.90	TJ	0.016	0.051	0.106	0.010	0.049	0.010
	ATJ50	0.012	0.048	0.110	0.007	0.046	0.098
	BJ	0.013	0.063	0.108	0.009	0.053	0.103
	ABJ50	0.012	0.044	0.096	0.008	0.047	0.095

Table 4.9 reports the percentages of 1000 tests associated with all methods (BJ, TJ, ABJ50,ATJ50) to falsely reject null hypothesis $H_0 : \beta_{\tau 2} = 0$ at nominal levels $\alpha = (0.01, 0.05, 0.1)$ with the sample sizes $n = (200, 500)$. The rejection rates of the test associated with all methods are close to the corresponding nominal levels when small size $n = 200$ is used. After increasing the sample size to $n = 500$, the rejection rates of the test associated with all methods are enhanced and they are closer to corresponding nominal levels.

Next, to check the power of the Wald tests, we set different effect sizes $\beta_1 = (0.2, 0.4, 0.6, 0.8, 1.0)$.

Table 4.10: Power of $S = 1000$ hypotheses test (i.e., $1 - \beta$) as related to the QR parameters $\beta_{\tau 1}$ are reported by for different QR methods TJ, ATJ50, BJ, and ABJ50. Significance level $\alpha = 0.05$ and sample size $n = 200$.

τ	Method	$\beta_{\tau 1} = 0.2$	$\beta_{\tau 1} = 0.4$	$\beta_{\tau 1} = 0.6$	$\beta_{\tau 1} = 0.8$	$\beta_{\tau 1} = 1.0$
Case 1. $\varepsilon \sim \mathcal{N}(0, 1)$						
0.10	TJ	0.985	0.960	1.000	1.000	1.000
	ATJ50	1.000	0.995	1.000	1.000	1.000
	BJ	0.985	0.990	1.000	1.000	1.000

Continued on next page

Table 4.10 – continued from previous page

τ	Method	$\beta_{\tau 1} = 0.2$	$\beta_{\tau 1} = 0.4$	$\beta_{\tau 1} = 0.6$	$\beta_{\tau 1} = 0.8$	$\beta_{\tau 1} = 1.0$
	ABJ50	1.000	1.000	1.000	1.000	1.000
0.25	TJ	1.000	1.000	1.000	1.000	1.000
	ATJ50	1.000	1.000	1.000	1.000	1.000
	BJ	1.000	1.000	1.000	1.000	1.000
	ABJ50	1.000	1.000	1.000	1.000	1.000
0.50	TJ	1.000	1.000	1.000	1.000	1.000
	ATJ50	1.000	1.000	1.000	1.000	1.000
	BJ	1.000	1.000	1.000	1.000	1.000
	ABJ50	1.000	1.000	1.000	1.000	1.000
0.75	TJ	1.000	1.000	1.000	1.000	1.000
	ATJ50	1.000	1.000	1.000	1.000	1.000
	BJ	1.000	1.000	1.000	1.000	1.000
	ABJ50	1.000	1.000	1.000	1.000	1.000
0.90	TJ	0.989	1.000	1.000	1.000	1.000
	ATJ50	1.000	1.000	1.000	1.000	1.000
	BJ	0.990	1.000	1.000	1.000	1.000
	ABJ50	0.999	1.000	1.000	1.000	1.000
Case 2. $\varepsilon \sim \mathcal{T}_2$						
0.10	TJ	0.538	0.943	1.000	1.000	1.000
	ATJ50	0.623	0.956	1.000	1.000	1.000
	BJ	0.543	0.945	1.000	1.000	1.000
	ABJ50	0.632	0.958	1.000	1.000	1.000
0.25	TJ	0.973	1.000	1.000	1.000	1.000

Continued on next page

Table 4.10 – continued from previous page

τ	Method	$\beta_{\tau 1} = 0.2$	$\beta_{\tau 1} = 0.4$	$\beta_{\tau 1} = 0.6$	$\beta_{\tau 1} = 0.8$	$\beta_{\tau 1} = 1.0$
	ATJ50	0.999	1.000	1.000	1.000	1.000
	BJ	0.977	1.000	1.000	1.000	1.000
	ABJ50	0.999	1.000	1.000	1.000	1.000
0.50	TJ	0.922	1.000	1.000	1.000	1.000
	ATJ50	1,000	1.000	1.000	1.000	1.000
	BJ	0.933	1.000	1.000	1.000	1.000
	ABJ50	1.000	1.000	1.000	1.000	1.000
0.75	TJ	0.932	1.000	1.000	1.000	1.000
	ATJ50	0.974	1.000	1.000	1.000	1.000
	BJ	0.921	1.000	1.000	1.000	1.000
	ABJ50	0.978	1.000	1.000	1.000	1.000
0.90	TJ	0.542	0.932	0.975	0.996	1.000
	ATJ50	0.569	0.948	0.986	1.000	1.000
	BJ	0.542	0.934	0.977	0.997	1.000
	ABJ50	0.571	0.951	0.987	1.000	1.000
Case 3. $\varepsilon \sim \mathcal{X}_3^2$						
0.10	TJ	1.000	1.000	1.000	1.000	1.000
	ATJ50	1.000	1.000	1.000	1.000	1.000
	BJ	0.998	1.000	1.000	1.000	1.000
	ABJ50	1.000	1.000	1.000	1.000	1.000
0.25	TJ	1.000	1.000	1.000	1.000	1.000
	ATJ50	1.000	1.000	1.000	1.000	1.000
	BJ	0.999	1.000	1.000	1.000	1.000

Continued on next page

Table 4.10 – continued from previous page

τ	Method	$\beta_{\tau 1} = 0.2$	$\beta_{\tau 1} = 0.4$	$\beta_{\tau 1} = 0.6$	$\beta_{\tau 1} = 0.8$	$\beta_{\tau 1} = 1.0$
	ABJ50	1.000	1.000	1.000	1.000	1.000
0.50	TJ	0.973	1.000	1.000	1.000	1.000
	ATJ50	0.990	1.000	1.000	1.000	1.000
	BJ	0.978	1.000	1.000	1.000	1.000
	ABJ50	0.992	1.000	1.000	1.000	1.000
0.75	TJ	0.613	0.983	1.000	1.000	1.000
	ATJ50	0.627	1.000	1.000	1.000	1.000
	BJ	0.604	0.988	1.000	1.000	1.000
	ABJ50	0.639	1.000	1.000	1.000	1.000
0.90	TJ	0.335	0.683	0.934	0.977	0.995
	ATJ50	0.358	0.737	0.938	0.983	1.000
	BJ	0.337	0.686	0.935	0.978	0.997
	ABJ50	0.364	0.734	0.943	0.979	1.000

Table 4.10 and Table 4.9A report the percentages of 1000 tests associated with all methods for correctly rejecting null hypothesis $\beta_{\tau 1} = 0$ when it is not true with samples sizes $n = (200, 500)$. For normal distributed error, the powers of test related to $\beta_{\tau 1}$ on five quantiles are greater than 80%. When error follows t distribution, the power of test associated with all methods become worse only on lower and upper quantiles of data due to the presence of outliers and extreme values. The probability mass in the tail is spread out over a wide range of values, making the test focusing on extreme quantile harder to detect an effect. Similarly, the powers of test on upper quantiles are getting lower since the test is influenced by the sknewness of Chi-squared distribution. After we increase the sample, Table 4.9A finds that the powers of test

related to the regression parameter are mostly greater than 80% on five quantiles of data. The only exception is that on upper quantile ($\tau = 0.9$) of data with Chi-squared distributed error when effect size equals 0.2.

In general, the performance of hypothesis testing related to a particular QR parameter is good.

4.4 Application: Fish Abundance and Commercial Fishing Activities

In this section we illustrate our methods of QR model for counts by conducting an analysis on 'fishing' data available in R package 'MASS' (Venables & Ripley, 2002), reported by Bailey et al. (2009). The primary intention to collect the fishing data is to examine the impact of commercial fishing in catching areas on the populations of deep-sea fish, particularly in the deeper water over time. It is important to study the fish abundance before and after the commercial activities that helps assess the ecosystem and manage fisheries sustainably. By fitting different regression models to the fishing abundance data across two periods, we illustrate the application of our statistical model in the discipline of fishery science.

4.4.1 Real Data and Explanations

Table 4.11: Overview of the Fishing Dataset

Variable	**Description**	**Type/Attributes**
totabund	Total number of fish per site.	Integer
meandepth	Mean water depth per site in meters.	Integer

Continued on next page

Table 4.11 – continued from previous page

Variable	Description	Type/Attributes
sweptarea	Adjusted area of the site in square meters.	Numeric
density	Foliage density index.	Numeric
site	Catch site identifier.	Categorical
year	Year of data collection (1977-2002).	Integer
period	Time period of data (0=1977-1989, 1=2000+).	Categorical

Table 4.11 gives an overview of the fishing data set. A total of 147 catch sites are researched to record the population of deep-sea fish (totabund), mean water depth (meandepth), adjusted site area in (sweptarea), foliage density index under water (density) from 1977 to 2002 (year). The commercial fishing activities began in the year of 2000. Therefore, period distinguishes between data collected where 0 = before 2000, 1 = after 2000. Bailey et al. (2009) initially studied it to understand the long term change of population of deep-water fish in the North East Atlantic. Hilbe (2014) then studied the data for the same purpose, using a generalized linear mixed model and proposed flexible mixture models. The results found that the fish abundance keeps falling from 800 to 2500 meters, even deeper than the maximum depth of legal commercial fishing (1600 meters). The findings indicate that the effect of fishery has already transmitted to very deep offshore areas, leading to concern of over-fishing and marine reverse management (Bailey et al., 2009). The ecological system is complex (Pepin et al., 2022), and area and under-water environment are expected to influence fish populations. We analyze the data using our proposed QR method. The purpose of this application is bi-fold. We illustrate how the proposed method are used and at

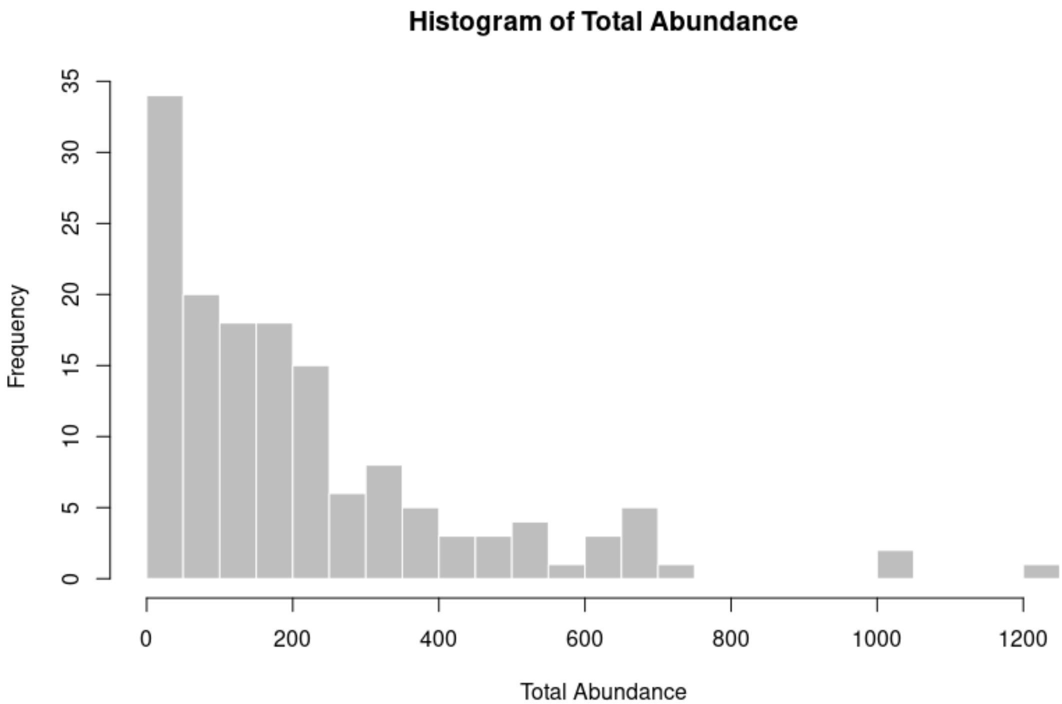


Figure 4.3: Histogram of fish populations for all 147 catching sites.

the same time try to see if the effect of commercial fishing activities varies over the distribution of fish abundance. Figure 4.3 displays the histogram of fish populations for 147 catching sites. The maximum fish abundance is 1230 while the lowest is only 2, showing a significant disparity. Also, the distribution of fish populations is seriously right-skewed, meaning that most of catching site have very low fish counts, and the majority of populations concentrates in only a few sites. The fishing data might contain outliers or extreme values for cites whose abundance exceeds 1000. Focusing on the change of fish abundance due to commercial fishing activities, Figure 4.4 includes the box plots for the total fish abundances, and abundance before and after commercial fishing. The group of box plots shows that the median and upper quantile of abundance during 1977-1989 is the highest. The wide range of upper and lower quantiles suggest a large variability within the data. Moreover, the median and upper quantile of abundance significantly dropped since the start of commercial fish-

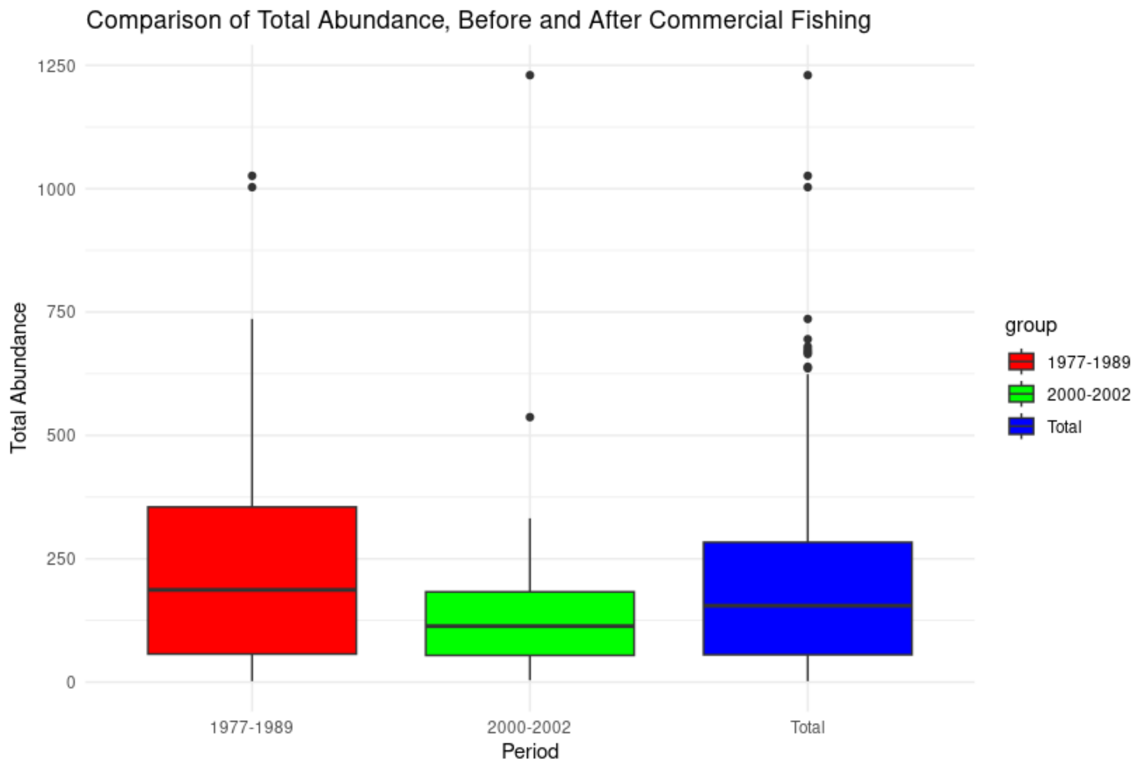


Figure 4.4: Box plots of fish abundance before and after commercial fishing activities.

ing. Overall, the total abundance exhibits high variability, long tail of distribution, and notable outliers present in total abundance notably. Taking both skewness and outliers into account, we expect that traditional mean regression model may provide limited information about impact of fishing activities.

4.4.2 Data Pre-processing and Regression Models

To comprehensively analyze the fish abundance before and after commercial fishing, we introduce interaction terms $Density \times Period$, $Meandepth \times Period$ and $Sweptarea \times Period$ into the regression models. Let $y_i, i = 1, ..., 147$, be the abundance of the i^{th} catching site and $\boldsymbol{x}_i = (Density_i, Meandepth_i, Sweptarea_i, Period_i,$ $Density_i \times Period_i, Meandepth \times Period_i, Sweptarea_i \times Period_i)$ be a vector of predictors. We fit QR model using Beta distributed noises u_i, where the QR response is $z_i = y_i + u_i$. According to the equivariance properties of QR model, logarithmic transformation is applied to the jittered response such that $T(z_i;\tau) = \log(z_i - Q_U(\tau))$ if $z_i - Q_U(\tau) \geq \varsigma$ or equals to $\log(\varsigma)$ otherwise. Also, since the value of adjusted area is too large, we scale it by converting the unit from square kilometer (km^2) to mega square kilometer (Mkm^2) (e.g., 1 Mkm^2 equals 1,000,000 km^2). The following QR model will be used to fit the fish abundance data

$$\begin{aligned} Q_{T(z_i;\tau)}(\tau|\boldsymbol{x}_i) =& \beta_{\tau 0} + \beta_{\tau 1} Density_i + \beta_{\tau 2} Meandepth_i + \beta_{\tau 3} Sweptarea_i + \beta_{\tau 4} Period_i \\ & + \beta_{\tau 5} Density_i \times Period_i + \beta_{\tau 6} Meandepth \times Period_i \\ & + \beta_{\tau 7} Sweptarea_i \times Period_i. \end{aligned} \tag{4.13}$$

In order to compare the results and better understand the effect of the commercial fishing, mean regression model, Poisson regression model and ordinary QR model are fit based on same predictors and interaction terms. QR models are fit on the $100\% \times \tau = (0.25, 0.5, 0.75)$-th percentile of fishing data.

4.4.3 Results and Interpretations

Table 4.12 and Table 4.13 report the estimated parameters, estimated standard errors, corresponding 95% confidence intervals and p-values of parameters with significance code (*) at nominal level 0.05, respectively by fitting mean regression model, Poisson regression model, proposed quantile regression model and ordinary QR model. First of all, the results obtained from mean regression and Poisson regression provide comparable estimated parameters. The negative $\hat{\beta}_2$ from both models indicate a negative effect of water depth on the mean value of fish population. Density of foliage plays an important role in growth of mean fish population. It is interesting that the estimated β_4 is positive. It implies that the start of commercial fishing does not directly lead to reduction of fish abundance. However, if we take interaction terms into account, negative estimated β_5, β_6 and β_7 shows that the start of commercial fishing activities in those catching sites negatively affects all catching sites. The estimated parameters suggests that the fish populations in catching sites with higher density of foliage and larger area are affected heavily, and the commercial fishing only slightly aggravates the decrease of fish in deep water. The estimates $\hat{\beta}_3(0.25) > \hat{\beta}_3(0.5) > \hat{\beta}_3(0.75) > 0$ are the three largest, revealing that density of foliage is still demanded the most for the fish abundance, but the demand decreases with the quantiles of fish abundance. Conversely, the need for aquatic area to support fish life increases with the quantiles of fish abundance as $0 < \hat{\beta}_3(0.25) < \hat{\beta}_3(0.5) < \hat{\beta}_3(0.75)$. The estimates $\hat{\beta}_2(0.25) \approx \hat{\beta}_3(0.5) \approx \hat{\beta}_3(0.75) \approx 0$ indicates the water depth almost has no effect on abundance. Moreover, the commercial fishing is harmful to more densely populated catching sites as $\hat{\beta}_4(0.75) < 0$. This might be because the fish population are likely concentrated in highly dense area, which makes the commercial fishing cause more decrease of fish abundance. Now we study the estimated parameters of interaction terms on different quantile. The estimates $\hat{\beta}_7(0.75) < \hat{\beta}_7(0.5) < \hat{\beta}_7(0.25) < 0$ implies that the impact of commercial fishing in large water area increases with fish abundance. On the other hand, the estimated parameter of $\beta_{\tau 6}$ roughly equals zero on three quantiles, indicating the impact of commercial fishing on deep-water fish is almost negligible, regardless of fish abundance. For the interaction between foliage density index and period, $\hat{\beta}_5(0.25) < \hat{\beta}_5(0.5) < 0$ suggest that those smaller and medium fish abundance in dense foliage are vulnerable to commercial fishing. Considering foliage as a part of aquatic system providing food source and shelters for fish productivity (Massicotte et al., 2015), fishing activities seem harmful to fish abundance. But an interesting finding is shown on the high quantile of abundance. The estimate $\beta_5(0.75) = 29.7746 > 0$ implies that the commercial fishing increasese the fish abundance at high quantile and in dense foliage water, challenge to our common assumption. This interesting phenomenon may be explained by theorem of resilience and transition under ecological change, which introduced resilience or aiding adaption to changes induced by environmental or social pressure (Woods et al., 2021). If we assume that the foliage growth are disturbed by commercial fishing, then Woods et al. (2021) showed that ecological system dependent on fisheries will be resilient or remain viable to coop with negative environmental change. Ecological disturbances somehow make the remaining water areas more stabilized after fishing activities, allowing large fish population to utilize the remaining habitat structure to flourish and grow. Therefore, consider all above, in a long term aspect we suggest to avoid commercial fishing in water with large area to prevent fish species from depopulation. Fishing activities can be held in dense foliage water with a good maintenance of under water plants, unless the initial fish population is sufficiently large. Choosing deep water for commercial fishing seems much safer for maintaining the fish abundance. Therefore, monitoring commercial fishing activities and implementing marine reverse management are actionable actions to stabilize the fish abundance.

Overall, given that the distribution of fish populations in 147 catching sites is right-skewed, the results from QR models are much more representative. Also according to our study, focusing on only mean fish population provides limited information. Our model provides a comprehensive overview of impact on fish population made by commercial fishing while giving a smaller standard errors.

Table 4.12: Estimated parameters (EP), their standard errors (SE), corresponding 95% confidence intervals (CI) and p-values at nominal level $\alpha = 0.05$ from fitting both the mean regression model and Poisson regression model.

	Mean Regression				Poisson Regression			
β_τ	EP	SE	CI	p-value	EP	SE	CI	p-value
β_0	5.1320	0.2878	(4.5626, 5.701)	< 0.001*	5.2860	0.0260	(5.2346, 5.3366)	< 0.001*
β_1	118.60	17.5201	(83.9159, 153.2144)	< 0.001*	88.2401	1.1260	(86.0350, 90.4485)	< 0.001*
β_2	-0.0005	0.0001	(-0.0007, -0.0004)	< 0.001*	-0.0005	0.0000	(-0.4929, -0.0005)	< 0.001*
β_3	8.0120	3.1921	(1.7003, 14.3227)	0.013*	8.4780	0.2892	(7.9066, 9.0403)	< 0.001*
β_4	0.2314	0.04230	(-0.6050, 1.0677)	0.585	-0.0089	0.0440	(-0.5168, 0.0773)	0.839
β_5	-16.2300	30.0001	(-75.5446, 43.0895)	0.589	-12.4500	1.6190	(-1.5629, -9.2822)	< 0.001*
β_6	-0.0001	0.0002	(-0.0002, -0.0004)	0.511	-0.0002	0.0000	(-0.0003, -0.0001)	< 0.001*
β_7	-8.9260	6.1790	(-21.1416, 3.2904)	0.150	-8.7690	0.8055	(-1.0356, -7.1986)	< 0.001*

Table 4.13: Estimated parameters (EP), their standard errors (SE), corresponding 95% confidence intervals (CI) and p-values at nominal level $\alpha = 0.05$ from fitting both the proposed quantile regression model and ordinary quantile regression model at three quantiles $\tau = (0.25, 0.50, 0.75)$.

		Proposed Quantile Regression				Ordinary Quantile Regression			
τ	β_τ	EP	SE	CI	p-value	EP	SE	CI	p-value
0.25	$\beta_{\tau 0}$	5.0764	0.0348	(5.0075, 5.1453)	< 0.001*	5.0751	0.4160	(4.2527, 5.8975)	< 0.001*
	$\beta_{\tau 1}$	112.7303	0.8676	(111.0149, 114.4458)	< 0.001*	112.8124	25.3330	(62.7276, 162.8972)	< 0.001*
	$\beta_{\tau 2}$	-0.0006	0.0000	(-0.0006, -0.0005)	< 0.001*	-0.0006	0.0001	(-0.0008, -0.0003)	< 0.001*
	$\beta_{\tau 3}$	7.6326	0.2799	(7.0791, 8.1862)	< 0.001*	7.6398	4.6143	(-1.4829, 16.7626)	0.100
	$\beta_{\tau 4}$	0.0564	0.0418	(-0.0263, 0.1391)	0.179	0.0591	0.6115	(-1.1499, 1.2681)	0.923
	$\beta_{\tau 5}$	-27.4575	1.0666	(-29.5665, -25.3487)	< 0.001*	-27.4473	43.3684	(-113.1891, 58.2944)	0.5278
	$\beta_{\tau 6}$	-0.0000	0.0000	(-0.0001, 0.0001)	0.435	-0.0000	0.0002	(-0.0005, 0.0004)	0.865
	$\beta_{\tau 7}$	-5.0724	1.366	(-7.7738, -2.3711)	< 0.001*	-5.0369	8.9315	(-22.6950, 12.6212)	0.573
0.50	$\beta_{\tau 0}$	5.2413	0.2677	(4.7120, 5.7705)	< 0.001*	5.2463	0.2215	(4.8084, 5.6842)	< 0.001*
	$\beta_{\tau 1}$	105.5414	10.6538	(84.4783, 126.6045)	< 0.001*	105.3466	13.4892	(78.6777, 132.0155)	< 0.001*
	$\beta_{\tau 2}$	-0.0005	0.0000	(-0.0006, -0.0005)	< 0.001*	-0.0005	0.0001	(-0.0006, -0.0003)	< 0.001*
	$\beta_{\tau 3}$	7.6711	1.7951	(4.1220, 11.2201)	< 0.001*	7.6316	2.4570	(2.7740, 12.4892)	0.0023

Continued on next page

Table 4.13 continued from previous page

τ	β_τ	Proposed Quantile Regression				Ordinary Quantile Regression			
		EP	SE	CI	p-value	EP	SE	CI	p-value
	$\beta_{\tau 4}$	0.5289	0.2772	(-0.0190, 1.0769)	0.062	0.5198	0.3256	(-0.1239, 1.1636)	0.113
	$\beta_{\tau 5}$	-39.1218	10.6911	(-60.2586, -17.9850)	< 0.001*	-38.8658	23.0926	(-84.5212, 6.7896)	0.095
	$\beta_{\tau 6}$	-0.0000	0.0000	(-0.00005, 0.00007)	0.659	-0.0000	0.0001	(-0.0002, 0.0002)	0.922
	$\beta_{\tau 7}$	-9.9233	2.5511	(-14.9669, -4.8797)	< 0.001*	-9.8604	4.7558	(-19.2629, -0.4579)	0.040*
0.75	$\beta_{\tau 0}$	5.2345	0.4251	(4.3942, 6.0749)	< 0.001*	5.2244	0.1470	(4.9338, 5.5150)	< 0.001*
	$\beta_{\tau 1}$	98.1531	15.6282	(67.2553, 129.0510)	< 0.001*	98.4594	8.9517	(80.7615, 116.1574)	< 0.001*
	$\beta_{\tau 2}$	-0.0006	0.0001	(-0.0007, -0.0005)	< 0.001*	-0.0006	0.0000	(-0.0007, -0.0005)	< 0.001*
	$\beta_{\tau 3}$	16.1530	2.0311	(12.1374, 20.1686)	< 0.001*	16.1526	1.6305	(12.9290, 19.3762)	< 0.001*
	$\beta_{\tau 4}$	-0.1610	0.4393	(-1.0295, 0.7075)	0.715	-0.1493	0.2161	(-0.5765, 0.2779)	0.491
	$\beta_{\tau 5}$	29.7746	16.2565	(-2.3653, 61.9145)	0.069	29.3029	15.3247	(-0.9948, 59.6006)	0.058
	$\beta_{\tau 6}$	0.0004	0.0001	(0.0003, 0.0006)	< 0.001*	0.0004	0.0001	(0.0003, 0.0006)	< 0.001*
	$\beta_{\tau 7}$	-17.3879	2.0762	(-21.4926, -13.2831)	<0.001	-17.3844	3.1560	(-23.6241, -11.1447)	< 0.001*

Chapter 5

Concluding Remarks

Traditional QR method is not directly applicable to count data due to discreteness and non-differentiability of sample objective function. Systematic bias and poor inference could occur, if QR is naively used to model count data. We are inspired by Machado and Silva (2005)'s uniform jittering method and proposed an optimized jittering method using non-negative continuous variable (Tweedie and Beta) to smooth the counts. The parameters of noise distribution can be pre-determined by utilizing the ALD. Technique of averaging noises out is applicable to this alternative jittering method, which further reduces the variability of QR estimate. The MCEM and MH algorithms can be used for maximizing the likelihood of the ALD, reducing the computational cost. QR estimator obtained by using this alternative jittering method is consistent and asymptotically normally distributed, facilitating the statistical inference on QR model. The simulations in Chapter 4 show that our QR method performs well. The variability of estimator is much smaller than uniform jittering method.

As a future work, we will discuss the asymptotic equivalence of hypothesis tests based on continuous data and truncated data. Two hypothesis tests are asymptotically equivalent if their difference in performance of rejection rate or power of the test vanishes as the sample size tends to infinity. Again we let continuous y^c and truncated $y = \text{trunc}(y^c)$ follow the same data generation process as in simulation **II**. Consider two QR models as follows,

$$\begin{aligned} &\text{Model 1: } Q_{T(z;\tau)}(\tau|\boldsymbol{x}) = \boldsymbol{x}'\boldsymbol{\beta}_\tau, \text{ and} \\ &\text{Model 2: } Q_{y^c}(\tau|\boldsymbol{x}_i) = \boldsymbol{x}'\boldsymbol{\gamma}_\tau, \end{aligned} \tag{5.1}$$

where $T(z_i;\tau) = z_i - Q_U(\tau)$ if $z_i - Q_U(\tau) > \varsigma$ and equal to ς otherwise, $\boldsymbol{\beta}_\tau$ and $\boldsymbol{\gamma}_\tau$ are true parameters of corresponding QR models at τ-th quantile. To evaluate the asymptotic equivalence of the tests, we examine the powers of the test related to $\beta_{\tau 1}$ and $\gamma_{\tau 1}$ by comparing the percentages for the two tests to reject the null hypotheses

$$\begin{aligned} &H_0 : \beta_{\tau 1} = 0 \quad \text{vs.} \quad H_1 : \beta_{\tau 1} \neq 0, \text{ and} \\ &H_0 : \gamma_{\tau 1} = 0 \quad \text{vs.} \quad H_1 : \gamma_{\tau 1} \neq 0, \end{aligned} \tag{5.2}$$

when the null hypothesis is true or not. The corresponding test statistics for j-th estimator of two models are

$$t^{\text{Model 1}} = \frac{\hat{\beta}_{\tau j} - 0}{\text{s.e.}(\beta_{\tau j})} \quad \text{and} \quad t^{\text{Model 2}} = \frac{\hat{\gamma}_{\tau j} - 0}{\text{s.e.}(\gamma_{\tau j})}, \tag{5.3}$$

respectively. The preliminary results observed from this simulation has been established. We will further investigate in the future for a theoretical framework to understand asymptotic equivalence of tests.

Moreover, we would like to further extend the proposed QR model to fit to zero-inflated data more accurately. Chapter 4 prompts some possible developments of our jittering approach. In simulation I, jittering performs worse when there are excessive zeros, even with a zero-inflation of 20%. Ling et al. (2022) proposed a two-part QR model with logistic regression to model the zero part. It seems interesting to mix our proposed QR model with logistic regression, which allows us to model zero part with logistic regression and non-zero part with proposed QR model. Moreover, in the empirical study, the fishing population is assumed to be independent of one another for illustration purpose. However, geographically if fishing sites share the same body of water system and environmental conditions or are monitored under same policies, the fish population in one site will possibly depend on one another. To enhance the applicability of alternative jittering to real data, we have to modify the proposed jittering method to adapt dependent counts. Corresponding model assumptions, estimation and inference will be derived in the future.

Bibliography

Andrews, D. (1986). Empirical process methods in econometrics. In R. F. Engle & D. McFadden (Eds.), *Handbook of econometrics* (1st ed., pp. 2247–2294, Vol. 4). Elsevier. https://EconPapers.repec.org/RePEc:eee:ecochp:4-37

Bailey, D., Collins, M., Gordon, J., Zuur, A., & Priede, I. (2009). Long-term changes in deep-water fish populations in the northeast atlantic: A deeper reaching effect of fisheries? *Proceedings. Biological sciences / The Royal Society*, *276*, 1965–9. https://doi.org/10.1098/rspb.2009.0098

Chen, C. W., Gerlach, R., Hwang, B. B., & McAleer, M. (2012). Forecasting value-at-risk using nonlinear regression quantiles and the intra-day range. *International Journal of Forecasting*, *28*(3), 557–574. https://doi.org/https://doi.org/10.1016/j.ijforecast.2011.12.004

Chen, C. (2005). Growth charts of body mass index (bmi) with quantile regression. *Proceedings of the 2005 International Conference on Algorithmic Mathematics and Computer Science, AMCS'05*, 114–120.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. L. Erlbaum Associates. https://books.google.ca/books?id=gA04ngAACAAJ

Davino, C., Furno, M., & Vistocco, D. (2013). Quantile regression: Theory and applications. *Quantile Regression: Theory and Applications*. https://doi.org/10.1002/9781118752685

De Backer, M., Ghouch, A., & Keilegom, I. (2020). Linear censored quantile regression: A novel minimum-distance approach. *Scandinavian Journal of Statistics*, *47*. https://doi.org/10.1111/sjos.12475

Ding, D. S., Patel, A. K., Singhania, R. R., Chen, C. W., & Dong, C. D. (2022). Effects of temperature and salinity on growth, metabolism and digestive enzymes synthesis of goniopora columna. *Biology*, *11*(3), 436. https://doi.org/10.3390/biology11030436

Dunn, P. K. (2022). *Tweedie: Evaluation of tweedie exponential family models* [R package version 2.3.5].

Dunn, P. K., & Smyth, G. K. (2005). Series evaluation of tweedie exponential dispersion models. *Statistics and Computing*, *15*(4), 267–280.

Efron, B. (1992). Poisson overdispersion estimates based on the method of asymmetric maximum likelihood. *Journal of the American Statistical Association*, *87*(417), 98–107. https://doi.org/10.1080/01621459.1992.10475180

Efron, B., & Tibshirani, R. J. (1998). *Introduction to the bootstrap*. Chapman & Hall/CRC.

Erdman, D., & Sinko, A. (2008). Zero-inflated poisson and zero-inflated negative binomial models using the countreg procedure.

Friederichs, P., & Hense, A. (2007). Statistical downscaling of extreme precipitation events using censored quantile regression. *Monthly Weather Review - MON WEATHER REV*, *135*, 2365–2378. https://doi.org/10.1175/MWR3403.1

Geraci, M. (2014). Linear quantile mixed models: The lqmm package for laplace quantile regression. *Journal of Statistical Software*, *57*(13), 1–29. https://doi.org/10.18637/jss.v057.i13

Geraci, M., & Bottai, M. (2006). Quantile regression for longitudinal data using the asymmetric Laplace distribution. *Biostatistics*, *8*(1), 140–154. https://doi.org/10.1093/biostatistics/kxj039

Geraci, M., & Farcomeni, A. (2022). Mid-quantile regression for discrete responses. *Statistical Methods in Medical Research*, *31*, 096228022110605. https://doi.org/10.1177/09622802211060525

Gujarati, D. N. (2003). *Basic econometrics* (International Edition). McGraw-Hill.

Hajdu, T., & Hajdu, G. (2014). Income and subjective well-being: How important is the methodology? *Hungarian Statistical Review*, *92*, 110–128.

Harding, M., & Lamarche, C. (2019). Penalized estimation of a quantile count model for panel data. *Annals of Economics and Statistics*, (134), 177–206. Retrieved April 24, 2024, from https://www.jstor.org/stable/10.15609/annaeconstat2009.134.0177

Hastings, W. K. (1970). Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, *57*(1), 97–109. https://doi.org/10.1093/biomet/57.1.97

Hilbe, J. M. (2014). *Modeling count data*. Cambridge University Press.

Horowitz, J. (1992). A smoothed maximum score estimator for the binary response model. *Econometrica*, *60*(3), 505–31. https://EconPapers.repec.org/RePEc:ecm:emetrp:v:60:y:1992:i:3:p:505-31

Jentsch, C., & Leucht, A. (2015). Bootstrapping sample quantiles of discrete data. *Annals of the Institute of Statistical Mathematics*, *68*. https://doi.org/10.1007/s10463-015-0503-3

Johnson, N. L., Kotz, S., & Balakrishnan, N. (1994 - 1995). Continuous univariate distributions (2nd ed.).

Jorgensen, B. (1987). Exponential dispersion models. *Journal of the Royal Statistical Society. Series B (Methodological)*, *49*(2), 127–162. Retrieved April 23, 2024, from http://www.jstor.org/stable/2345415

Kim, J., & Pollard, D. (1990). Cube root asymptotics. *The Annals of Statistics*, *18*(1), 191–219. Retrieved April 25, 2024, from http://www.jstor.org/stable/2241541

Kocherginsky, M., & He, X. (2007). Extensions of the markov chain marginal bootstrap [Silver Jubilee Issue Dedicated to Richard A. Johnson on his 70th birthday]. *Statistics & Probability Letters*, *77*(12), 1258–1268. https://doi.org/https://doi.org/10.1016/j.spl.2007.03.010

Koenker, R. (2004). Quantreg: An r package for quantile regression and related methods.

Koenker, R. (2005). *Quantile regression*. Cambridge University Press.

Koenker, R., & Bassett, G. (1978). Regression quantiles. *Econometrica*, *46*(1), 33–50. Retrieved April 21, 2024, from http://www.jstor.org/stable/1913643

Koenker, R., & Machado, J. A. F. (1999). Goodness of fit and related inference processes for quantile regression. *Journal of the American Statistical Association*, *94*(448), 1296–1310. Retrieved April 22, 2024, from http://www.jstor.org/stable/2669943

Kuan, C.-M. (2007). *An introduction to quantile regression* (Research Report). Institute of Economics, Academia Sinica.

Lancaster, H. O. (1961). Significance tests in discrete distributions. *Journal of the American Statistical Association*, *56*(294), 223–234. https://doi.org/10.1080/01621459.1961.10482105

Lee, M.-j. (1992). Median regression for ordered discrete response. *Journal of Econometrics*, *51*(1), 59–77. https://doi.org/https://doi.org/10.1016/0304-4076(92)90029-Q

Levine, R., & Casella, G. (2001). Implementations of the mcem algorithm. *Journal of Computational and Graphical Statistics - J COMPUT GRAPH STAT*, *10*. https://doi.org/10.1198/106186001317115045

Ling, W., Cheng, B., Wei, Y., Willey, J., & Cheung, Y. (2022). Statistical inference in quantile regression for zero-inflated outcomes. *Statistica Sinica*, *32*. https://doi.org/10.5705/ss.202020.0368

Lu, X., & Fan, Z. (2020). Generalized linear mixed quantile regression with panel data. *PLOS ONE*, *15*(8), 1–16. https://doi.org/10.1371/journal.pone.0237326

Ma, Y., Genton, M., & Parzen, E. (2011). Asymptotic properties of sample quantiles of discrete distributions. *Annals of the Institute of Statistical Mathematics*, *63*, 227–243. https://doi.org/10.1007/s10463-008-0215-z

Machado, J. A. F., & Silva, J. M. C. S. (2005). Quantiles for counts. *Journal of the American Statistical Association*, *100*(472), 1226–1237. https://doi.org/10.1198/016214505000000330

Manning, W. G. (1998). The logged dependent variable, heteroscedasticity, and the retransformation problem. *Journal of Health Economics*, *17*(3), 283–295. https://ideas.repec.org/a/eee/jhecon/v17y1998i3p283-295.html

Manski, C. F. (1975). Maximum score estimation of the stochastic utility model of choice. *Journal of Econometrics*, *3*(3), 205–228. https://doi.org/https://doi.org/10.1016/0304-4076(75)90032-9

Manski, C. F. (1985). Semiparametric analysis of discrete response: Asymptotic properties of the maximum score estimator. *Journal of Econometrics*, *27*(3), 313–333. https://doi.org/https://doi.org/10.1016/0304-4076(85)90009-0

Massicotte, P., Bertolo, A., Brodeur, P., Hudon, C., Mingelbier, M., & Magnan, P. (2015). Influence of the aquatic vegetation landscape on larval fish abundance. *Journal of Great Lakes Research*, *41*. https://doi.org/10.1016/j.jglr.2015.05.010

Millard, S. P. (2013). *Envstats: An r package for environmental statistics*. Springer. https://www.springer.com

Muller, J.-M., Brisebarre, N., Dinechin, F., Jeannerod, C.-P., Lefèvre, V., Melquiond, G., Revol, N., Stehlé, D., & Torres, S. (2010, January). *Handbook of floating-point arithmetic*. https://doi.org/10.1007/978-0-8176-4705-6

Newey, W., & Powell, J. (1987). Asymmetric least squares estimation and testing. *Econometrica*, *55*, 819–47. https://doi.org/10.2307/1911031

Padellini, T., & Rue, H. (2018). Model-based quantile regression for discrete data.

Paris, R. (2011, January). Incomplete gamma and related functions.

Parzen, E. (1993). Change pp plot and continous sample quantile function. *Communications in Statistics - Theory and Methods*, *22*(12), 3287–3304. https://doi.org/10.1080/03610929308831216

Pearson, E. S. (1950). On questions raised by the combination of tests based on discontinuous distributions. *Biometrika*, *37*(3/4), 383–398. Retrieved April 22, 2024, from http://www.jstor.org/stable/2332389

Peng, L. (2021). Quantile regression for survival data. *Annual Review of Statistics and Its Application*, *8*, 413–437. https://doi.org/10.1146/annurev-statistics-042720-020233

Pepin, P., King, J., Holt, C., Gurney-Smith, H., Shackell, N., Hedges, K., & Bundy, A. (2022). Incorporating knowledge of changes in climatic, oceanographic and ecological conditions in canadian stock assessments. *Fish and Fisheries*, *23*. https://doi.org/10.1111/faf.12692

Perillo, A., Mazzoni, L. G., Passos, L., Goulart, V., Duca, C., & Young, R. (2017). Anthropogenic noise reduces bird species richness and diversity in urban parks. *Ibis*, *159*. https://doi.org/10.1111/ibi.12481

Powell, J. (1986). Censored regression quantiles. *Journal of Econometrics*, *32*(1), 143–155. https://EconPapers.repec.org/RePEc:eee:econom:v:32:y:1986:i:1:p:143-155

Powell, J. L. (1984). Least absolute deviations estimation for the censored regression model. *Journal of Econometrics*, *25*(3), 303–325. https://ideas.repec.org/a/eee/econom/v25y1984i3p303-325.html

Rampichini, C., Grilli, L., & Varriale, R. (2015). Statistical modelling of gained university credits to evaluate the role of pre-enrolment assessment tests: An approach based on quantile regression for counts. *Statistical Modelling*, *16*. https://doi.org/10.1177/1471082X15596087

Saha, D., Alluri, P., Dumbaugh, E., & Gan, A. (2020). Application of the poisson-tweedie distribution in analyzing crash frequency data. *Accident Analysis & Prevention*, *137*, 105456. https://doi.org/https://doi.org/10.1016/j.aap.2020.105456

Siddiqui, M. (1960). Distributions of quantiles from a bivariate population. *Journal of Research of the National Bureau of Standards*, *64*, 145–150.

Smyth, G. K. (2007). Regression analysis of quantity data with exact zeroes ∗. https://api.semanticscholar.org/CorpusID:971001

Stevens, W. L. C. (1950). Fiducial limits of the parameter of a discontinuous distribution. *Biometrika*, *37 1-2*, 117–29. https://api.semanticscholar.org/CorpusID:2069401

Stigler, S. M. (1977). Fractional order statistics, with applications. *Journal of the American Statistical Association*, *72*(359), 544–550. https://doi.org/10.1080/01621459.1977.10480611

Venables, W. N., & Ripley, B. D. (2002). *Modern applied statistics with s* (Fourth) [ISBN 0-387-95457-0]. Springer. https://www.stats.ox.ac.uk/pub/MASS4/

Wagner, H. M. (1959). Linear programming techniques for regression analysis. *Journal of the American Statistical Association*, *54*(285), 206–212. https://doi.org/10.1080/01621459.1959.10501506

Walker, A. M. (1968). A Note on the Asymptotic Distribution of Sample Quantiles. *Journal of the Royal Statistical Society: Series B (Methodological)*, *30*(3), 570–575. https://doi.org/10.1111/j.2517-6161.1968.tb00757.x

Woods, P. J., Macdonald, J. I., Bárðarson, H., Bonanomi, S., Boonstra, W. J., Cornell, G., Cripps, G., Danielsen, R., Färber, L., Ferreira, A. S. A., Ferguson, K., Holma, M., Holt, R. E., Hunter, K. L., Kokkalis, A., Langbehn, T. J., Ljungström, G., Nieminen, E., Nordström, M. C., … Yletyinen, J. (2021). A review of adaptation options in fisheries management to support resilience and transition under socio-ecological change. *ICES Journal of Marine Science*, *79*(2), 463–479. https://doi.org/10.1093/icesjms/fsab146

Appendix

Table 4.1A: Estimates $\hat{\beta}$, biases ($bias$) and relative efficiencies (REF) to the estimators of β_0, β_1, and β_2 using different methods (ORD, UJ, AUJ50, TJ, ATJ50, ATJ150, BJ, ABJ50, ABJ150) at three quantiles $\tau = (0.25, 0.5, 0.75)$ based on sample size $n = 500$.

		β_0		β_1		β_2	
τ	Method	Bias	MSE	Bias	MSE	Bias	MSE
Case 1. *Poisson*							
0.25	ORD	0.3001	0.8273	-0.1818	0.3931	-0.0491	0.1903
	UJ	0.0188	0.0073	-0.0103	0.0040	-0.0103	0.0081
	AUJ50	0.0208	0.0048	-0.0090	0.0021	-0.0125	0.0061
	TJ	-0.0078	0.0069	0.0035	0.0026	0.0022	0.0068
	ATJ50	-0.0040	0.0047	0.0022	0.0018	0.0177	0.0051
	ATJ150	-0.0035	0.0046	0.0021	0.0017	-0.0168	0.0050
	BJ	0.0007	0.0062	-0.0064	0.0032	0.0021	0.0073
	ABJ50	-0.0013	0.0041	-0.0044	0.0018	0.0000	0.0053
	ABJ150	-0.0011	0.0041	-0.0050	0.0017	0.0005	0.0052
0.5	ORD	-0.0017	0.0067	0.0034	0.0044	-0.0039	0.0068
	UJ	0.0081	0.0039	-0.0037	0.0023	-0.0090	0.0046
	AUJ50	0.0039	0.0029	-0.0029	0.0021	-0.0076	0.0041

Continued on next page

		β_0		β_1		β_2	
τ	Method	Bias	MSE	Bias	MSE	Bias	MSE
	TJ	-0.0069	0.0037	-0.0002	0.0020	0.0026	0.0042
	ATJ50	0.0020	0.0022	-0.0019	0.0014	-0.0027	0.0034
	ATJ150	0.0014	0.0021	-0.0010	0.0013	-0.0025	0.0032
	BJ	0.0089	0.0033	0.0008	0.0019	-0.0124	0.0041
	ABJ50	0.0080	0.0028	0.0009	0.0015	-0.0098	0.0035
	ABJ150	0.0075	0.0026	0.0004	0.0014	-0.0076	0.0035
0.75	ORD	0.0010	0.0053	-0.0021	0.0036	0.0000	0.0055
	UJ	-0.0058	0.0026	0.0049	0.0014	0.0054	0.0035
	AUJ50	-0.0072	0.0021	0.0039	0.0011	0.0103	0.0030
	TJ	-0.0081	0.0027	0.0005	0.0012	0.0114	0.0053
	ATJ50	0.0070	0.0020	-0.0021	0.0011	0.0003	0.0033
	ATJ150	-0.0067	0.0019	-0.0007	0.0010	0.0021	0.0031
	BJ	-0.0085	0.0025	0.0023	0.0014	0.0069	0.0033
	ABJ50	-0.0065	0.0022	0.0067	0.0011	0.0040	0.0027
	ABJ150	-0.0055	0.0021	0.0017	0.0010	0.0037	0.0025
Case 2. *Zero-inflated Poisson*							
0.25	ORD	-1.5599	9.3840	-1.3825	6.2528	-0.0249	0.5656
	UJ	-0.0056	0.0566	-0.0590	0.0493	-0.0535	0.0945
	AUJ50	-0.0258	0.0480	-0.0833	0.0443	-0.0473	0.0755
	TJ	-0.0310	0.0521	-0.0814	0.0460	-0.0635	0.0899
	ATJ50	-0.0411	0.0475	-0.0909	0.0423	-0.0711	0.0830
	ATJ150	-0.0298	0.0469	-0.0786	0.0415	-0.056	0.0823
	BJ	-0.0457	0.0521	-0.0778	0.0511	-0.0333	0.0773

Continued on next page

Table 4.1A – continued from previous page

		β_0		β_1		β_2	
τ	Method	Bias	MSE	Bias	MSE	Bias	MSE
	ABJ50	-0.0500	0.0390	-0.0656	0.0340	-0.0232	0.0530
	ABJ150	-0.0401	0.0371	-0.0551	0.0325	-0.0214	0.0515
0.5	ORD	-0.0926	0.0764	0.0596	0.0494	0.0343	0.0243
	UJ	-0.0161	0.0104	0.0120	0.0049	-0.0055	0.0097
	AUJ50	-0.0160	0.0060	0.0054	0.0027	0.0050	0.0079
	TJ	0.0022	0.0065	0.0025	0.0037	0.0018	0.0080
	ATJ50	0.0115	0.0042	-0.0035	0.0020	-0.0200	0.0062
	ATJ150	0.0078	0.0040	-0.0019	0.0018	-0.0020	0.0059
	BJ	0.0024	0.0070	0.0027	0.0030	-0.0019	0.0097
	ABJ50	-0.0038	0.0041	0.0088	0.0019	0.0005	0.0065
	ABJ150	-0.0018	0.0038	0.0058	0.0017	0.0010	0.0062
0.75	ORD	-0.0073	0.0061	0.0143	0.0044	0.01433	0.0076
	UJ	-0.0109	0.0044	0.0106	0.0024	0.0044	0.0054
	AUJ50	-0.0107	0.0032	0.0117	0.0021	-0.0005	0.0041
	TJ	-0.0142	0.0046	0.0066	0.0021	0.0112	0.0049
	ATJ50	-0.0158	0.0029	0.0092	0.0018	0.0165	0.0045
	ATJ150	-0.0124	0.0027	0.0061	0.0018	0.0125	0.0045
	BJ	-0.0094	0.0038	0.0114	0.0019	0.0147	0.0043
	ABJ50	-0.0185	0.0030	0.0109	0.0015	0.0150	0.0037
	ABJ150	-0.0164	0.0029	0.0102	0.0014	0.0142	0.0036
Case 3. *Negative binomial*							
0.25	ORD	-0.1426	0.3563	0.0178	0.0562	0.1538	0.6351
	UJ	-0.0323	0.0489	0.0075	0.0313	-0.0044	0.0774

Continued on next page

τ	Method	β_0 Bias	β_0 MSE	β_1 Bias	β_1 MSE	β_2 Bias	β_2 MSE
	AUJ50	-0.0036	0.0310	0.0062	0.0152	-0.0303	0.0599
	TJ	-0.0024	0.0313	0.0013	0.0216	-0.0103	0.0393
	ATJ50	-0.0051	0.0149	0.0082	0.0100	-0.0355	0.0284
	ATJ50	-0.0053	0.0145	0.0087	0.0097	-0.0361	0.0280
	BJ	0.0042	0.0195	0.0050	0.0056	-0.0191	0.0198
	ABJ50	0.0011	0.0185	0.0063	0.0029	-0.0147	0.0142
	ABJ150	0.0021	0.0184	0.0065	0.0028	-0.0137	0.0140
0.5	ORD	0.0053	0.0323	-0.0137	0.0222	-0.0137	0.0306
	UJ	0.0023	0.0101	-0.0096	0.0050	0.0101	0.0183
	AUJ50	0.0163	0.0071	-0.0066	0.0039	0.0012	0.0143
	TJ	0.0230	0.0078	-0.0102	0.0044	-0.0186	0.0151
	ATJ50	0.0204	0.0070	-0.0103	0.0034	-0.0128	0.0125
	ATJ150	0.0196	0.0068	-0.0098	0.0033	-0.0106	0.0123
	BJ	-0.0037	0.0075	-0.0080	0.0043	-0.0119	0.0140
	ABJ50	-0.0042	0.0062	-0.0110	0.0035	-0.0054	0.0137
	ABJ150	0.0044	0.0061	-0.0090	0.0034	-0.0037	0.0137
0.75	ORD	-0.0078	0.0109	-0.0178	0.0084	0.0084	0.0162
	UJ	-0.0143	0.0060	-0.0108	0.0042	0.0123	0.0108
	AUJ50	-0.0248	0.0056	0.0003	0.0036	0.0239	0.0101
	TJ	-0.0169	0.0061	0.0004	0.0040	0.0058	0.0102
	ATJ50	-0.0072	0.0053	0.0070	0.0032	-0.0028	0.0095
	ATJ150	-0.0055	0.0052	0.0073	0.0031	-0.0021	0.0093
	BJ	-0.0071	0.0050	-0.0015	0.0048	0.0064	0.0088

Continued on next page

Table 4.1A – continued from previous page

τ	Method	β_0 Bias	β_0 MSE	β_1 Bias	β_1 MSE	β_2 Bias	β_2 MSE
	ABJ50	-0.0060	0.0043	-0.0031	0.0029	0.0012	0.0087
	ABJ150	-0.0053	0.0042	-0.0035	0.0027	-0.0011	0.0086
Case 4. *Zero-inflated negative binomial*							
0.25	ORD	-	-	-	-	-	-
	UJ	-0.0083	0.0574	-0.0104	0.0469	-0.0157	0.0681
	AUJ50	-0.0044	0.0405	-0.0369	0.0384	-0.0458	0.0310
	TJ	0.0534	0.0287	0.0744	0.0275	-0.0399	0.0512
	ATJ50	0.0346	0.0284	-0.475	0.0258	-0.0387	0.0281
	ATJ150	0.0268	0.0283	-0.356	0.0256	-0.0356	0.0288
	BJ	0.0555	0.0289	0.0734	0.0274	-0.0368	0.0510
	ABJ50	0.0272	0.0283	0.0471	0.0254	-0.0466	0.0277
	ABJ150	0.0188	0.0282	0.0378	0.0253	-0.0474	0.0270
0.5	ORD	0.3406	1.8217	-0.2705	1.0512	-0.0989	0.3232
	UJ	0.0156	0.0200	0.0194	0.0122	-0.0324	0.0347
	AUJ50	0.0397	0.0191	-0.0038	0.0075	-0.0179	0.0327
	TJ	-0.0326	0.0183	0.0091	0.0114	-0.0252	0.0412
	ATJ50	-0.0386	0.0172	0.0165	0.0079	-0.0076	0.0328
	ATJ150	-0.0386	0.0172	0.0165	0.0079	-0.0076	0.0328
	BJ	-0.0188	0.0199	0.0078	0.0102	0.0185	0.0356
	ABJ50	-0.0194	0.0167	0.0017	0.0071	-0.0218	0.0311
	ABJ50	-0.0197	0.0165	0.0014	0.0069	-0.0223	0.0309
0.75	ORD	-0.0250	0.0162	0.0013	0.0093	-0.0150	0.0260
	UJ	-0.0229	0.0110	-0.0051	0.0068	0.0172	0.0171

Continued on next page

Table 4.1A – continued from previous page

τ	Method	β_0 Bias	β_0 MSE	β_1 Bias	β_1 MSE	β_2 Bias	β_2 MSE
	AUJ50	-0.0336	0.0096	-0.0003	0.0060	0.0201	0.0167
	TJ	-0.0159	0.0094	0.0192	0.0076	0.0099	0.0167
	ATJ50	-0.013	0.0091	0.0172	0.0067	0.0050	0.0153
	ATJ150	-0.010	0.0090	0.0170	0.0066	0.0039	0.0152
	BJ	-0.0210	0.0079	0.0010	0.0042	0.0124	0.0152
	ABJ50	-0.0192	0.0057	0.0112	0.0041	0.0070	0.0115
	ABJ150	-0.0179	0.0056	0.0121	0.0041	0.0065	0.0113

Table 4.2A: Standard deviations ($s.d.$) of 1000 estimates of quantile regression parameters $\beta_{\tau 0}$, $\beta_{\tau 1}$ and $\beta_{\tau 2}$ and 1000 sample averages of estimated asymptotic standard errors ($s.e.$) when counts follow different Poisson, negative binomial and zero-inflated distributions are reported by using different methods (TJ, ATJ50, BJ, ABJ50) based on sample size $n = 200$.

τ	Method	$\beta_{\tau 0}$ s.d.	$\beta_{\tau 0}$ s.e.	$\beta_{\tau 1}$ s.d.	$\beta_{\tau 1}$ s.e.	$\beta_{\tau 2}$ s.d.	$\beta_{\tau 2}$ s.e.
Case 1. *Poisson*							
0.25	TJ	0.1369	0.1406	0.0999	0.1058	0.1404	0.1484
	ATJ50	0.1065	0.1116	0.0723	0.0812	0.1208	0.1307
	BJ	0.1356	0.1391	0.1008	0.1048	0.1383	0.1467
	ABJ50	0.1032	0.1129	0.0696	0.0708	0.1187	0.1323
0.50	TJ	0.0906	0.0949	0.0735	0.0725	0.0985	0.1080
	ATJ50	0.0802	0.0864	0.0619	0.0658	0.0908	0.1023
	BJ	0.0922	0.0958	0.0703	0.0733	0.0961	0.1095
	ABJ50	0.0809	0.0862	0.0625	0.0655	0.0914	0.1017

Continued on next page

Table 4.2A – continued from previous page

τ	Method	$\beta_{\tau 0}$ s.d.	$\beta_{\tau 0}$ s.e.	$\beta_{\tau 1}$ s.d.	$\beta_{\tau 1}$ s.e.	$\beta_{\tau 2}$ s.d.	$\beta_{\tau 2}$ s.e.
0.75	TJ	0.0818	0.0831	0.0616	0.0637	0.0922	0.1015
	ATJ50	0.0712	0.0754	0.0531	0.0570	0.0815	0.0938
	BJ	0.0770	0.0833	0.0623	0.0628	0.0909	0.1000
	ABJ50	0.0712	0.0748	0.0542	0.0571	0.0845	0.0942
Case 2. *Zero-inflated Poisson*							
0.25	TJ	0.3808	0.4406	0.3399	0.3859	0.4788	0.5737
	ATJ50	0.3340	0.3904	0.2738	0.3538	0.3898	0.4187
	BJ	0.3876	0.4442	0.3417	0.3859	0.4985	0.5917
	ABJ50	0.3340	0.3989	0.2790	0.3656	0.3916	0.5218
0.50	TJ	0.1404	0.1511	0.0994	0.1100	0.1665	0.1708
	ATJ50	0.1227	0.1369	0.0866	0.0973	0.1475	0.1648
	BJ	0.1483	0.1523	0.0997	0.1098	0.1529	0.1730
	ABJ50	0.1212	0.1371	0.0863	0.0981	0.1462	0.1648
0.75	TJ	0.0679	0.0658	0.0481	0.0498	0.0691	0.0721
	ATJ50	0.0639	0.0617	0.0452	0.0478	0.0637	0.0722
	BJ	0.0566	0.0617	0.0431	0.0466	0.0688	0.0723
	ABJ50	0.0934	0.1005	0.0713	0.0786	0.1155	0.1224
Case 3. *Negative binomial*							
0.25	TJ	0.1854	0.1955	0.1409	0.1528	0.2599	0.2737
	ATJ50	0.1543	0.1648	0.1007	0.1198	0.2113	0.2388
	BJ	0.1891	0.1967	0.1484	0.1533	0.2503	0.2728
	ABJ50	0.1569	0.1661	0.1126	0.1214	0.2173	0.2405
0.50	TJ	0.1430	0.1584	0.1135	0.1241	0.1921	0.2165

Continued on next page

Table 4.2A – continued from previous page

τ	Method	$\beta_{\tau 0}$ s.d.	$\beta_{\tau 0}$ s.e.	$\beta_{\tau 1}$ s.d.	$\beta_{\tau 1}$ s.e.	$\beta_{\tau 2}$ s.d.	$\beta_{\tau 2}$ s.e.
	ATJ50	0.1336	0.1494	0.1054	0.1163	0.1825	0.2078
	BJ	0.1421	0.1584	0.1197	0.1213	0.1975	0.2154
	ABJ50	0.1381	0.1469	0.1081	0.1173	0.1879	0.2080
0.75	TJ	0.1313	0.1446	0.1086	0.1156	0.1799	0.2005
	ATJ50	0.1209	0.1337	0.0981	0.1128	0.1710	0.1913
	BJ	0.1273	0.1395	0.1051	0.1090	0.1731	0.1972
	ABJ50	0.1221	0.1315	0.0933	0.1133	0.1734	0.1911
Case 4. *Zero-inflated negative binomial*							
0.25	TJ	0.3264	0.3413	0.3065	0.3140	0.4450	0.4764
	ATJ50	0.2677	0.2809	0.2559	0.2608	0.3651	0.3845
	BJ	0.3212	0.3427	0.2912	0.3152	0.4458	0.4749
	ABJ50	0.2688	0.2896	0.2576	0.2676	0.3614	0.3886
0.50	TJ	0.2091	0.2164	0.1436	0.1597	0.2816	0.3085
	ATJ50	0.1821	0.2081	0.1423	0.1485	0.2718	0.2958
	BJ	0.2045	0.2182	0.1549	0.1643	0.2825	0.3080
	ABJ50	0.1893	0.2083	0.1481	0.1513	0.2810	0.3038
0.75	TJ	0.1560	0.1631	0.1254	0.1307	0.2010	0.2268
	ATJ50	0.1530	0.1641	0.1245	0.1325	0.2193	0.2312
	BJ	0.1583	0.1606	0.1259	0.1311	0.2109	0.2264
	ABJ50	0.1594	0.1614	0.1163	0.1376	0.2089	0.2352

Table 4.3A: Converage probabilities of 1000 estimates of quantile regression parameters $\beta_{\tau 0}$, $\beta_{\tau 1}$ and $\beta_{\tau 2}$ at levels of confidence $\alpha = (0.01, 0.05, 0.1)$ are reported by using different quantile regression methods different methods (TJ, ATJ50, BJ, ABJ50) based on sample size $n = 200$.

		$\beta_{\tau 0}$			$\beta_{\tau 1}$			$\beta_{\tau 2}$		
τ	Method	$P_{0.99}$	$P_{0.95}$	$P_{0.90}$	$P_{0.99}$	$P_{0.95}$	$P_{0.90}$	$P_{0.99}$	$P_{0.95}$	$P_{0.90}$
Case 1. *Poisson*										
0.25	TJ	0.990	0.953	0.917	0.989	0.959	0.914	0.996	0.966	0.905
	ATJ50	0.979	0.946	0.911	0.993	0.955	0.914	0.984	0.961	0.911
	BJ	0.987	0.945	0.897	0.983	0.947	0.902	0.985	0.946	0.892
	ABJ50	0.990	0.952	0.912	0.992	0.961	0.911	0.994	0.954	0.912
0.50	TJ	0.988	0.950	0.916	0.988	0.946	0.891	0.985	0.954	0.910
	ATJ50	0.993	0.954	0.918	0.994	0.955	0.910	0.991	0.953	0.921
	BJ	0.985	0.942	0.898	0.988	0.952	0.910	0.987	0.960	0.914
	ABJ50	0.990	0.958	0.911	0.990	0.949	0.903	0.993	0.963	0.920
0.75	TJ	0.991	0.940	0.881	0.982	0.946	0.946	0.986	0.951	0.889
	ATJ50	0.987	0.941	0.889	0.983	0.942	0.904	0.987	0.955	0.914
	BJ	0.984	0.944	0.890	0.983	0.939	0.894	0.981	0.933	0.900
	ABJ50	0.997	0.959	0.904	0.984	0.939	0.890	0.982	0.941	0.908
Case 2. *Zero-inflated Poisson*										
0.25	TJ	0.986	0.956	0.916	0.978	0.936	0.905	0.936	0.971	0.927
	ATJ50	0.983	0.951	0.921	0.977	0.937	0.910	0.991	0.975	0.948
	BJ	0.985	0.957	0.911	0.974	0.937	0.904	0.989	0.960	0.928
	ABJ50	0.991	0.952	0.915	0.993	0.953	0.907	0.995	0.966	0.912
0.50	TJ	0.986	0.955	0.914	0.992	0.951	0.915	0.986	0.941	0.904
	ATJ50	0.994	0.949	0.902	0.992	0.956	0.915	0.988	0.947	0.906

Continued on next page

τ	Method	$\beta_{\tau 0}$ $P_{0.99}$	$\beta_{\tau 0}$ $P_{0.95}$	$\beta_{\tau 0}$ $P_{0.90}$	$\beta_{\tau 1}$ $P_{0.99}$	$\beta_{\tau 1}$ $P_{0.95}$	$\beta_{\tau 1}$ $P_{0.90}$	$\beta_{\tau 2}$ $P_{0.99}$	$\beta_{\tau 2}$ $P_{0.95}$	$\beta_{\tau 2}$ $P_{0.90}$
	BJ	0.989	0.952	0.904	0.993	0.954	0.905	0.981	0.944	0.896
	ABJ50	0.987	0.951	0.906	0.985	0.951	0.916	0.988	0.939	0.887
0.75	TJ	0.988	0.945	0.891	0.981	0.943	0.892	0.996	0.957	0.913
	ATJ50	0.989	0.945	0.898	0.985	0.946	0.888	0.990	0.947	0.895
	BJ	0.989	0.951	0.903	0.986	0.954	0.909	0.981	0.941	0.900
	ABJ50	0.987	0.933	0.884	0.990	0.958	0.900	0.992	0.947	0.947
Case 3. *Negative binomial*										
0.25	TJ	0.995	0.957	0.908	0.987	0.941	0.890	0.990	0.943	0.895
	ATJ50	0.988	0.941	0.893	0.992	0.965	0.915	0.992	0.966	0.912
	BJ	0.983	0.948	0.891	0.988	0.942	0.892	0.989	0.960	0.913
	ABJ50	0.998	0.968	0.916	0.994	0.958	0.906	0.996	0.952	0.906
0.50	TJ	0.985	0.948	0.889	0.984	0.952	0.912	0.982	0.945	0.897
	ATJ50	0.984	0.952	0.916	0.990	0.956	0.912	0.993	0.956	0.911
	BJ	0.984	0.942	0.896	0.986	0.946	0.892	0.990	0.936	0.894
	ABJ50	0.989	0.954	0.905	0.989	0.953	0.911	0.989	0.963	0.911
0.75	TJ	0.985	0.936	0.897	0.982	0.946	0.896	0.990	0.956	0.904
	ATJ50	0.986	0.946	0.893	0.986	0.946	0.910	0.990	0.957	0.914
	BJ	0.987	0.949	0.909	0.986	0.934	0.895	0.984	0.944	0.900
	ABJ50	0.985	0.944	0.889	0.990	0.957	0.911	0.987	0.946	0.906
Case 4. *Zero-inflated negative binomial*										
0.25	TJ	0.983	0.945	0.911	0.982	0.948	0.907	0.996	0.963	0.923
	ATJ50	0.980	0.935	0.894	0.981	0.942	0.890	0.995	0.968	0.923
	BJ	0.977	0.941	0.902	0.987	0.947	0.894	0.993	0.964	0.923

Continued on next page

Table 4.3A – continued from previous page

τ	Method	$\beta_{\tau 0}$ $P_{0.99}$	$\beta_{\tau 0}$ $P_{0.95}$	$\beta_{\tau 0}$ $P_{0.90}$	$\beta_{\tau 1}$ $P_{0.99}$	$\beta_{\tau 1}$ $P_{0.95}$	$\beta_{\tau 1}$ $P_{0.90}$	$\beta_{\tau 2}$ $P_{0.99}$	$\beta_{\tau 2}$ $P_{0.95}$	$\beta_{\tau 2}$ $P_{0.90}$
	ABJ50	0.986	0.945	0.900	0.995	0.950	0.909	0.990	0.960	0.920
0.50	TJ	0.986	0.944	0.896	0.990	0.955	0.916	0.986	0.950	0.905
	ATJ50	0.985	0.939	0.888	0.980	0.950	0.885	0.995	0.955	0.905
	BJ	0.984	0.950	0.897	0.992	0.951	0.912	0.988	0.953	0.908
	ABJ50	0.976	0.937	0.896	0.985	0.957	0.911	0.987	0.938	0.897
0.75	TJ	0.991	0.952	0.907	0.982	0.942	0.891	0.996	0.958	0.912
	ATJ50	0.992	0.952	0.904	0.983	0.938	0.898	0.990	0.945	0.905
	BJ	0.983	0.942	0.887	0.982	0.943	0.897	0.981	0.937	0.893
	ABJ50	0.985	0.954	0.903	0.986	0.957	0.914	0.985	0.944	0.895

Table 4.5A: Power of test of 1000 estimates of quantile regression parameters $\beta_{\tau 2}$ at nominal level $\alpha = 0.05$ are reported by using different quantile regression methods (TJ, ATJ50, BJ, and ABJ50) based on sample size $n = 500$.

τ	Method	$\beta_{\tau 1} = 0.2$	$\beta_{\tau 1} = 0.4$	$\beta_{\tau 1} = 0.6$	$\beta_{\tau 1} = 0.8$	$\beta_{\tau 1} = 1.0$
Case 1. *Poisson*						
0.25	TJ	0.996	1.000	1.000	1.000	1.000
	ATJ50	1.000	1.000	1.000	1.000	1.000
	BJ	0.998	1.000	1.000	1.000	1.000
	ABJ50	1.000	1.000	1.000	1.000	1.000
0.50	TJ	1.000	1.000	1.000	1.000	1.000
	ATJ50	1.000	1.000	1.000	1.000	1.000
	BJ	1.000	1.000	1.000	1.000	1.000

Continued on next page

Table 4.5A – continued from previous page

τ	Method	$\beta_{\tau 1} = 0.2$	$\beta_{\tau 1} = 0.4$	$\beta_{\tau 1} = 0.6$	$\beta_{\tau 1} = 0.8$	$\beta_{\tau 1} = 1.0$
	ABJ50	1.000	1.000	1.000	1.000	1.000
0.75	TJ	1.000	1.000	1.000	1.000	1.000
	ATJ50	1.000	1.000	1.000	1.000	1.000
	BJ	1.000	1.000	1.000	1.000	1.000
	ABJ50	1.000	1.000	1.000	1.000	1.000
Case 2. *Zero-inflated Poisson*						
0.25	TJ	0.330	0.666	0.837	0.909	0.910
	ATJ50	0.385	0.747	0.915	0.926	0.937
	BJ	0.346	0.696	0.834	0.904	0.909
	ABJ50	0.389	0.754	0.918	0.932	0.943
0.50	TJ	1.000	1.000	1.000	1.000	1.000
	ATJ50	1.000	1.000	1.000	1.000	1.000
	BJ	1.000	1.000	1.000	1.000	1.000
	ABJ50	1.000	1.000	1.000	1.000	1.000
0.75	TJ	1.000	1.000	1.000	1.000	1.000
	ATJ50	1.000	1.000	1.000	1.000	1.000
	BJ	1.000	1.000	1.000	1.000	1.000
	ABJ50	1.000	1.000	1.000	1.000	1.000
Case 3. *Negative binomial*						
0.25	TJ	0.744	0.898	0.991	0.997	1.000
	ATJ50	0.845	0.899	0.999	1.000	1.000
	BJ	0.738	0.896	0.993	0.998	1.000
	ABJ50	0.854	0.886	0.998	1.000	1.000

Continued on next page

τ	Method	$\beta_{\tau 1} = 0.2$	$\beta_{\tau 1} = 0.4$	$\beta_{\tau 1} = 0.6$	$\beta_{\tau 1} = 0.8$	$\beta_{\tau 1} = 1.0$
0.50	TJ	1.000	1.000	1.000	1.000	1.000
	ATJ50	1.000	1.000	1.000	1.000	1.000
	BJ	1.000	1.000	1.000	1.000	1.000
	ABJ50	1.000	1.000	1.000	1.000	1.000
0.75	TJ	1.000	1.000	1.000	1.000	1.000
	ATJ50	1.000	1.000	1.000	1.000	1.000
	BJ	1.000	1.000	1.000	1.000	1.000
	ABJ50	1.000	1.000	1.000	1.000	1.000
Case 4. *Zero-inflated Negative binomial*						
0.25	TJ	0.185	0.580	0.885	0.975	0.981
	ATJ50	0.301	0.892	0.998	1.000	1.000
	BJ	0.189	0.593	0.889	0.964	0.983
	ABJ50	0.306	0.899	0.998	1.000	1.000
0.50	TJ	0.673	0.995	1.000	1.000	1.000
	ATJ50	0.743	0.999	1.000	1.000	1.000
	BJ	0.650	0.994	1.000	1.000	1.000
	ABJ50	0.739	1.000	1.000	1.000	1.000
0.75	TJ	0.883	1.000	1.000	1.000	1.000
	ATJ50	0.876	1.000	1.000	1.000	1.000
	BJ	0.868	1.000	1.000	1.000	1.000
	ABJ50	0.889	1.000	1.000	1.000	1.000

Table 4.6A: Biases ($Bias$), mean squared errores (MSE) and relative efficiencies (REF) of estimators $\beta_{\tau 0}$, $\beta_{\tau 1}$, and $\beta_{\tau 2}$ and bias ($Bias^{*}$) of estimators $\gamma_{\tau 0}$, $\gamma_{\tau 1}$, and $\gamma_{\tau 2}$ on five quantiles $\tau = (0.10, 0.25, 0.5, 0.75, 0.90)$ of data when error follows normal, student's t and Chi-squared are reported using different methods (ORD, MID, UJ, AUJ50, TJ, ATJ50, BJ, ABJ50) based on sample size $n = 500$.

		$\beta_{\tau 0}$				$\beta_{\tau 1}$				$\beta_{\tau 2}$			
τ	Method	Bias	Bias*	MSE	REF	Bias	Bias*	MSE	REF	Bias	Bias*	MSE	REF
Case 1. $\varepsilon \sim \mathcal{N}(0,1)$													
0.10	ORD	-0.0644	-0.0590	0.0886		0.0107	0.0011	0.0007		-0.0348	0.0221	0.0382	
	MID	-	-	-	-	-	-	-	-	-	-	-	-
	UJ	0.0541	-0.0590	0.0867		-0.0043	0.0011	0.0007		0.0367	0.0221	0.0454	
	AUJ50	-0.0158	-0.0590	0.0692		0.0013	0.0011	0.0006		0.0070	0.0221	0.0237	
	TJ	-0.0592	-0.0590	0.9742		0.0016	0.0011	0.0006		-0.0342	0.0221	0.0345	
	ATJ50	-0.0533	-0.0590	0.0756		0.0035	0.0011	0.0004		-0.0322	0.0221	0.0301	
	BJ	-0.0519	-0.0590	0.0721		0.0018	0.0011	0.0009		-0.0113	0.0221	0.0477	
	ABJ50	-0.0531	-0.0590	0.0615		0.0065	0.0011	0.0005		-0.0240	0.0221	0.0317	
0.25	ORD	-0.0427	0.0118	0.0469		0.0041	-0.0019	0.0008		0.0130	-0.0040	0.0264	
	MID	0.0105	0.0118	0.0554		0.0009	-0.0019	0.0005		0.0496	-0.0040	0.0346	

Continued on next page

Table 4.6A – continued from previous page

		$\beta_{\tau 0}$				$\beta_{\tau 1}$				$\beta_{\tau 2}$			
τ	Method	Bias	Bias*	MSE	REF	Bias	Bias*	MSE	REF	Bias	Bias*	MSE	REF
	UJ	-0.0246	0.0118	0.0669		0.0025	-0.0019	0.0005		0.0204	-0.0040	0.0260	
	AUJ50	-0.0258	0.0118	0.0421		0.0024	-0.0019	0.0003		0.0130	-0.0040	0.0213	
	TJ	0.0156	0.0118	0.0578		-0.0032	-0.0019	0.0006		-0.0089	-0.0040	0.0258	
	ATJ50	-0.0112	0.0118	0.0502		0.0010	-0.0019	0.0005		0.0055	-0.0040	0.0234	
	BJ	0.0178	0.0118	0.0556		-0.0023	-0.0019	0.0005		-0.0082	-0.0040	0.0254	
	ABJ50	-0.0116	0.0118	0.0495		0.0008	-0.0019	0.0004		0.0040	-0.0040	0.0206	
0.50	ORD	-0.0325	0.0231	0.0467		0.0038	-0.0014	0.0004		0.0376	-0.0027	0.0190	
	MID	0.0093	0.0231	0.0045		-0.0035	-0.0014	0.0004		0.0058	-0.0027	0.0039	
	UJ	-0.0376	0.0231	0.0582		0.0048	-0.0014	0.0007		0.0285	-0.0027	0.0229	
	AUJ50	-0.0198	0.0231	0.0346		0.0022	-0.0014	0.0005		0.0141	-0.0027	0.0207	
	TJ	-0.0256	0.0231	0.0687		-0.0004	-0.0014	0.0004		0.0026	-0.0027	0.0239	
	ATJ50	-0.0161	0.0231	0.0388		0.0022	-0.0014	0.0003		-0.0110	-0.0027	0.0163	
	BJ	-0.0312	0.0231	0.0667		-0.0003	-0.0014	0.0004		-0.0015	-0.0027	0.0204	
	ABJ50	-0.0317	0.0231	0.0443		0.0003	-0.0014	0.0004		0.0056	-0.0027	0.0188	

Continued on next page

Table 4.6A – continued from previous page

		$\beta_{\tau 0}$				$\beta_{\tau 1}$				$\beta_{\tau 2}$			
τ	Method	Bias	Bias*	MSE	REF	Bias	Bias*	MSE	REF	Bias	Bias*	MSE	REF
0.75	ORD	0.0427	-0.0596	0.0553		-0.0044	-0.0012	0.0006		-0.0098	-0.0045	0.0179	
	MID	0.0372	-0.0596	0.0363		-0.0039	-0.0012	0.0003		0.0441	-0.0045	0.0205	
	UJ	0.0434	-0.0596	0.0599		-0.0043	-0.0012	0.0006		-0.0287	-0.0045	0.0286	
	AUJ50	0.0611	-0.0596	0.0491		-0.0053	-0.0012	0.0004		-0.0238	-0.0045	0.0236	
	TJ	0.0232	-0.0596	0.0592		-0.0020	-0.0012	0.0005		-0.0060	-0.0045	0.0252	
	ATJ50	-0.0157	-0.0596	0.0407		0.0009	-0.0012	0.0003		-0.0120	-0.0045	0.0200	
	BJ	0.0235	-0.0596	0.0507		0.0018	-0.0012	0.0004		-0.0232	-0.0045	0.0264	
	ABJ50	-0.0134	-0.0596	0.0478		0.0005	-0.0012	0.0004		0.0040	-0.0045	0.0209	
0.90	ORD	-0.0158	-0.0151	0.0862		0.0022	-0.0014	0.0010		-0.0049	0.0162	0.0457	
	MID	-0.3240	-0.0151	0.1359		-0.0238	-0.0014	0.0012		-0.0173	0.0162	0.0692	
	UJ	-0.0549	-0.0151	0.0972		0.0042	-0.0014	0.0009		-0.0103	0.0162	0.0422	
	AUJ50	-0.0025	-0.0151	0.0659		0.0029	-0.0014	0.0006		-0.0118	0.0162	0.0298	
	TJ	0.0169	-0.0151	0.0945		-0.0039	-0.0014	0.0009		0.0372	0.0162	0.0451	
	ATJ50	-0.0031	-0.0151	0.0612		-0.0015	-0.0014	0.0006		0.0225	0.0162	0.0301	

Continued on next page

Table 4.6A – continued from previous page

τ	Method	$\beta_{\tau 0}$ Bias	$\beta_{\tau 0}$ Bias*	$\beta_{\tau 0}$ MSE	$\beta_{\tau 0}$ REF	$\beta_{\tau 1}$ Bias	$\beta_{\tau 1}$ Bias*	$\beta_{\tau 1}$ MSE	$\beta_{\tau 1}$ REF	$\beta_{\tau 2}$ Bias	$\beta_{\tau 2}$ Bias*	$\beta_{\tau 2}$ MSE	$\beta_{\tau 2}$ REF
	BJ	0.0164	-0.0151	0.0950		-0.0031	-0.0014	0.0008		0.0369	0.0162	0.0431	
	ABJ50	-0.0028	-0.0151	0.0608		-0.0014	-0.0014	0.0005		0.0214	0.0162	0.0294	
Case 2. $\varepsilon \sim \mathcal{T}_2$													
0.10	ORD	-0.1603	0.0379	0.4796		0.0143	-0.0036	0.0045		-0.0430	-0.0300	0.2071	
	MID	-0.0541	0.0379	0.3827		0.0052	-0.0036	0.0030		0.1586	-0.0300	0.1941	
	UJ	0.0051	0.0379	0.4195		-0.0038	-0.0036	0.0040		0.0037	-0.0300	0.1931	
	UJ50	-0.0935	0.0379	0.3830		0.0071	-0.0036	0.0035		-0.0512	-0.0300	0.1818	
	TJ	-0.0591	0.0379	0.3732		0.0021	-0.0036	0.0038		0.0699	-0.0300	0.2201	
	ATJ50	-0.0201	0.0379	0.3099		0.0029	-0.0036	0.0029		0.0510	-0.0300	0.1540	
	BJ	-0.0618	0.0379	0.3705		0.0016	-0.0036	0.0034		0.0729	-0.0300	0.2182	
	ABJ50	-0.0196	0.0379	0.3097		0.0027	-0.0036	0.0029		0.0509	-0.0300	0.1573	
0.25	ORD	-0.0571	-0.0347	0.1090		0.0048	-0.0024	0.0011		-0.0034	0.0191	0.0467	
	MID	-0.0086	-0.0347	0.1134		-0.0013	-0.0024	0.0008		-0.0066	0.0191	0.0518	
	UJ	-0.0622	-0.0347	0.0865		0.0024	-0.0024	0.0011		0.0183	0.0191	0.0474	

Continued on next page

Table 4.6A – continued from previous page

τ	Method	$\beta_{\tau 0}$				$\beta_{\tau 1}$				$\beta_{\tau 2}$			
		Bias	Bias*	MSE	REF	Bias	Bias*	MSE	REF	Bias	Bias*	MSE	REF
	UJ50	-0.0211	-0.0347	0.0955		0.0038	-0.0024	0.0008		-0.0010	0.0191	0.0423	
	TJ	-0.0368	-0.0347	0.1043		0.0043	-0.0024	0.0010		0.0323	0.0191	0.0452	
	ATJ50	0.0367	-0.0347	0.0879		-0.0024	-0.0024	0.0008		-0.0111	0.0191	0.0419	
	BJ	-0.0364	-0.0347	0.1030		0.0040	-0.0024	0.0009		0.0318	0.0191	0.0443	
	ABJ50	0.0344	-0.0347	0.0850		-0.0021	-0.0024	0.0007		-0.0235	0.0191	0.0416	
0.50	ORD	0.0362	0.0156	0.0598		-0.0032	0.0021	0.0006		-0.0486	-0.0236	0.0320	
	MID	0.0522	0.0156	0.0119		-0.0055	0.0021	0.0001		0.0089	-0.0236	0.0066	
	UJ	-0.0136	0.0156	0.0624		-0.0023	0.0021	0.0006		-0.0040	-0.0236	0.0328	
	UJ50	0.0331	0.016	0.0449		0.0029	0.0021	0.0005		-0.0489	-0.0236	0.0286	
	TJ	-0.0399	0.0156	0.0592		0.0027	0.0021	0.0005		0.0110	-0.0236	0.0311	
	ATJ50	0.0327	0.0156	0.0535		-0.0019	0.0021	0.0005		-0.0101	-0.0236	0.0277	
	BJ	-0.0400	0.0156	0.0587		0.0025	0.0021	0.0005		0.0103	-0.0236	0.0306	
	ABJ50	0.0154	0.0156	0.0533		-0.0024	0.0021	0.0004		-0.0101	-0.0236	0.0272	
0.75	OR	0.1224	-0.0205	0.1248		-0.0111	0.0013	0.0010		-0.0407	0.0102	0.0640	

Continued on next page

Table 4.6A – continued from previous page

τ	Method	$\beta_{\tau 0}$ Bias	$\beta_{\tau 0}$ Bias*	$\beta_{\tau 0}$ MSE	$\beta_{\tau 0}$ REF	$\beta_{\tau 1}$ Bias	$\beta_{\tau 1}$ Bias*	$\beta_{\tau 1}$ MSE	$\beta_{\tau 1}$ REF	$\beta_{\tau 2}$ Bias	$\beta_{\tau 2}$ Bias*	$\beta_{\tau 2}$ MSE	$\beta_{\tau 2}$ REF
	MID	0.0908	-0.0205	0.1853		-0.0108	0.0013	0.0015		0.0163	0.0102	0.0924	
	UJ	0.1306	-0.0205	0.1395		-0.0091	0.0013	0.0012		0.0476	0.0102	0.0563	
	UJ50	0.0856	-0.0205	0.1303		-0.0065	0.0013	0.0011		0.0301	0.0102	0.0530	
	TJ	-0.0133	-0.0205	0.0998		0.0024	0.0013	0.0009		0.0123	0.0102	0.0514	
	ATJ50	-0.0286	-0.0205	0.0798		0.0013	0.0013	0.0006		0.0157	0.0102	0.0495	
	BJ	-0.0124	-0.0205	0.0987		0.0010	0.0013	0.0009		0.0149	0.0102	0.0550	
	ABJ50	-0.0259	-0.0205	0.0785		0.0019	0.0013	0.0006		-0.0182	0.0102	0.0496	
0.90	OR	0.0718	0.0503	0.5355		-0.0045	-0.0011	0.0046		-0.0410	0.0313	0.2342	
	MID	-0.1702	0.0503	0.4709		0.0099	-0.0011	0.0039		0.1491	0.0313	0.2608	
	UJ	0.0310	0.0503	0.4750		-0.0016	-0.0011	0.0039		-0.0116	0.0313	0.2381	
	UJ50	0.0699	0.0503	0.4390		-0.0044	-0.0011	0.0036		-0.0385	0.0313	0.2013	
	TJ	0.0642	0.0503	0.4102		-0.0015	-0.0030	0.0037		-0.0271	0.0313	0.2011	
	ATJ50	0.0180	0.0503	0.3255		-0.0011	-0.0011	0.0029		0.0271	0.0313	0.2029	
	BJ	0.0611	0.0503	0.4099		-0.0013	-0.0011	0.0037		-0.0299	0.0313	0.2000	

Continued on next page

Table 4.6A – continued from previous page

		$\beta_{\tau 0}$				$\beta_{\tau 1}$				$\beta_{\tau 2}$			
τ	Method	Bias	Bias*	MSE	REF	Bias	Bias*	MSE	REF	Bias	Bias*	MSE	REF
	ABJ50	0.0640	0.0503	0.3135		0.0012	-0.0011	0.0028		0.0280	0.0313	0.1842	
Case 3. $\varepsilon \sim \mathcal{X}_3^2$													
0.10	OR	-0.0173	-0.0071	0.0792		0.0031	0.0013	0.0007		-0.0102	0.0093	0.0398	
	MID	0.0480	-0.0071	0.0880		-0.0028	0.0013	0.0007		-0.0122	0.0093	0.0326	
	UJ	0.0097	-0.0071	0.0783		-0.0024	0.0013	0.0006		0.0374	0.0093	0.0335	
	UJ50	0.0072	-0.0071	0.0538		0.0014	0.0013	0.0005		-0.0049	0.0093	0.0281	
	TJ	0.0167	-0.0071	0.0786		-0.0011	0.0013	0.0005		0.0125	0.0093	0.0329	
	ATJ50	-0.0047	-0.0071	0.0467		0.0018	0.0013	0.0003		-0.0067	0.0093	0.0212	
	BJ	0.0170	-0.0071	0.0792		-0.0013	0.0013	0.0006		0.0133	0.0093	0.0339	
	ABJ50	-0.0058	-0.0071	0.0479		0.0019	0.0013	0.0004		-0.0092	0.0093	0.0216	
0.25	OR	-0.0492	0.0161	0.1071		0.0080	0.0015	0.0010		0.0128	0.0122	0.0501	
	MID	-0.0904	0.0161	0.0713		0.0196	0.0015	0.0008		-0.0010	0.0122	0.0414	
	UJ	-0.0515	0.0161	0.0954		0.0070	0.0015	0.0009		0.0239	0.0122	0.0505	
	UJ50	-0.0356	0.0161	0.0898		0.0064	0.0015	0.0008		0.0190	0.0122	0.0424	

Continued on next page

Table 4.6A – continued from previous page

		$\beta_{\tau 0}$				$\beta_{\tau 1}$				$\beta_{\tau 2}$			
τ	Method	Bias	Bias*	MSE	REF	Bias	Bias*	MSE	REF	Bias	Bias*	MSE	REF
	TJ	0.0174	0.0161	0.0904		-0.0006	0.0015	0.0008		0.0112	0.0122	0.0439	
	ATJ50	-0.0105	0.0161	0.0840		0.0015	0.0015	0.0006		0.0162	0.0122	0.0405	
	BJ	-0.0315	0.0161	0.0917		0.0018	0.0015	0.0008		0.0246	0.0122	0.0508	
	ABJ50	0.0114	0.0161	0.0676		-0.0019	0.0015	0.0007		0.0152	0.0122	0.0346	
0.50	OR	0.0670	-0.0111	0.2119		-0.0088	-0.0016	0.0020		0.0979	0.0102	0.0957	
	MID	0.0102	-0.0111	0.1361		0.0009	-0.0016	0.0011		0.0613	0.0102	0.1025	
	UJ	0.0009	-0.0111	0.1971		0.0021	-0.0016	0.0020		0.0391	0.0102	0.1155	
	UJ50	-0.0043	-0.0111	0.1659		-0.0019	-0.0016	0.0017		0.0310	0.0102	0.0855	
	TJ	0.0020	-0.0111	0.1939		-0.0013	-0.0016	0.0017		0.0279	0.0102	0.0823	
	ATJ50	-0.0089	-0.0111	0.1649		0.0010	-0.0016	0.0014		-0.0182	0.0102	0.0829	
	BJ	-0.0235	-0.0111	0.2120		0.0026	-0.0016	0.0019		0.0296	0.0102	0.0949	
	ABJ50	-0.0174	-0.0111	0.1630		0.0025	-0.0016	0.0014		0.0182	0.0102	0.0728	
0.75	OR	-0.1158	-0.0403	0.5035		0.0151	0.0014	0.0054		0.2190	0.0043	0.2319	
	MID	0.0785	-0.0403	0.4660		-0.0258	0.0014	0.0045		0.0545	0.0043	0.2224	

Continued on next page

Table 4.6A – continued from previous page

		$\beta_{\tau 0}$				$\beta_{\tau 1}$				$\beta_{\tau 2}$			
τ	Method	Bias	Bias*	MSE	REF	Bias	Bias*	MSE	REF	Bias	Bias*	MSE	REF
	UJ	0.1789	-0.0403	0.5301		-0.0195	0.0014	0.0048		0.0516	0.0043	0.2254	
	UJ50	-0.0687	-0.0403	0.4764		0.0090	0.0014	0.0041		0.0519	0.0043	0.2621	
	TJ	-0.0379	-0.0403	0.4994		0.0012	0.0014	0.0044		0.0679	0.0043	0.2129	
	ATJ50	-0.0786	-0.0403	0.4094		0.0084	0.0014	0.0037		0.0012	0.0043	0.2018	
	BJ	-0.1007	-0.0403	0.4882		0.0112	0.0014	0.0044		-0.0490	0.0043	0.2290	
	ABJ50	0.0504	-0.0403	0.4192		0.0029	0.0014	0.0038		-0.0040	0.0043	0.1933	
0.90	OR	0.0497	0.1030	1.4362		-0.0110	0.0058	0.0130		0.2307	0.0252	0.6046	
	MID	-0.1942	0.1030	1.3269		0.0200	0.0058	0.0112		-0.0155	0.0252	0.5340	
	UJ	0.1823	0.1030	1.6021		-0.0262	0.0058	0.0141		0.1870	0.0252	0.7870	
	UJ50	0.1258	0.1030	1.3352		-0.0156	0.0058	0.0122		0.2213	0.0252	0.6469	
	TJ	-0.0975	0.1030	1.3519		0.0206	0.0058	0.0111		-0.0119	0.0252	0.5637	
	ATJ50	-0.0914	0.1030	1.1702		0.0054	0.0058	0.0106		-0.0201	0.0252	0.5948	
	BJ	-0.0531	0.1030	1.1995		0.0156	0.0058	0.0108		0.0045	0.0252	0.6230	
	ABJ50	0.0909	0.1030	1.0761		-0.0069	0.0058	0.0097		-0.0225	0.0252	0.5963	

Table 5.7A: Standard deviations ($s.d.$) of 1000 estimates of quantile regression parameters $\beta_{\tau 0}$, $\beta_{\tau 1}$ and $\beta_{\tau 2}$ and 1000 sample averages of estimated asymptotic standard errors ($s.e.$) based on truncated data by using different quantile regression methods and Standard deviations ($s.d.^{*}$) of 1000 estimates of quantile regression parameters $\gamma_{\tau 0}$, $\gamma_{\tau 1}$ and $\gamma_{\tau 2}$ based on original continuous data are reported. Quantile levels $\tau = (0.1, 0.25, 0.5, 0.75, 0.9)$ and sample size $n = 500$ are used.

		$\beta_{\tau 0}$			$\beta_{\tau 1}$			$\beta_{\tau 2}$		
τ	Method	s.d.	s.e.	s.d.*	s.d.	s.e.	s.d.*	s.d.	s.e.	s.d.*
Case 1. $\varepsilon \sim \mathcal{N}(0,1)$										
0.1	TJ	0.3288	0.3241	0.2987	0.0302	0.0299	0.0271	0.2235	0.2148	0.2065
	ATJ50	0.2488	0.2598	0.2817	0.0238	0.0253	0.0265	0.1799	0.1796	0.1869
	BJ	0.3161	0.3160	0.2921	0.0298	0.0298	0.0275	0.2188	0.2178	0.2032
	ABJ50	0.2429	0.2570	0.2717	0.0230	0.0243	0.0258	0.1770	0.1745	0.1824
0.25	TJ	0.2412	0.2401	0.2432	0.0232	0.0235	0.0231	0.1599	0.1606	0.1600
	ATJ50	0.2351	0.2132	0.2245	0.0214	0.0205	0.0215	0.1468	0.1489	0.1459
	BJ	0.2358	0.2397	0.2339	0.0227	0.0228	0.0225	0.1598	0.1646	0.1569
	ABJ50	0.2231	0.2053	0.2185	0.0207	0.0195	0.0208	0.1439	0.1438	0.1491
0.50	TJ	0.2183	0.2356	0.2027	0.0207	0.0224	0.0195	0.1550	0.1410	0.1345

Continued on next page

Table 4.7A – continued from previous page

		$\beta_{\tau 0}$			$\beta_{\tau 1}$			$\beta_{\tau 2}$		
τ	Method	s.d.	s.e.	s.d.*	s.d.	s.e.	s.d.*	s.d.	s.e.	s.d.*
	ATJ50	0.1967	0.1941	0.2027	0.0184	0.0185	0.0195	0.1374	0.1619	0.1345
	BJ	0.2108	0.2108	0.2027	0.0207	0.0221	0.0195	0.1433	0.1599	0.1345
	ABJ50	0.2057	0.2190	0.2027	0.0201	0.0208	0.0195	0.1374	0.1481	0.1345
0.75	TJ	0.2266	0.2343	0.2167	0.0219	0.0228	0.0209	0.1683	0.1698	0.1498
	ATJ50	0.2015	0.2133	0.2167	0.0194	0.0202	0.0209	0.1411	0.1450	0.1498
	BJ	0.2258	0.2415	0.2167	0.0217	0.0229	0.0209	0.1631	0.1657	0.1498
	ABJ50	0.2028	0.2031	0.2167	0.0189	0.0194	0.0209	0.1460	0.1407	0.1498
0.9	TJ	0.3089	0.3018	0.2801	0.0291	0.0293	0.0268	0.2055	0.2061	0.1901
	ATJ50	0.2487	0.2578	0.2788	0.0248	0.0248	0.0272	0.1745	0.1787	0.1977
	BJ	0.3086	0.3013	0.2754	0.0289	0.0289	0.0259	0.2048	0.2060	0.1853
	ABJ50	0.2473	0.2558	0.2777	0.0241	0.0244	0.0265	0.1707	0.1754	0.1946
Case 2. $\varepsilon \sim t_2$										
0.10	TJ	0.6088	0.6101	0.5974	0.0599	0.0593	0.0598	0.4658	0.4689	0.4420
	ATJ50	0.5587	0.6097	0.6066	0.0556	0.0598	0.0587	0.3956	0.4169	0.4199

Continued on next page

Table 4.7A – continued from previous page

τ	Method	$\beta_{\tau 0}$ s.d.	$\beta_{\tau 0}$ s.e.	$\beta_{\tau 0}$ s.d.*	$\beta_{\tau 1}$ s.d.	$\beta_{\tau 1}$ s.e.	$\beta_{\tau 1}$ s.d.*	$\beta_{\tau 2}$ s.d.	$\beta_{\tau 2}$ s.e.	$\beta_{\tau 2}$ s.d.*
	BJ	0.6071	0.6078	0.5847	0.0591	0.0583	0.0579	0.4626	0.4660	0.4373
	ABJ50	0.5576	0.6084	0.6171	0.0543	0.0582	0.0587	0.3944	0.4191	0.4225
0.25	TJ	0.3201	0.3256	0.2998	0.0305	0.0326	0.0295	0.2076	0.2202	0.2053
	ATJ50	0.2787	0.2934	0.2933	0.0273	0.0289	0.0281	0.1945	0.2097	0.2087
	BJ	0.3197	0.3297	0.2992	0.0300	0.0315	0.0285	0.2070	0.2206	0.2066
	ABJ50	0.2767	0.2947	0.2944	0.0264	0.0284	0.0276	0.2033	0.2170	0.2048
0.50	TJ	0.2401	0.2585	0.2301	0.0237	0.0252	0.0224	0.1764	0.1771	0.1635
	ATJ50	0.2298	0.2314	0.2324	0.0212	0.0223	0.0216	0.1614	0.1598	0.1601
	BJ	0.2396	0.2576	0.2279	0.0233	0.0246	0.0217	0.1753	0.1757	0.1702
	ABJ50	0.2293	0.2301	0.2301	0.0208	0.0219	0.0211	0.1608	0.1588	0.1586
0.75	TJ	0.3102	0.3127	0.2985	0.0298	0.0301	0.0298	0.2270	0.2282	0.2068
	ATJ50	0.2798	0.2856	0.2912	0.0264	0.0272	0.0273	0.2125	0.2085	0.2076
	BJ	0.3084	0.3154	0.2980	0.0299	0.0303	0.0285	0.2125	0.2265	0.2048
	ABJ50	0.2771	0.2821	0.2908	0.0259	0.0268	0.0270	0.2101	0.2087	0.2092

Continued on next page

Table 4.7A – continued from previous page

τ	Method	$\beta_{\tau 0}$ s.d.	$\beta_{\tau 0}$ s.e.	$\beta_{\tau 0}$ s.d.*	$\beta_{\tau 1}$ s.d.	$\beta_{\tau 1}$ s.e.	$\beta_{\tau 1}$ s.d.*	$\beta_{\tau 2}$ s.d.	$\beta_{\tau 2}$ s.e.	$\beta_{\tau 2}$ s.d.*
0.90	TJ	0.6401	0.6592	0.6238	0.0615	0.0631	0.0594	0.0431	0.4601	0.4164
	ATJ50	0.6221	0.6516	0.5994	0.0630	0.0584	0.0614	0.4376	0.4528	0.4158
	BJ	0.6390	0.6580	0.6188	0.0610	0.0626	0.0586	0.4302	0.4595	0.4132
	ABJ50	0.6191	0.6189	0.6127	0.0585	0.0597	0.0605	0.4303	0.4521	0.4190
Case 3. $\varepsilon \sim \mathcal{X}_3^2$										
0.10	TJ	0.2810	0.2701	0.2128	0.0201	0.0251	0.0202	0.1834	0.1846	0.1452
	ATJ50	0.2186	0.2198	0.2128	0.0203	0.0201	0.0202	0.1465	0.1489	0.1452
	BJ	0.2813	0.2715	0.2128	0.0264	0.0257	0.0202	0.1840	0.1857	0.1452
	ABJ50	0.2195	0.2218	0.2128	0.0209	0.0208	0.0202	0.1474	0.1500	0.1452
0.25	TJ	0.3003	0.3063	0.3153	0.0285	0.0292	0.0279	0.2097	0.2116	0.2019
	ATJ50	0.2898	0.2713	0.3153	0.0267	0.0260	0.0279	0.1938	0.1915	0.2019
	BJ	0.3013	0.3067	0.3153	0.0283	0.0293	0.0279	0.2123	0.2129	0.2019
	ABJ50	0.2692	0.2749	0.3153	0.0278	0.0262	0.0279	0.1859	0.1881	0.2019
0.50	TJ	0.4406	0.4432	0.4323	0.0420	0.0423	0.0412	0.2858	0.3050	0.2945

Continued on next page

Table 4.7A – continued from previous page

		$\beta_{\tau 0}$			$\beta_{\tau 1}$			$\beta_{\tau 2}$		
τ	Method	s.d.	s.e.	s.d.*	s.d.	s.e.	s.d.*	s.d.	s.e.	s.d.*
	ATJ50	0.4062	0.4156	0.4295	0.0386	0.0396	0.0409	0.2874	0.2822	0.2945
	BJ	0.4602	0.4462	0.4320	0.0436	0.0427	0.0421	0.3069	0.3090	0.2976
	ABJ50	0.4037	0.4081	0.4286	0.0385	0.0389	0.0411	0.2701	0.2808	0.2873
0.75	TJ	0.7060	0.6990	0.7092	0.0671	0.0666	0.0672	0.4666	0.4786	0.4648
	ATJ50	0.6554	0.6674	0.6792	0.0620	0.0635	0.0645	0.4595	0.4665	0.4682
	BJ	0.6918	0.6926	0.6702	0.0660	0.0659	0.0640	0.4763	0.4768	0.4693
	ABJ50	0.6478	0.6678	0.6726	0.0618	0.0635	0.0644	0.4399	0.4628	0.4601
0.90	TJ	1.0592	0.9758	0.1085	0.0942	0.1103	0.1070	0.7511	0.7870	0.7812
	ATJ50	1.0790	1.1191	1.1330	0.1032	0.1062	0.1082	0.7520	0.7847	0.7570
	BJ	1.0945	0.9751	1.1173	0.1029	0.0932	0.1064	0.7731	0.7897	0.6877
	ABJ50	1.0360	1.1187	0.0989	0.1085	0.1072	0.1021	0.7738	0.7854	0.7422

Table 4.8A: Coverage probabilities of 1000 estimates of quantile regression parameters $\beta_{\tau 0}$, $\beta_{\tau 1}$ and $\beta_{\tau 2}$ are reported by using different quantile regression methods TJ, ATJ50, BJ, and ABJ50, respectively. Nominal levels $\alpha = (0.01, 0.05, 0.1)$ and sample size $n = 500$ are used.

		$\beta_{\tau 0}$			$\beta_{\tau 1}$			$\beta_{\tau 2}$		
τ	Method	$P_{0.99}$	$P_{0.95}$	$P_{0.90}$	$P_{0.99}$	$P_{0.95}$	$P_{0.90}$	$P_{0.99}$	$P_{0.95}$	$P_{0.90}$
Case 1. $\varepsilon \sim \mathcal{N}(0,1)$										
0.10	TJ	0.980	0.945	0.896	0.990	0.949	0.899	0.991	0.952	0.905
	ATJ50	0.992	0.948	0.916	0.994	0.945	0.887	0.992	0.948	0.910
	BJ	0.985	0.945	0.890	0.980	0.940	0.890	0.985	0.945	0.895
	ABJ50	0.990	0.955	0.915	0.990	0.945	0.885	0.995	0.940	0.895
0.25	TJ	0.990	0.950	0.910	0.985	0.945	0.898	0.985	0.959	0.900
	ATJ50	0.994	0.945	0.896	0.982	0.947	0.905	0.984	0.955	0.909
	BJ	0.985	0.950	0.915	0.990	0.940	0.905	0.990	0.955	0.915
	ABJ50	0.990	0.940	0.890	0.980	0.940	0.890	0.985	0.945	0.900
0.50	TJ	0.988	0.946	0.905	0.990	0.944	0.895	0.993	0.945	0.893
	ATJ50	0.989	0.954	0.903	0.985	0.943	0.908	0.989	0.939	0.897
	BJ	0.980	0.950	0.910	0.985	0.954	0.913	0.985	0.955	0.905
	ABJ50	0.985	0.953	0.901	0.985	0.947	0.905	0.980	0.937	0.890
0.75	TJ	0.991	0.958	0.900	0.990	0.942	0.899	0.990	0.954	0.903
	ATJ50	0.990	0.948	0.898	0.991	0.951	0.902	0.992	0.952	0.905
	BJ	0.995	0.960	0.905	0.993	0.953	0.905	0.990	0.960	0.900
	ABJ50	0.985	0.945	0.895	0.990	0.935	0.885	0.995	0.952	0.900
0.90	TJ	0.991	0.942	0.899	0.986	0.941	0.991	0.981	0.939	0.888
	ATJ50	0.990	0.950	0.905	0.993	0.955	0.903	0.993	0.948	0.897
	BJ	0.990	0.943	0.897	0.987	0.943	0.900	0.989	0.954	0.904

Continued on next page

τ	Method	$\beta_{\tau 0}$ $P_{0.99}$	$\beta_{\tau 0}$ $P_{0.95}$	$\beta_{\tau 0}$ $P_{0.90}$	$\beta_{\tau 1}$ $P_{0.99}$	$\beta_{\tau 1}$ $P_{0.95}$	$\beta_{\tau 1}$ $P_{0.90}$	$\beta_{\tau 2}$ $P_{0.99}$	$\beta_{\tau 2}$ $P_{0.95}$	$\beta_{\tau 2}$ $P_{0.90}$
	ABJ50	0.991	0.951	0.907	0.992	0.942	0.907	0.995	0.943	0.892
Case 2. $\varepsilon \sim \mathcal{T}_2$										
0.10	TJ	0.990	0.957	0.905	0.988	0.946	0.895	0.982	0.949	0.897
	ATJ50	0.992	0.954	0.903	0.993	0.954	0.908	0.988	0.952	0.901
	BJ	0.985	0.950	0.902	0.988	0.954	0.903	0.982	0.935	0.885
	ABJ50	0.990	0.953	0.908	0.991	0.953	0.903	0.986	0.950	0.895
0.25	TJ	0.986	0.946	0.894	0.985	0.952	0.912	0.981	0.936	0.910
	ATJ50	0.987	0.945	0.904	0.989	0.951	0.902	0.990	0.954	0.901
	BJ	0.990	0.955	0.909	0.990	0.960	0.915	0.985	0.951	0.909
	ABJ50	0.986	0.946	0.902	0.986	0.950	0.900	0.990	0.960	0.908
0.50	TJ	0.991	0.952	0.902	0.994	0.961	0.914	0.985	0.945	0.893
	ATJ50	0.987	0.956	0.904	0.993	0.958	0.903	0.987	0.941	0.895
	BJ	0.990	0.950	0.905	0.988	0.952	0.918	0.986	0.946	0.894
	ABJ50	0.985	0.952	0.902	0.991	0.952	0.909	0.987	0.939	0.889
0.75	TJ	0.987	0.947	0.895	0.983	0.943	0.904	0.993	0.953	0.906
	ATJ50	0.993	0.953	0.906	0.992	0.958	0.895	0.994	0.960	0.908
	BJ	0.985	0.944	0.891	0.985	0.965	0.902	0.986	0.953	0.906
	ABJ50	0.987	0.948	0.894	0.990	0.954	0.891	0.993	0.959	0.893
0.90	TJ	0.991	0.954	0.905	0.988	0.951	0.897	0.990	0.953	0.908
	ATJ50	0.991	0.955	0.908	0.989	0.950	0.896	0.989	0.950	0.904
	BJ	0.985	0.956	0.905	0.986	0.950	0.905	0.995	0.957	0.910
	ABJ50	0.993	0.954	0.904	0.990	0.946	0.909	0.985	0.950	0.908
Case 3. $\varepsilon \sim \mathcal{T}_2$										

Continued on next page

Table 4.8A – continued from previous page

τ	Method	$\beta_{\tau 0}$			$\beta_{\tau 1}$			$\beta_{\tau 2}$		
		$P_{0.99}$	$P_{0.95}$	$P_{0.90}$	$P_{0.99}$	$P_{0.95}$	$P_{0.90}$	$P_{0.99}$	$P_{0.95}$	$P_{0.90}$
0.10	TJ	0.991	0.952	0.894	0.991	0.951	0.902	0.993	0.950	0.897
	ATJ50	0.992	0.954	0.895	0.991	0.942	0.893	0.986	0.952	0.903
	BJ	0.990	0.952	0.882	0.986	0.946	0.892	0.992	0.952	0.898
	ABJ50	0.993	0.956	0.890	0.993	0.935	0.894	0.987	0.946	0.895
0.25	TJ	0.985	0.941	0.898	0.990	0.943	0.900	0.987	0.942	0.892
	ATJ50	0.985	0.944	0.983	0.984	0.940	0.897	0.986	0.945	0.897
	BJ	0.987	0.950	0.902	0.988	0.962	0.904	0.986	0.951	0.887
	ABJ50	0.990	0.942	0.892	0.982	0.938	0.894	0.986	0.942	0.896
0.50	TJ	0.983	0.947	0.904	0.987	0.944	0.895	0.984	0.951	0.908
	ATJ50	0.986	0.955	0.909	0.990	0.957	0.903	0.986	0.946	0.897
	BJ	0.987	0.943	0.897	0.981	0.942	0.886	0.982	0.951	0.902
	ABJ50	0.983	0.946	0.906	0.984	0.946	0.903	0.990	0.949	0.896
0.75	TJ	0.988	0.947	0.904	0.988	0.950	0.896	0.992	0.954	0.905
	ATJ50	0.992	0.945	0.905	0.987	0.947	0.899	0.988	0.947	0.897
	BJ	0.988	0.946	0.893	0.986	0.949	0.895	0.991	0.956	0.898
	ABJ50	0.992	0.953	0.905	0.992	0.950	0.900	0.994	0.953	0.906
0.90	TJ	0.991	0.948	0.905	0.991	0.953	0.901	0.991	0.952	0.904
	ATJ50	0.984	0.956	0.914	0.984	0.958	0.912	0.994	0.951	0.903
	BJ	0.992	0.958	0.903	0.991	0.952	0.902	0.991	0.942	0.897
	ABJ50	0.992	0.958	0.908	0.996	0.958	0.914	0.994	0.956	0.908

Table 4.9A: Power of $S = 1000$ hypotheses test (i.e., $1 - \beta$) as related to the quantile regression parameters $\beta_{\tau 1}$ are reported by using different quantile regression methods TJ, ATJ50, BJ, and ABJ50. Nominal level $\alpha = 0.05$ and sample size $n = 500$ are used.

τ	Method	$\beta_{\tau 1} = 0.2$	$\beta_{\tau 1} = 0.4$	$\beta_{\tau 1} = 0.6$	$\beta_{\tau 1} = 0.8$	$\beta_{\tau 1} = 1.0$
Case 1. $\varepsilon \sim \mathcal{N}(0,1)$						
0.10	TJ	0.995	0.960	1.000	1.000	1.000
	ATJ50	1.000	0.995	1.000	1.000	1.000
	BJ	0.995	0.990	1.000	1.000	1.000
	ABJ50	1.000	1.000	1.000	1.000	1.000
0.25	TJ	1.000	1.000	1.000	1.000	1.000
	ATJ50	1.000	1.000	1.000	1.000	1.000
	BJ	1.000	1.000	1.000	1.000	1.000
	ABJ50	1.000	1.000	1.000	1.000	1.000
0.50	TJ	1.000	1.000	1.000	1.000	1.000
	ATJ50	1.000	1.000	1.000	1.000	1.000
	BJ	1.000	1.000	1.000	1.000	1.000
	ABJ50	1.000	1.000	1.000	1.000	1.000
0.75	TJ	1.000	1.000	1.000	1.000	1.000
	ATJ50	1.000	1.000	1.000	1.000	1.000
	BJ	1.000	1.000	1.000	1.000	1.000
	ABJ50	1.000	1.000	1.000	1.000	1.000
0.90	TJ	1.000	1.000	1.000	1.000	1.000
	ATJ50	1.000	1.000	1.000	1.000	1.000
	BJ	1.000	1.000	1.000	1.000	1.000

Continued on next page

Table 4.9A – continued from previous page

τ	Method	$\beta_{\tau 1} = 0.2$	$\beta_{\tau 1} = 0.4$	$\beta_{\tau 1} = 0.6$	$\beta_{\tau 1} = 0.8$	$\beta_{\tau 1} = 1.0$
	ABJ50	1.000	1.000	1.000	1.000	1.000
Case 2. $\varepsilon \sim \mathcal{T}_2$						
0.10	TJ	0.855	1.000	1.000	1.000	1.000
	ATJ50	0.917	1.000	1.000	1.000	1.000
	BJ	0.858	1.000	1.000	1.000	1.000
	ABJ50	0.921	1.000	1.000	1.000	1.000
0.25	TJ	0.998	1.000	1.000	1.000	1.000
	ATJ50	1.000	1.000	1.000	1.000	1.000
	BJ	0.998	1.000	1.000	1.000	1.000
	ABJ50	1.000	1.000	1.000	1.000	1.000
0.50	TJ	1.000	1.000	1.000	1.000	1.000
	ATJ50	1.000	1.000	1.000	1.000	1.000
	BJ	1.000	1.000	1.000	1.000	1.000
	ABJ50	1.000	1.000	1.000	1.000	1.000
0.75	TJ	0.999	1.000	1.000	1.000	1.000
	ATJ50	1.000	1.000	1.000	1.000	1.000
	BJ	0.999	1.000	1.000	1.000	1.000
	ABJ50	1.000	1.000	1.000	1.000	1.000
0.90	TJ	0.904	0.998	1.000	1.000	1.000
	ATJ50	0.923	1.000	1.000	1.000	1.000
	BJ	0.908	0.999	1.000	1.000	1.000
	ABJ50	0.919	1.000	1.000	1.000	1.000
Case 3. $\varepsilon \sim \mathcal{X}_3^2$						

Continued on next page

Table 4.9A – continued from previous page

τ	Method	$\beta_{\tau 1} = 0.2$	$\beta_{\tau 1} = 0.4$	$\beta_{\tau 1} = 0.6$	$\beta_{\tau 1} = 0.8$	$\beta_{\tau 1} = 1.0$
0.10	TJ	1.000	1.000	1.000	1.000	1.000
	ATJ50	1.000	1.000	1.000	1.000	1.000
	BJ	1.000	1.000	1.000	1.000	1.000
	ABJ50	1.000	1.000	1.000	1.000	1.000
0.25	TJ	1.000	1.000	1.000	1.000	1.000
	ATJ50	1.000	1.000	1.000	1.000	1.000
	BJ	1.000	1.000	1.000	1.000	1.000
	ABJ50	1.000	1.000	1.000	1.000	1.000
0.50	TJ	1.000	1.000	1.000	1.000	1.000
	ATJ50	1.000	1.000	1.000	1.000	1.000
	BJ	0.999	1.000	1.000	1.000	1.000
	ABJ50	1.000	1.000	1.000	1.000	1.000
0.75	TJ	0.932	0.997	1.000	1.000	1.000
	ATJ50	0.997	1.000	1.000	1.000	1.000
	BJ	0.936	0.998	1.000	1.000	1.000
	ABJ50	0.998	1.000	1.000	1.000	1.000
0.90	TJ	0.612	0.963	0.999	1.000	1.000
	ATJ50	0.657	0.983	1.000	1.000	1.000
	BJ	0.617	0.966	1.000	1.000	1.000
	ABJ50	0.664	0.987	1.000	1.000	1.000

Table 5.9A: Estimates $\hat{\beta}$, biases ($bias$) and relative efficiencies (REF) to the estimators of β_0, β_1, and β_2 using different methods (OR, MID, TJ, ATJ50, ATJ150, BJ, ABJ50, ABJ150) at three quantiles (0.25, 0.5, 0.75). Sample size $n = 500$.

		$\beta_{\tau 0}$			$\beta_{\tau 1}$			$\beta_{\tau 2}$		
τ	Method	Bias	MSE	REF	Bias	MSE	REF	Bias	MSE	REF
Case 1. a) $\varepsilon \sim \mathcal{N}(0,2)$										
0.50	OLS	0.0449	0.1097	1.630	-0.0016	0.0010	1.600	0.0053	0.0477	1.918
	ORD	-0.0107	0.1788	1.000	0.0008	0.0016	1.000	0.01904	0.0915	1.000
	TJ	0.1047	0.1750	1.022	-0.0115	0.0016	1.000	-0.0858	0.0845	1.083
	ATJ50	0.1432	0.1476	1.211	-0.0153	0.0013	1.231	-0.0630	0.0675	1.356
	BJ	0.1443	0.1871	0.956	-0.0148	0.0015	1.067	-0.0914	0.0955	0.958
	ABJ50	0.1403	0.1456	1.228	-0.0150	0.0012	1.333	-0.0990	0.0765	1.196
Case 1. b) $\varepsilon \sim \mathcal{N}(2,4)$										
0.50	OLS	-0.0614	0.1080	1.599	0.0007	0.0040	1.725	-0.0540	0.2109	1.644
	ORD	-0.0412	0.1805	1.000	0.0298	0.0069	1.000	-0.0231	0.3468	1.000
	TJ	-0.0400	0.1721	0.998	0.0352	0.0074	0.932	0.0085	0.3028	1.145
	ATJ50	-0.0112	0.1725	1.111	0.0340	0.0066	1.045	-0.0002	0.2857	1.214

Continued on next page

Table 5.9A – continued from previous page

τ	Method	$\beta_{\tau 0}$			$\beta_{\tau 1}$			$\beta_{\tau 2}$		
		Bias	MSE	REF	Bias	MSE	REF	Bias	MSE	REF
	BJ	-0.0677	0.1756	1.061	0.0396	0.0080	0.863	0.0129	0.3052	1.136
	ABJ50	-0.0616	0.1704	1.164	0.0228	0.0068	1.015	-0.0247	0.2701	1.284
Case 2. a) $\varepsilon \sim \mathcal{S}tudent's\ t_3$										
0.50	OLS	0.0826	0.0923	0.728	-0.0036	0.0007	0.714	-0.0017	0.0360	0.742
	ORD	0.0500	0.0672	1.000	-0.0056	0.0005	1.000	-0.0201	0.0267	1.000
	TJ	0.0258	0.0598	1.124	-0.0030	0.0005	1.000	0.0081	0.0253	1.055
	ATJ50	0.0602	0.0447	1.503	-0.0063	0.0004	1.250	-0.0011	0.0190	1.405
	BJ	0.0570	0.0604	1.113	-0.0056	0.0005	1.000	-0.0182	0.0288	0.927
	ABJ50	0.0743	0.0585	1.149	-0.0070	0.0004	1.250	-0.0073	0.0197	1.355
Case 2. b) $\varepsilon \sim \mathcal{S}tudent's\ t_2$										
0.50	OLS	-0.0533	0.0901	0.639	-0.0019	0.0007	0.714	0.0050	0.0349	0.811
	ORD	-0.0398	0.0576	1.000	0.0048	0.0005	1.000	0.0532	0.0283	1.000
	TJ	-0.0606	0.0555	1.038	0.0069	0.0005	1.000	0.0259	0.0258	1.097
	ATJ50	-0.0437	0.0533	1.081	0.0050	0.0004	1.250	0.0295	0.0226	1.142

Continued on next page

Table 5.9A – continued from previous page

τ	Method	$\beta_{\tau 0}$			$\beta_{\tau 1}$			$\beta_{\tau 2}$		
		Bias	MSE	REF	Bias	MSE	REF	Bias	MSE	REF
	BJ	-0.0217	0.0551	1.045	0.0036	0.0005	1.000	0.0196	0.0299	1.252
	ABJ50	-0.0454	0.0515	1.118	0.0047	0.0004	1.250	0.0363	0.0218	1.298
Case 3. a) $\varepsilon \sim \mathcal{X}_2^2$										
0.50	OLS	-0.0913	0.1211	1.042	0.0002	0.0012	0.917	0.0028	0.0521	0.925
	ORD	-0.0585	0.1263	1.000	0.0016	0.0011	1.000	0.0411	0.0482	1.000
	TJ	-0.0501	0.1261	1.002	0.0077	0.0011	1.000	0.0252	0.0458	1.052
	ATJ50	-0.0440	0.1029	1029	0.0065	0.0009	1.222	0.0164	0.0381	1.265
	BJ	-0.0518	0.1131	1.117	0.0067	0.0009	1.222	0.0176	0.0475	1.015
	ABJ50	-0.0471	0.0988	1.278	0.0058	0.0008	1.375	0.0367	0.0390	1.236
Case 3. b) $\varepsilon \sim \mathcal{X}_1^2$										
0.50	OLS	0.1123	0.0712	0.494	-0.0026	0.0005	0.4	0.0040	0.0271	0.557
	ORD	0.0633	0.0352	1.000	-0.0057	0.0002	1.000	0.0271	0.0151	1.000
	TJ	0.0415	0.0331	1.063	-0.0038	0.0002	1.000	0.0277	0.0142	1.063
	ATJ50	0.0406	0.0232	1.517	-0.0042	0.0001	2.000	0.0320	0.0087	1.736

Continued on next page

Table 5.9A – continued from previous page

τ	Method	$\beta_{\tau 0}$			$\beta_{\tau 1}$			$\beta_{\tau 2}$		
		Bias	MSE	REF	Bias	MSE	REF	Bias	MSE	REF
	BJ	0.0566	0.0326	1.080	-0.0054	0.0002	1.000	0.0238	0.0151	1.000
	ABJ50	0.0610	0.0244	1.443	-0.0062	0.0001	2.000	0.0319	0.0106	1.425