Quantile Regression for Longitudinal Data

by

©Xiaoming Lu

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

Master of Science

Department of Mathematics and Statistics

Memorial University of Newfoundland

05/2014

St. John's Newfoundland

Abstract

Quantile regression is a powerful statistical methodology that complements the classical linear regression by examining how covariates influence the location, scale, and shape of the entire response distribution and offering a global view of the statistical landscape. In this thesis we propose a new quantile regression model for longitudinal data. The proposed approach incorporates the correlation structure between repeated measures to enhance the efficiency of the inference. In order to use the Newton-Raphson iteration method to obtain convergent estimates, the estimating functions are redefined as smoothed functions which are differentiable with respect to regression parameters. Our proposed method for quantile regression provides consistent estimates with asymptotically normal distributions. Simulation studies are carried out to evaluate the performance of the proposed method. As an illustration, the proposed method was applied to a real-life data that contains self-reported labor pain for women in two groups.

Acknowledgments

I would like to express the sincere appreciation to my supervisor, Dr. Zhaozhi Fan, for his guidance through the learning and writing process of this thesis. Furthermore I would like to thank my cousin Yunqi Ji and sister-in-law Bingrui Sun for giving me constant encouragement and persistent help in academic study and daily life. I am also thankful to many faculties in our Department, Dr. Brajendra Sutradhar, Dr. J.C Loredo-Osti, Dr. Alwell Oyet, Dr. Hong Wang, for introducing me to the statistical world as well for the support all the way. I would like to thank all my loved ones, my parents, Aunt Xu Fengqiu, and other relatives and friends who have supported me throughout these two years' graduate study, both by calling me and leaving messages on my social networks. I am grateful to all others who have helped.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

Longitudinal data are very common in many areas of applied studies. Such data are repeatedly collected from independent subjects over time and correlation arises between measures from the same subject. One advantage of longitudinal study is that, additional to modeling the cohort effect, one can still specify the individual patterns of change. In order to take the correlation into consideration to not only avoid loss of efficiency in estimation but also make correct statistical inference, a number of methods are developed to evaluate covariate effects on the mean of a response variable (Liang and Zeger, 1986; Qu et al., 2000; Jung and Ying, 2003). Sutradhar (2003) has proposed a generalization of the quasi-likelihood estimation approach to model the conditional mean of the response by solving the generalized quasi-likelihood (GQL) estimating equations. A general stationary auto-correlation matrix is used in this method, which, in fact, represents the correlations of many stationary dynamic models, such as stationary auto-regressive order 1 (AR(1)), stationary moving average order 1 (MA(1)), and stationary equi-correlation (EQC) models.

Quantile regression (Koenker and Bassett Jr, 1978) has become a widely used technique in applications. The effects of covariates are modeled through conditional quantiles of the response variable, rather than the conditional mean, which makes it possible to characterize any arbitrary point of a distribution and thus provide a complete description of the entire response distribution. Compared to the classical mean regression, quantile regression is more robust to outliers and the error patterns do not need to be specified. Therefore, quantile regression has been widely used, (see Chen et al., 2004; Koenker, 2005; Reich et al., 2010; Farcomeni, 2012, among others).

Recently quantile regression has been extended to longitudinal data analysis. A simple way to do so is to assume working independence that ignores correlations between repeated measures, which, of course, may cause loss of efficiency, see Wei and He (2006); Wang and He (2007); Mu and Wei (2009); Wang and Fygenson (2009); Wang (2009); Wang et al. (2009a). Jung (1996) firstly developed a quasi-likelihood method for median regression which incorporates correlations between repeated measures for longitudinal data. This method requires estimation of the correlation matrix. Based on Jung's work, Lipsitz et al. (1997) proposed a weighted GEE model. Koenker (2004) considered a random effect model for estimating quantile functions with subject specific fixed effects and based inference on a penalized likelihood. Karlsson (2008) suggested a weighted approach for a nonlinear quantile regression estimation of longitudinal data. Geraci and Bottai (2007) made inferences by using a random intercept to account for the within-subject correlations and proposed a method using the asymmetric Laplace distribution (ALD). Geraci and Bottai's work was generalized by Liu and Bottai (2009), who gave a linear mixed effect quantile regression model using a multivariate Laplace distribution. Farcomeni (2012) proposed a linear quantile regression model allowing time-varying random effects and modeled subject-specific parameters through a latent Markov chain. To reduce the loss of efficiency in inferences of quantile regression, Tang and Leng (2011) incorporate the within-subject correlations through a specified conditional mean model.

Unlike in the classical linear regression, it is difficult to account for the correlations between repeated measures in quantile regression. Misspecification of the correlation structure in GEE method also leads to loss of inferential efficiency. Moreover, the approximating algorithms for computing estimates could be very complicated, and computational problems could occur when statistical software is applied to do intensive re-samplings in the inference procedure. To overcome these problems, Fu and Wang (2012) proposed a combination of the between- and within-subject estimating equations for parameter estimation. By combining multiple sets of estimating equations, Leng and Zhang (2012) developed a new quantile regression model which produces efficient estimates. Those two papers extend the induced smoothing method (Brown and Wang, 2005) to quantile regression, and thus obtained smoothed objective functions which allow the application of Newton-Raphson iteration, and the latter automatically gives both the estimates of parameters and the sandwich estimate of their covariance matrix.

In this thesis, we propose a more general quantile regression model by appropriately incorporating a correlation structure between repeated measures in longitudinal data. By employing a general stationary auto-correlation matrix, we avoid the specification of any particular correlation structure. The correlation coefficients can be iteratively estimated in the process of the regression estimation. By using the induced smoothed estimating functions, we can obtain estimates of parameters and their asymptotic covariance matrix by using Newton-Raphson algorithm. The estimators obtained using our proposed method are consistent and asymptotically normal. The results of the intensive simulation studies reveal that our proposed method outperforms those methods based on working independence assumption. Furthermore, our approach is simpler and more general than other quantile regression methods for longitudinal data on theoretical derivation, practical application and statistical programming.

The remainder of this thesis proceeds as follows: in Chapter 2 we give an introduction to quantile regression and list some of its properties. Chapter 3 develops the proposed quantile regression method and the algorithm of parameter estimation. Asymptotic properties of parameter estimators are derived as well in this chapter. In Chapter 4, to illustrate the performance of the proposed method, we carry out extensive simulation studies and apply the method to the labor pain data. Finally, the thesis is concluded in Chapter 5.

Chapter 2

Quantile Regression

In traditional mean regression, the mean and the standard deviation are two essential measures used to describe a distribution. The mean describes the central location of one distribution, and the standard deviation describes the dispersion. However, focusing on the mean and standard deviation alone will lead us to ignore other important properties which offer more insights into the distribution. Self-thinning of tropical plants (Cade and Guo, 2000) is a very interesting example, where the effects of increasing germination densities of seedlings on the reduction in densities of mature plants were best revealed at the higher plant densities with intense intraspecific competition. Also, in social science, researchers often have data sets with skewed distribution which could not be well characterized by the mean and the standard deviation. To describe the distributional attributes of asymmetric response data sets, this chapter develops quantile-based measures of location and shape of a distribution. It also redefines a quantile as a solution to a certain minimization problem and, finally, introduces the quantile regression which examines how covariates influence the location, scale, and shape of the entire response distribution.

2.1 Quantiles and Quantile Functions

2.1.1 CDFs and Quantiles

To characterize any real-valued random variable X, we can use its cumulative distribution function(cdf):

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

A cumulative distribution function (cdf) has two important properties: monotonicity, i.e., $F(x_1) \leq F(x_2)$ whenever $x_1 \leq x_2$, and its behavior at infinity, $\lim_{x\to-\infty} F(x) = 0$ and $\lim_{x\to+\infty} F(x) = 1$. The cdf for the standard normal distribution is shown in Figure 2.1.

For a continuous random variable X, we can also represent its distribution using a probability density function (pdf) $f(x)$ defined as the function with the property that

$$P(a \leq X \leq b) = \int_a^b f(x)dx,$$

for any a and b. Hence, the relation between cdf and pdf is obvious, $F(x) = \int_{-\infty}^{x} f(u)du$.

The location and spread, or scale are usually considered as measures of a distribution. A shift in the location of one distribution while keeping the shape could be expressed by $F_1(x) = F(x + \Delta)$, where Δ is the change in location. A change of distribution may exist in both location and scale, so that the relationship between the original distribution and the distribution after the change takes the general form $F_2(x) = F(ax - c)$ for constants a and c ($a > 0$).

However, when distributions become more asymmetrical, more complex characteristics are needed.

Suppose a distribution has cdf $F(x)$ with the form in (2.1), the τth quantile of this

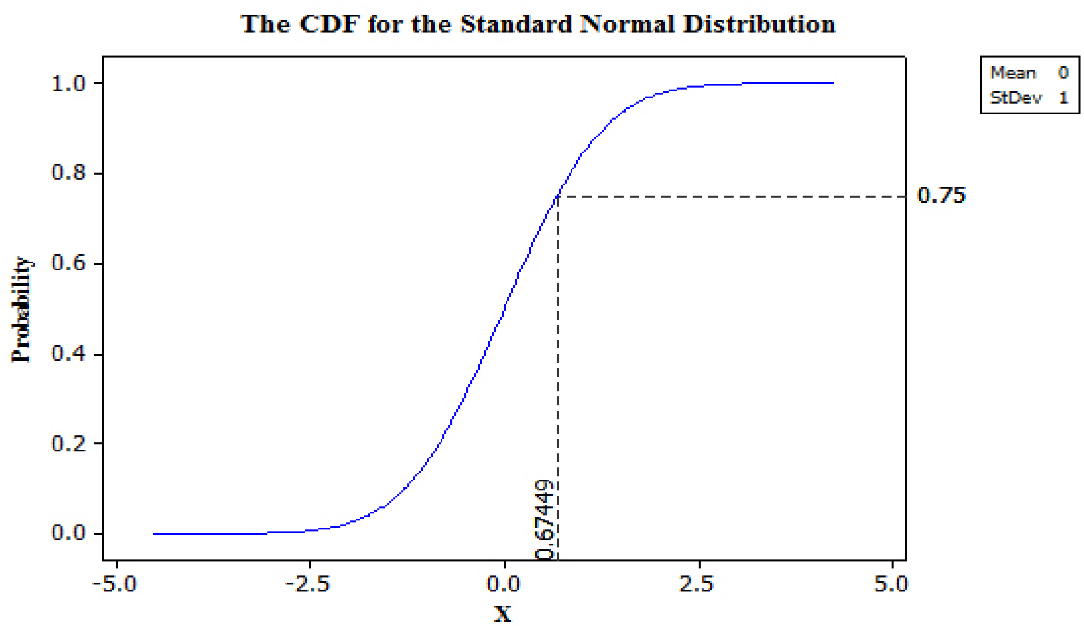

Figure 2.1: CDF for Standard Normal Distribution

distribution is denoted by

$$Q_\tau(X) = F^{-1}(\tau) = \inf\{x : F(x) \geq \tau\}. \tag{2.2}$$

Thus, the τth quantile of X is the smallest value of x such that $F(x) = \tau$, which indicates the probability of the population with X less than Q_τ is τ. For example, in the standard normal distribution case, as shown in Figure 2.1, $F(0.67449) = 0.75$, so $Q_{0.75} = 0.67449$.

When F is replaced by the empirical distribution function

$$F_n(x) = n^{-1} \sum_{i=1}^{n} I(X_i \leq x), \tag{2.3}$$

we obtain the τth *sample* quantile as

$$\hat{Q}_\tau(X) = F_n^{-1}(\tau) = \inf\{x : F_n(x) \geq \tau\}. \tag{2.4}$$

2.1.2 Quantile Functions

Definition 2.1.1. *We denote $Q_{(\cdot)}$, the function of τ, as the quantile function of F or the sample quantile function of the corresponding empirical cdf F_n.*

Figure 2.2 shows an example of a cdf and the corresponding quantile function when X is a continuous random variable. We can observe that both the cdf and the quantile function are monotonic nondecreasing continuous functions.

For an empirical distribution, there is a strong connection between sample quantiles and *order statistics*. Suppose we rank a given sample $x_1, \ldots, x_n$ from the smallest value to the largest value as $x_{(1)}, \ldots, x_{(n)}$, where $x_{(1)} \leq x_{(2)} \leq \ldots \leq x_{(n)}$. Then, if we refer to $x_{(i)}$ as the ith-order statistic corresponding to the sample, the (k/n)th sample quantile could be obtained as the kth-order sample statistic, $x_{(k)}$.

If the sample size is large enough, say $x_1, \ldots, x_n$ is a large sample drawn from a distribution with probability density function $f(X)$ and quantile function $Q_{(\cdot)}$, the distribution of sample quantile $\hat{Q}_\tau$ is approximately normal. That is

$$\hat{Q}_\tau \sim N(Q_\tau, \sigma_\tau),$$

where the mean Q_τ is the value of quantile function $Q_{(\cdot)}$ at the point τ, and the variance $\sigma_\tau = \frac{\tau(1-\tau)}{n} \times \frac{1}{f(Q_\tau)^2}$ (Walker, 1968). Using σ_τ to estimate the quantile sampling variability requires a way of estimating the unknown probability density function f. To do this, we make use of the inverse density function: $1/f(Q_\tau) = dQ_\tau/d\tau$, which can be approximated by $(\hat{Q}_{\tau+h} - \hat{Q}_{\tau-h})/2h$ for some small value of h.

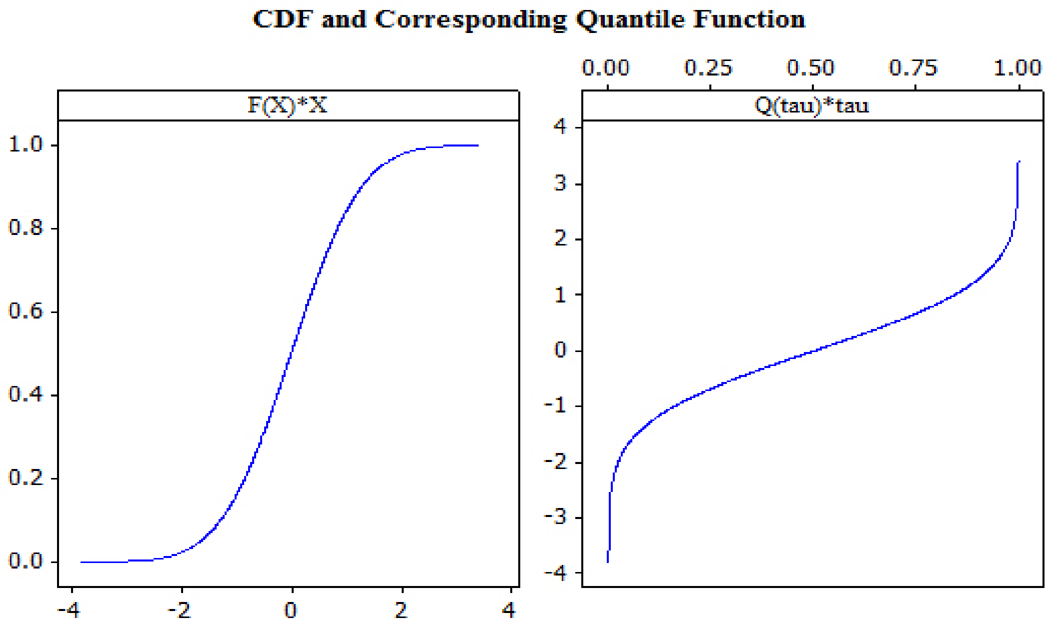

Figure 2.2: A CDF and the Corresponding Quantile Function

2.1.3 Quantile-Based Measures of Location and Shape

To study the characteristics of a distribution, it is always natural to firstly measure its location. Like using the mean to indicate the center of a symmetric distribution, and the median has been used as a *quantile-based measure* of the center of a skewed distribution. Including the median, *quantile-based location* gives a complete view of the location of the distribution not just the center. For example, one may be interested in examining a location at the lower tail (e.g., 0.1th quantile) or upper tail (e.g., 0.9th quantile) of a distribution.

Some measures used to describe the shape of a distribution are scale, skewness, kurtosis etc. Scale and skewness are more commonly used. The scale of a symmetric distribution relies on the *standard deviation*, but when a distribution becomes highly asymmetric or heavy-tailed, the standard deviation may no longer perfectly interpret the scale. To characterize the scale of a skewed or heavy tailed distribution, we use the following *quantile-based scale* measure (QSC) at a selected τ:

$$QSC_\tau = Q_{1-\tau} - Q_\tau \tag{2.5}$$

for $\tau \leq 0.5$. Therefore, we can obtain the spread of any desirable middle $100(1-2\tau)\%$, for example, the spread of the middle 95% of the population between $Q_{0.025}$ and $Q_{0.975}$, or the middle 50% of the population between $Q_{0.25}$ and $Q_{0.75}$ (the conventional interquartile range).

Another measure of the shape of a distribution is skewness, which has value zero when the distribution is symmetric, a negative value indicates left skewness and a positive value indicates right skewness. We can describe the skewness as an imbalance between the scales above and below the median. The upper and lower scales can be characterized by the quantile function. The quantile function of a symmetric distribution should be symmetric itself about the median (0.5th quantile). By contrast, the quantile function for a skewed distribution is asymmetric about the median. This can be observed from Figure 2.3 by comparing the slope of the quantile function at the 0.1th quantile with the slope at the 0.9th quantile. We can see that, for the quantile function of the standard normal distribution (the upper plot in Figure 2.3), the slope at 0.1th quantile is the same as the slope at 0.9th quantile. This is true for any pair of corresponding quantiles (Q_τ and $Q_{1-\tau}$) around the median. However, this property breaks down when the distribution becomes less symmetric. As shown in the lower plot in Figure 2.3, the quantile function of a right-skewed distribution has very different slopes at the 0.1th quantile and the 0.9th quantile.

Hence, we denote a measure of *quantile-based skewness* as an expression of a ratio

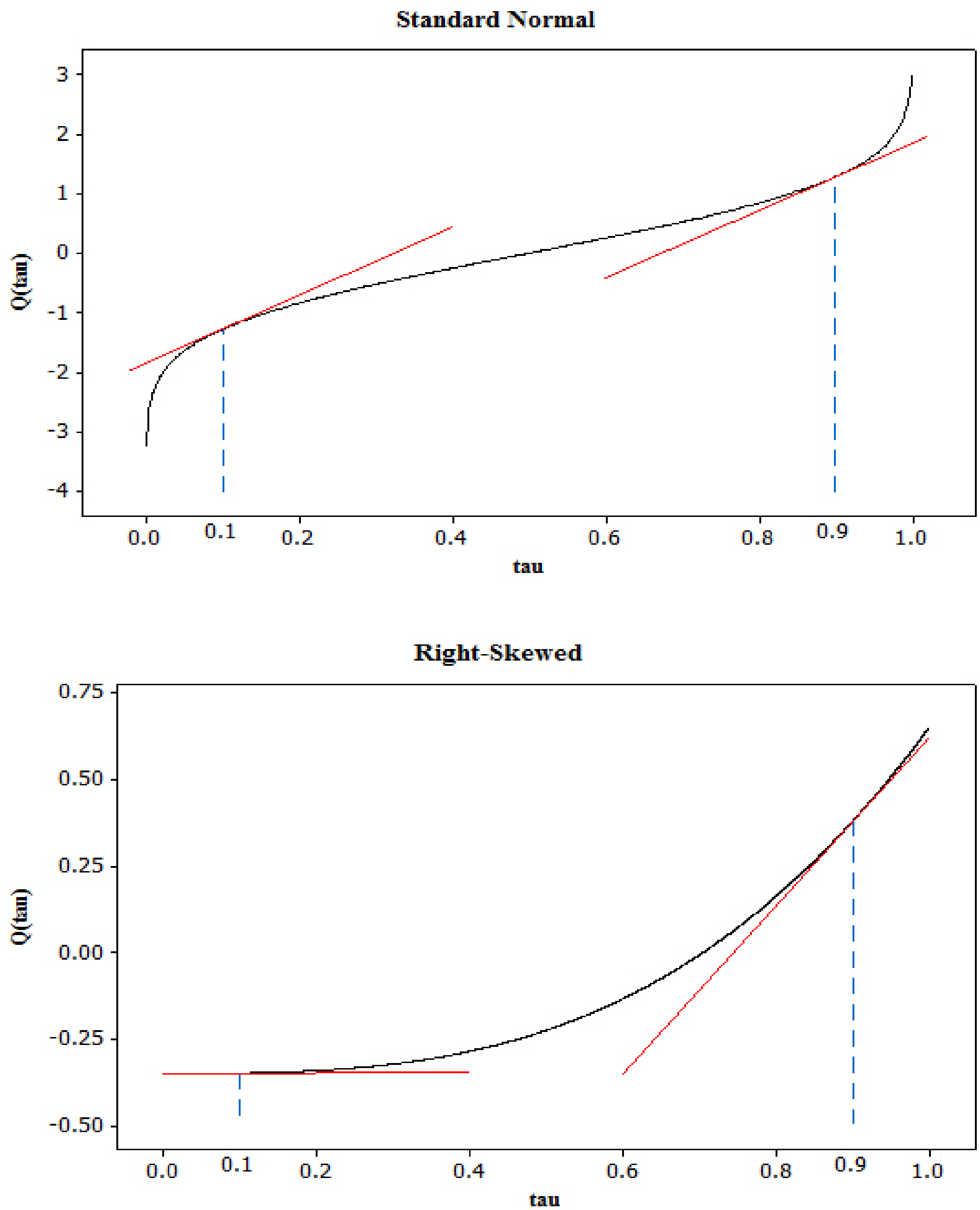

Figure 2.3: Quantile Functions for Standard Normal and a Skewed Distribution

of the upper scale to the lower scale:

$$QSK_\tau = \frac{Q_{1-\tau} - Q_{0.5}}{Q_{0.5} - Q_\tau} - 1 \tag{2.6}$$

for $\tau < 0.5$. The *minus one* (-1) part is to make sure that the quantity QSK_τ takes the value zero for a symmetric distribution, a negative value for a left-skewed distribution and a positive value for a right-skewed distribution.

2.2 Quantile Regression Model and Estimation

2.2.1 Quantile as a Solution to a Certain Minimization Problem

In previous section, we have defined quantiles in terms of cdf. However, this kind of definition does not contain useful information that can be applied to a regression problem. For this reason, we try to seek another expression of quantiles which can be used to construct regression models.

Suppose that the mean of the distribution of y can be obtained as the point μ at which the *mean squared deviation* $E[(Y-\mu)^2]$ is minimized. Since

$$\begin{aligned} E[(Y-\mu)^2] &= E[Y^2] - 2E[Y]\mu + \mu^2 \\ &= (\mu - E[Y])^2 + (E[Y^2] - (E[Y])^2) \\ &= (\mu - E[Y])^2 + Var(Y), \end{aligned}$$

and $Var(Y)$ is constant, we minimize $E[(Y-\mu)^2]$ by taking $\mu = E[Y]$. Similarly, the sample mean for a sample of size n can also be obtained by seeking the point μ that minimizes the mean squared distance:

$$\frac{1}{n}\sum_{i=1}^{n}(y_i - \mu)^2.$$

This kind of property can also be applied in finding the median of a distribution. Let $|Y - m|$ be the *absolute distance* from Y to m and $E|Y - m|$ be the *mean absolute distance*. The median is the solution of the minimization problem:

$$\min_{m \in \mathbb{R}} E|Y - m|$$

or

$$\min_{m \in \mathbb{R}} \frac{1}{n}\sum_{i=1}^{n}|y_i - m|$$

on the sample level.

To define quantiles as a solution to a minimization problem, we first define a check loss function:

$$\rho_\tau(u) = u(\tau - I(u < 0)) \tag{2.7}$$

for some $\tau \in (0, 1)$. This function means that, we give u a weight of τ if it is a positive value and a weight of $1 - \tau$ if it is a negative value. This is illustrated in Figure 2.4. Notice that, when $\tau = 0.5$, two times of the value given by the loss function is the absolute value of u. That is:

$$\begin{aligned} 2\rho_{0.5}(u) &= 2u(0.5 - I(u < 0)) \\ &= \begin{cases} -u & \text{if } u < 0 \\ u & \text{if } u > 0 \end{cases} \\ &= |u|. \end{aligned}$$

We seek to minimize the expected loss

$$
\begin{aligned}
E\rho_\tau(Y-\hat{y}) &= \int_{-\infty}^{+\infty} \rho_\tau(y-\hat{y})dF(y) \\
&= (\tau-1)\int_{-\infty}^{\hat{y}}(y-\hat{y})dF(y) + \tau\int_{\hat{y}}^{+\infty}(y-\hat{y})dF(y).
\end{aligned} \tag{2.8}
$$

Differentiating with respect to $\hat{y}$ and setting the partial derivative to zero will lead to the solution for the minimum. That is

$$
\begin{aligned}
\frac{\partial}{\partial\hat{y}}E\rho_\tau(Y-\hat{y}) &= \frac{\partial}{\partial\hat{y}}(\tau-1)\int_{-\infty}^{\hat{y}}(y-\hat{y})dF(y) + \frac{\partial}{\partial\hat{y}}\tau\int_{\hat{y}}^{+\infty}(y-\hat{y})dF(y). \\
&= (1-\tau)\int_{-\infty}^{\hat{y}} dF(y) + \tau\int_{\hat{y}}^{+\infty} dF(y) \\
&= \int_{-\infty}^{\hat{y}} dF(y) - \tau\{\int_{-\infty}^{\hat{y}} dF(y) + \int_{\hat{y}}^{+\infty} dF(y)\} \\
&= F(\hat{y}) - \tau\int_{-\infty}^{+\infty} dF(y) \\
&= F(\hat{y}) - \tau \\
&\overset{\text{set}}{=} 0.
\end{aligned} \tag{2.9}
$$

When the solution is unique, $\hat{y} = F^{-1}(\tau)$; otherwise, we choose the smallest value from a set of τth quantiles. Notice that, if $\tau = \frac{1}{2}$, the solution gives the median (0.5th quantile).

Similarly, the τth sample quantile is the value of $\hat{y}$ that minimizes the sample expected loss function, which may be written as the following problem

$$
\min_{\hat{y}\in\mathbb{R}} \sum_{i=1}^{n} \rho_\tau(y_i - \hat{y}). \tag{2.10}
$$

2.2.2 Quantile Regression

Before we introduce the *quantile regression modeling* (QRM) method, we take a look at the *linear regression modeling* (LRM) method. We know that the LRM is a widely

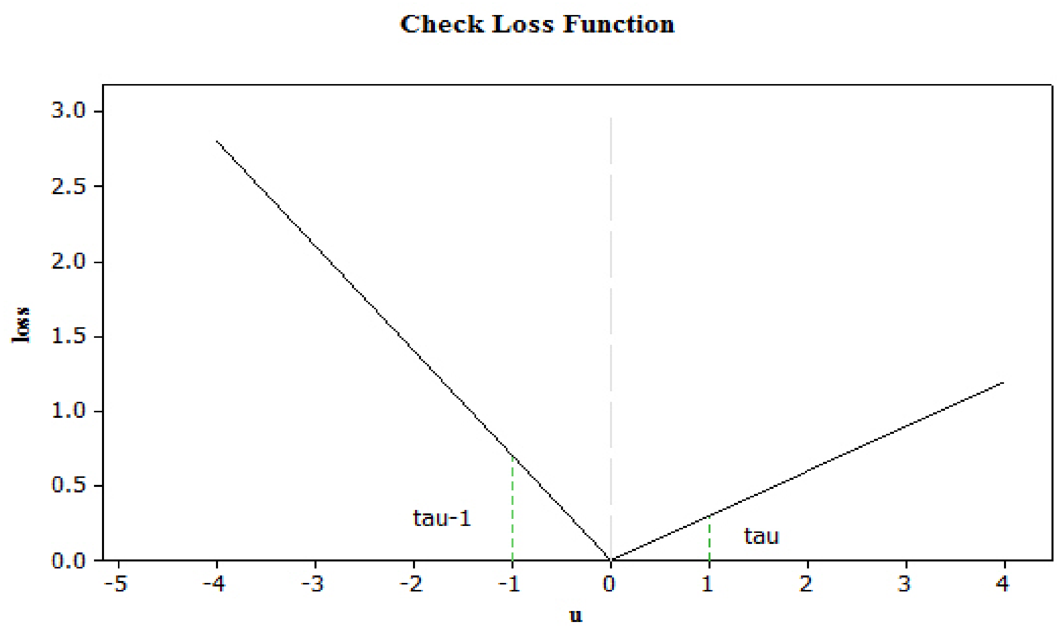

Figure 2.4: Check Loss Function for a Certain τ

used method which focuses on modeling the *conditional mean* of a response variable. Giving a sample with response variable y and covariate x, according to what we discussed in the previous section, the sample mean solves the problem

$$\min_{u \in \mathbb{R}} \sum_{i=1}^{n} (y_i - u)^2. \tag{2.11}$$

If the conditional mean of y given x is linear and expressed as $\mu(x) = x^T\beta$, then β can be estimated by solving

$$\min_{\beta \in \mathbb{R}^p} \sum_{i=1}^{n} (y_i - x_i^T\beta)^2, \tag{2.12}$$

which is the ordinary least squares solution of linear regression model. Similarly, since the τth sample quantile solves the problem in (2.10), we are willing to specify the τth *conditional quantile function* as $Q_\tau(y|x) = x^T\beta_\tau$, and to obtain $\hat{\beta}_\tau$ by solving

$$\min_{\beta\in\mathbb{R}^p} \sum_{i=1}^{n} \rho_\tau(y_i - x_i^T\beta_\tau). \tag{2.13}$$

This is the germ of the idea elaborated by Koenker and Bassett (1978).

In linear regression, we can use the least squares technique to obtain a closed-form estimate. Let $X = [x_1, \ldots, x_n]^T$, to minimize

$$\|y - \hat{y}(\beta)\|^2 = (y - X\beta)^T(y - X\beta),$$

we differentiate with respect to β to receive the estimating equations

$$\frac{\partial}{\partial\beta}\|y - \hat{y}(\beta)\|^2 = X^T(y - X\beta) = 0$$

and solve for $\hat{\beta}$. The estimating equations yield a unique solution if the design matrix X has full column rank.

In quantile regression, however, when we proceed in the same way, we should be more cautious about the differentiation part. The piecewise linear and continuous objective function,

$$R(\beta_\tau) = \sum_{i=1}^{n} \rho_\tau(y_i - x_i^T\beta_\tau),$$

is differentiable except at the points at which one or more residuals, $y_i - x_i^T\beta_\tau$, are zero. At such points, we try to use a directional derivative of $R(\beta_\tau)$ in a certain direction w. The directional derivative is given by

$$\begin{aligned}\frac{\partial}{\partial\beta_\tau}R(\beta_\tau, w) &\equiv \frac{\partial}{\partial t}R(\beta_\tau + tw)|_{t=0} \\ &= \frac{\partial}{\partial t}\sum_{i=1}^{n}(y_i - x_i^T\beta_\tau - x_i^T tw)[\tau - I(y_i - x_i^T\beta_\tau - x_i^T tw < 0)]|_{t=0} \\ &= -\sum_{i=1}^{n}\psi_\tau^*(y_i - x_i^T\beta_\tau, -x_i^T w)x_i^T w,\end{aligned}$$

where

$$\psi_\tau^*(u, v) = \begin{cases} \tau - I(u < 0) & \text{if } u \neq 0 \\ \tau - I(v < 0) & \text{if } u = 0. \end{cases} \tag{2.14}$$

If all the directional derivatives are nonnegative at a point $\hat{\beta}_\tau$ (i.e., $\frac{\partial}{\partial\beta_\tau}R(\hat{\beta}_\tau, w) \geq 0$ for all $w \in \mathbb{R}^p$ which $\|w\|= 1$), then $\hat{\beta}_\tau$ minimizes $R(\beta_\tau)$.

2.2.3 Equivariance and Transformation

Equivariance properties are often treated as an important aid in interpreting statistical results. Let a τth regression quantile based on observations (y, X) be denoted by $\hat{\beta}_\tau(y, X)$. We collect some basic properties in the following result:

Proposition 2.2.1 (Koenker and Bassett, 1978)**.** *Let A be any $p \times p$ nonsingular matrix, $\gamma \in \mathbb{R}^p$, and $a > 0$. Then, for any $\tau \in [0, 1]$,*

1. $\hat{\beta}_\tau(ay, X) = a\hat{\beta}_\tau(y, X)$
2. $\hat{\beta}_\tau(-ay, X) = -a\hat{\beta}_{1-\tau}(y, X)$
3. $\hat{\beta}_\tau(y + X\gamma, X) = \hat{\beta}_\tau(y, X) + \gamma$
4. $\hat{\beta}_\tau(y, XA) = A^{-1}\hat{\beta}_\tau(y, X)$.

Proof of Proposition 2.2.1. Let

$$\begin{aligned}\Psi_\tau(\hat{\beta}, y, X) &= \sum_{\{i: y_i > x_i^T\hat{\beta}\}} \tau|y_i - x_i^T\hat{\beta}| + \sum_{\{i: y_i < x_i^T\hat{\beta}\}} (1-\tau)|y_i - x_i^T\hat{\beta}| \\ &= \sum_{i=1}^{n} [\tau - \frac{1}{2} + \frac{1}{2}\operatorname{sgn}(y_i - x_i^T\hat{\beta})][y_i - x_i^T\hat{\beta}]\end{aligned}$$

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

1. $a\Psi_\tau(\hat{\beta}, y, X) = \Psi_\tau(a\hat{\beta}, ay, X)$
2. $a\Psi_\tau(\hat{\beta}, y, X) = \Psi_{1-\tau}(-a\hat{\beta}, -ay, X)$
3. $\Psi_\tau(\hat{\beta}, y, X) = \Psi_\tau(\hat{\beta} + \gamma, y + X\gamma, X)$
4. $\Psi_\tau(\hat{\beta}, y, X) = \Psi_\tau(A^{-1}\hat{\beta}, y, XA)$.

□

Remark. Properties 1 and 2 imply a form of scale equivariance, property 3 is usually called shift or regression equivariance, and property 4 is called equivariance to reparameterization of design.

Equivariance to monotone transformations is another critical property to understand the full potential of the quantile regression. Let $h(\cdot)$ be a monotone function on $\mathbb{R}$. Then, for any random variable Y,

$$Q_\tau(h(Y)) = h(Q_\tau(Y)). \tag{2.15}$$

This property follows immediately from the elementary fact that, for any nondecreasing function h,

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

An example is that a conditional quantile of $\log(Y)$ is the log of the conditional quantile of $Y(>0)$:

$$Q_\tau(\log(Y|X)) = \log(Q_\tau(Y|X)),$$

and equivalently,

$$Q_\tau(Y|X) = e^{Q_\tau(\log(Y)|X)}.$$

This property is particularly important for research involving skewed distributions.

2.3 Quantile Regression Inference

We have talked about the parameter estimation in Section 2.2.2. In this section, some inferential statistics, standard errors and confidence intervals will be discussed. The goodness of fit and model checking will also be discussed.

2.3.1 Standard Errors and Confidence Intervals

The coefficients in the quantile regression model, β_τ, can be expressed in the form $Q_\tau(y_i|x_i) = \sum_{j=1}^{p} \beta_{j\tau} x_{ij}$ which has an equivalent form $y_i = \sum_{j=1}^{p} \beta_{j\tau} x_{ij} + \varepsilon_{i\tau}$, where the $\varepsilon_{i\tau}$'s have a distribution whose τth quantile is zero. In order to make inferences for β_τ, we wish to obtain some measure of standard error $S_{\hat{\beta}_{j\tau}}$ of $\hat{\beta}_{j\tau}$. This standard error can be used to construct confidence intervals and hypothesis tests. Moreover, standard error $S_{\hat{\beta}_{j\tau}}$ must satisfy the property that asymptotically, the quantity $(\hat{\beta}_{j\tau} - \beta_{j\tau})/S_{\hat{\beta}_{j\tau}}$ has a standard normal distribution.

Under the i.i.d model ($\varepsilon_{i\tau}$'s are independent and identically distributed), the asymptotic covariance matrix for $\hat{\beta}_\tau$ is simply described as

$$\Sigma_{\hat{\beta}_\tau} = \frac{\tau(1-\tau)}{n} \cdot \frac{1}{f_{\varepsilon_\tau}(0)^2}(X^T X)^{-1}. \tag{2.16}$$

In this expression, $f_{\varepsilon_\tau}(0)$ is the probability density of the error term ε_τ at point 0. This density term is unknown and needs to be estimated. To do this, we adapt the inverse density function: $1/f_{\varepsilon_\tau} = dQ_\tau(\varepsilon_\tau)/d\tau$, which can be approximated by $(\hat{Q}_{\tau+h} - \hat{Q}_{\tau-h})/2h$ for some small value of h. Notice that, the sample quantiles, $\hat{Q}_{\tau+h}$ and $\hat{Q}_{\tau-h}$, are based on the residuals $\hat{\varepsilon}_{i\tau} = y_i - \sum_{j=1}^{p} \hat{\beta}_{j\tau} x_{ij}$, $i = 1, \ldots, n$.

In the non-i.i.d case, we introduce a *weighted version* (see D_1 below) of the $X^T X$ matrix. Koenker (2005) gives a multivariate normal approximation to the joint distribution of the coefficient estimates $\hat{\beta}_{j\tau}$. This distribution has a mean with components that are the true coefficients and a covariance matrix of the form:

$$\Sigma_{\hat{\beta}_\tau} = \frac{\tau(1-\tau)}{n} D_1^{-1} D_0 D_1^{-1}$$

where

$$D_0 = \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} x_i^T x_i, \text{ and}$$
$$D_1 = \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} w_i x_i^T x_i.$$

Here x_i is a $1 \times p$ vector, and the terms D_0 and D_1 are $p \times p$ matrices. The weight for the ith subject is $w_i = f_{\varepsilon_{i\tau}}(0)$. Because the density function is unknown, the weights need to be estimated. Two methods of estimating w_i's are given by Koenker (2005). Once the weights are estimated, under some mild conditions, the covariance matrix for $\hat{\beta}_\tau$ can be expressed as

$$\hat{\Sigma}_{\hat{\beta}_\tau} = \frac{\tau(1-\tau)}{n} \hat{D}_1^{-1} \hat{D}_0 \hat{D}_1^{-1}, \tag{2.17}$$

where

$$\hat{D}_0 = \frac{1}{n}\sum_{i=1}^{n} x_i^T x_i, \text{ and}$$
$$\hat{D}_1 = \frac{1}{n}\sum_{i=1}^{n} \hat{w}_i x_i^T x_i.$$

The square root of the diagonal elements of the estimated covariance matrix $\hat{\Sigma}_{\hat{\beta}_\tau}$ give the corresponding estimated standard errors, $S_{\hat{\beta}_{j\tau}}$, for coefficient estimators $\hat{\beta}_{j\tau}$, $j = 1, \ldots, p$. Now, we are able to test hypotheses about the effects of the covariates on the dependent variable, and to construct confidence intervals for the quantile regression coefficients. The $1-\alpha$ confidence interval of $\beta_{j\tau}$ is

$$C.I: \quad \hat{\beta}_{j\tau} \pm z_{\alpha/2} S_{\hat{\beta}_{j\tau}}.$$

Let the null hypothesis be

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

then H_0 is rejected if $|(\hat{\beta}_{j\tau} - 0)/S_{\hat{\beta}_{j\tau}}| > z_{\alpha/2}$.

2.3.2 Goodness of Fit

We know that, in linear regression models, the goodness of fit can be described by the R-squared (the coefficient of determination) method:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}. \tag{2.18}$$

The term $\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$ denotes the sum of squared distance between the observed responses y_i and the corresponding values $\hat{y}_i$ predicted by the fitted model. While the $\sum_{i=1}^{n}(y_i - \bar{y})^2$ part is the sum of squared distances between observed responses and the mean value of these responses. R^2 ranges from 0 to 1, with a larger value indicating a better model fit.

In order to measure the goodness of fit in quantile regression, Koenker and Machado (1999) give an R_τ measure denoted as

$$R_\tau = 1 - \frac{V_{1\tau}}{V_{0\tau}}. \tag{2.19}$$

In the above expression,

$$\begin{aligned} V_{1\tau} &= \sum_{i=1}^{n} \rho_\tau(y_i - \hat{y}_i), \text{ and} \\ V_{0\tau} &= \sum_{i=1}^{n} \rho_\tau(y_i - \hat{Q}_\tau), \end{aligned}$$

where $\rho_\tau(u) = u(\tau - I(u < 0))$ following Koenker and Bassett (1978), and $\hat{Q}_\tau$ is the τth sample quantile. Because $V_{1\tau}$ and $V_{0\tau}$ are nonnegative, and $V_{1\tau} \leq V_{0\tau}$, so R_τ lies between 0 and 1. Also, a larger R_τ indicates a better model fit.

Chapter 3

Proposed Quantile Regression Method for Longitudinal Data

In many fields of applied studies, the observed data are often from subjects measured repeatedly over time. This kind of data with repeated responses is known as longitudinal data. Usually it is assumed that responses measured from different individuals are independent of each other, while the observations from the same individual are correlated. In this chapter, we introduce several methods to estimate quantile regression models built on longitudinal datasets.

3.1 The Development of Proposed Quantile Regression Method for Longitudinal Data

In a longitudinal setup, we collect a small number of repeated responses along with certain multidimensional covariates from a large number of independent individuals. Let $y_{i1}, \ldots, y_{ij}, \ldots, y_{in_i}$ be $n_i \geq 2$ repeated measures observed from the ith subject, for $i = 1, \ldots, m$ where m is the number of subjects. Let $x_{ij} = (x_{ij1}, \ldots, x_{ijp})^T$ be the p-dimensional covariate vector corresponding to y_{ij}. Suppose that responses from different individuals are independent and those from the same subject are dependent. Under the mean regression, this type of data is usually modeled by a linear relationship

$$y_i = X_i\beta + \epsilon_i, \tag{3.1}$$

where $y_i = (y_{i1}, \ldots, y_{ij}, \ldots, y_{in_i})^T$ is the vector of repeated responses and $X_i = [x_{i1}, \ldots, x_{in_i}]^T$ is the $n_i \times p$ matrix of covariates for the ith individual. The parameter $\beta = (\beta_1, \ldots, \beta_p)^T$ in equation (3.1) denote the effects of the components of x_{ij} on y_{ij}, and $\epsilon_i = (\epsilon_{i1}, \ldots, \epsilon_{ij}, \ldots, \epsilon_{in_i})^T$ is the n_i-dimensional residual vector such that

$$\epsilon_i \sim (0, \Sigma_i),$$

where for all $i = 1, \ldots, m$, ϵ_i are independently distributed (id) with a 0 mean vector and $n_i \times n_i$ covariance matrix Σ_i. Here, the main interest is generally in finding a consistent and efficient estimate of β. In the following sections, we will discuss quantile regression models.

3.1.1 A Working Independence Quantile Regression Model

In order to start the quantile regression for longitudinal data, we review what we have discussed in Chapter 2. The model for the conditional quantile functions of the response y_{ij} is given by

$$Q_\tau(y_{ij}|x_{ij}) = x_{ij}^T\beta_\tau, \tag{3.2}$$

for a particular τ. In quantile regression we are interested in estimating β_τ consistently and as efficiently as possible.

In the longitudinal dataset, if we assume the working independence (WI) between repeated responses among each individual, we can apply the method given in Chapter 2. When the working independence is assumed, the repeated measures from the same individual are not correlated any more. That is all the $K = n_1 +, \ldots, +n_m$ responses from all the individuals are treated as independent observations. In this case, by applying the minimization problem (2.13), we can obtain $\hat{\beta}_{WI\tau}$, an estimate of β_τ, with some loss of efficiency by minimizing the following objective function

$$S(\beta_\tau) = \sum_{i=1}^{m} \sum_{j=1}^{n_i} \rho_\tau(y_{ij} - x_{ij}^T \beta_\tau), \tag{3.3}$$

where $\rho_\tau(u) = u(\tau - I(u \leq 0))$ (Koenker and Bassett, 1978).

Estimating equations can be derived from function 3.3 by setting the differentiation of $S(\beta_\tau)$ with respect to β_τ to be 0. That is

$$\begin{aligned} U_0(\beta_\tau) &= \frac{\partial S(\beta_\tau)}{\partial \beta_\tau} \\ &= \sum_{i=1}^{m} \sum_{j=1}^{n_i} x_{ij} \psi_\tau(y_{ij} - x_{ij}^T \beta_\tau) \\ &= \sum_{i=1}^{m} X_i^T \psi_\tau(y_i - X_i \beta_\tau) \\ &= 0, \end{aligned} \tag{3.4}$$

where $X_i = [x_{i1}, \ldots, x_{in_i}]^T$ is the $n_i \times p$ matrix of covariates, $y_i = (y_{i1}, \ldots, y_{in_i})^T$ is the $n_i \times 1$ vector of the variable of repeated measures for the ith individual, $\psi_\tau(u) = \rho'_\tau(u) = \tau - I(u < 0)$ is a discontinuous function, and $\psi_\tau(y_i - X_i\beta_\tau) = (\psi_\tau(y_{i1} - x_{i1}^T\beta_\tau), \ldots, \psi_\tau(y_{in_i} - x_{in_i}^T\beta_\tau))^T$ is a $n_i \times 1$ vector.

An efficient algorithm to obtain an estimate of β_τ by solving the equation (3.4), $U_0(\beta_\tau) = 0$, was given by Koenker and D'Orey (1987), which is available in statistical software R (package "quantreg"). Parameter estimators $\hat{\beta}_{WI\tau}$ are derived from estimating equations 3.4 under working independence assumption, therefore the efficiency of $\hat{\beta}_{WI\tau}$ may not be satisfactory. In order to obtain more efficient estimators, the correlations within repeated responses need to be taken into account.

3.1.2 Quasi-likelihood for Quantile Regression

To take the within correlations into the consideration when constructing quantile regression models for longitudinal data, a quasi-likelihood (QL) method was introduced by Jung (1996). Let $\varepsilon_i = (\varepsilon_{i1}, \ldots, \varepsilon_{ij}, \ldots, \varepsilon_{in_i})^T$, where $\varepsilon_{ij} = y_{ij} - x_{ij}^T \beta_\tau$ which is a continuous error term satisfying $P(\varepsilon_{ij} \leq 0) = \tau$ and with an unknown density function $f_{ij}(\cdot)$. In mean regression model, Bernoulli distributed $\psi_\tau(\varepsilon_i) = \psi_\tau(y_i - X_i\beta_\tau)$ can be treated as a random noise vector. Using this fact we may be able to generalize the QL to quantile regression by estimating the correlation matrix of $\psi_\tau(\varepsilon_i)$.

Let the covariance matrix of $\psi_\tau(\varepsilon_i)$ be denoted as

$$
\begin{aligned}
V_i &= \operatorname{cov}(\psi_\tau(y_i - X_i\beta_\tau)) \\
&= \operatorname{cov}\begin{pmatrix} \tau - I(\varepsilon_{i1} < 0) \\ \vdots \\ \tau - I(\varepsilon_{in_i} < 0) \end{pmatrix},
\end{aligned}
\tag{3.5}
$$

and

$$
\begin{aligned}
\Gamma_i &= \operatorname{diag}[f_{i1}(0), \ldots, f_{in_i}(0)] \\
&= \begin{pmatrix} f_{i1}(0) & & \\ & \ddots & \\ & & f_{in_i}(0) \end{pmatrix},
\end{aligned}
\tag{3.6}
$$

be an $n_i \times n_i$ diagonal matrix with jth diagonal element $f_{ij}(0)$.

Jung (1996) derived the derivative of the log-quasi-likelihood $l(\beta_\tau; y_i)$ with respect to β_τ, which can be used to estimate β_τ and written as

$$\frac{\partial l(\beta_\tau; y_i)}{\partial \beta_\tau} = X_i^T \mathit{\Gamma}_i V_i^{-1} \psi_\tau (y_i - X_i \beta_\tau). \tag{3.7}$$

Let $U_1(\beta_\tau) = \sum_{i=1}^m (\partial l(\beta_\tau; y_i)/\partial \beta_\tau)$. By solving the equation

$$U_1(\beta_\tau) = \sum_{i=1}^m X_i^T \mathit{\Gamma}_i V_i^{-1} \psi_\tau (y_i - X_i \beta_\tau) = 0 \tag{3.8}$$

the so-called maximum quasi-likelihood estimate ,$\hat{\beta}_\tau$, would be obtained.

In the estimating equation $U_1(\beta_\tau) = 0$, the term $\mathit{\Gamma}_i$ describes the dispersions in ε_{ij} and its diagonal elements can be well estimated by following Hendricks and Koenker (1992):

$$\hat{f}_{ij}(0) = 2h_n [x_{ij}^T (\hat{\beta}_{\tau+h_n} - \hat{\beta}_{\tau-h_n})]^{-1}, \tag{3.9}$$

where $h_n \to 0$ when $n \to \infty$ is a bandwidth parameter. In some cases when f_{ij} is difficult to estimate, $\mathit{\Gamma}_i$ can be simply treated as an identity matrix, with a slight loss of efficiency (Jung 1996).

However, the estimation of the covariance matrix V_i becomes much complicated when QL method is applied. Whatever correlation matrix that ε_i follows, the correlation matrix of $\psi_\tau(\varepsilon_i)$ is no longer the same one, and its correlation structure may be very difficult to specify. An easy way to avoid these difficulties is to assume working independence under the method of quasi-likelihood. When the repeated responses are assumed independent, so as the corresponding $\psi_\tau(\varepsilon_i)$. Hence, the matrix V_i becomes an $n_i \times n_i$ diagonal matrix with the jth element $\sigma_{ijj} = \text{var}(\psi_\tau(\varepsilon_{ij}))$. Using this type of V_i in equation (3.8) gives us a quasi-likelihood working independence (QLWI) model. Apparently, without considering the correlation within $\psi_\tau(\varepsilon_i)$, estimates from QLWI models may not be efficient.

3.1.3 The Proposed Quantile Regression Model

In order to obtain efficient estimates from the estimating equation (3.8), a better way to estimate covariance matrix of $\psi_\tau(\varepsilon_i)$ must be applied. Here, by taking the correlations between $\psi_\tau(\varepsilon_i)$ into consideration, we propose a new method solving the following estimating equations

$$U(\beta_\tau) = \sum_{i=1}^{m} X_i^T \Gamma_i \Sigma_i^{-1}(\rho)\psi_\tau(y_i - X_i\beta_\tau) = 0, \tag{3.10}$$

where $\Sigma_i(\rho)$ is the covariance matrix of $\psi_\tau(\varepsilon_i)$ that can be expressed as $\Sigma_i(\rho) = A_i^{\frac{1}{2}} C_i(\rho) A_i^{\frac{1}{2}}$, with $A_i = \text{diag}[\sigma_{i11}, \ldots, \sigma_{1n_in_i}]$ being an $n_i \times n_i$ diagonal matrix, $\sigma_{ijj} = \text{var}(\psi_\tau(\varepsilon_{ij}))$ and $C_i(\rho)$ as the correlation matrix of $\psi_\tau(\varepsilon_i)$, ρ being a correlation index parameter.

Suppose that the covariance matrix $\Sigma_i(\rho)$ in estimating equation (3.10) has a general stationary auto-correlation structure such that the correlation matrix $C_i(\rho)$ is given by

$$C_i(\rho) = \begin{pmatrix} 1 & \rho_1 & \rho_2 & \cdots & \rho_{n_i-1} \\ \rho_1 & 1 & \rho_1 & \cdots & \rho_{n_i-2} \\ \vdots & \vdots & \vdots & & \vdots \\ \rho_{n_i-1} & \rho_{n_i-2} & \rho_{n_i-3} & \cdots & 1 \end{pmatrix} \tag{3.11}$$

for all $i = 1, \ldots, m$, where ρ_ℓ can be estimated by

$$\hat{\rho}_\ell = \frac{\sum_{i=1}^{m} \sum_{j=1}^{n_i-\ell} \tilde{y}_{ij}\tilde{y}_{i,j+\ell}/m(n_i-\ell)}{\sum_{i=1}^{m} \sum_{j=1}^{n_i} \tilde{y}_{ij}^2/mn_i} \tag{3.12}$$

for $\ell = 1, \ldots, n_i - 1$ (Sutradhar and Kovacevic (2000, Eq. (2.18)), Sutradhar, 2003) with $\tilde{y}_{ij}$ defined as

$$\tilde{y}_{ij} = \frac{\psi_\tau(y_{ij} - x_{ij}^T\beta_\tau)}{\sqrt{\sigma_{ijj}}} \tag{3.13}$$

.

Now, the only thing unknown, except β_τ, is the variance of $\psi_\tau(\varepsilon_{ij})$. To estimate $\sigma_{ijj} = \text{var}(\psi_\tau(y_{ij} - x_{ij}^T\beta_\tau))$, we apply the fact that $\psi_\tau(\varepsilon_{ij}) = \psi_\tau(y_{ij} - x_{ij}^T\beta_\tau) = \tau - I(y_{ij} < x_{ij}^T\beta_\tau)$. Hence we have

$$\begin{aligned}\sigma_{ijj} &= \text{var}[\psi_\tau(\varepsilon_{ij})] \\ &= \text{var}[\tau - I(y_{ij} < x_{ij}^T\beta_\tau)] \\ &= \text{var}[I(y_{ij} < x_{ij}^T\beta_\tau)] \\ &= \Pr(y_{ij} < x_{ij}^T\beta_\tau)(1 - \Pr(y_{ij} < x_{ij}^T\beta_\tau)),\end{aligned} \tag{3.14}$$

where $\Pr(y_{ij} < x_{ij}^T\beta_\tau)$ is the probability of the event $\{y_{ij} < x_{ij}^T\beta_\tau\}$. If β_τ is the true parameter, we know that $x_{ij}^T\beta_\tau$ is exactly the τth quantile of the variable y_{ij}, hence $\Pr(y_{ij} < x_{ij}^T\beta_\tau) = \tau$, which leads to an estimator of σ_{ijj},

$$\tilde{\sigma}_{ijj} = \tau(1-\tau).$$

Consequently, A_i matrix can be estimated at the true β_τ as

$$\begin{aligned}\tilde{A}_i &= \text{diag}[\tilde{\sigma}_{i11}, \ldots, \tilde{\sigma}_{1n_in_i}] \\ &= \begin{pmatrix} \tau(1-\tau) & & \\ & \ddots & \\ & & \tau(1-\tau) \end{pmatrix}_{n_i \times n_i},\end{aligned} \tag{3.15}$$

and an estimator of the square root of A_i immediately follows as

$$\tilde{A}_i^{\frac{1}{2}} = \begin{pmatrix} \sqrt{\tau(1-\tau)} & & \\ & \ddots & \\ & & \sqrt{\tau(1-\tau)} \end{pmatrix}_{n_i \times n_i} \tag{3.16}$$

indicating constant diagonal matrices for a certain τ.

We can use the same method to estimate Γ_i as being discussed in QL models. Therefore, we can now obtain the proposed estimates by solving the estimating equations (3.10) where the estimator $\tilde{\sigma}_{ijj}$ is applied in estimating Σ_i. And we denote the parameter estimator obtained from the proposed quantile regression (PQR) model as $\hat{\beta}_{PQR\tau}$.

Notice that in the expression $\Sigma_i(\rho) = A_i^{\frac{1}{2}} C_i(\rho) A_i^{\frac{1}{2}}$, if we set A_i as the one at the true β_τ which is given by (3.15), $C_i(\rho)$ becomes the only part in $\Sigma_i(\rho)$ containing the information about the dataset and the parameter β_τ. However in practice, the estimated parameter may never exactly equal the true β_τ. Thus, the elements of the diagonal matrix A_i may differ from the constant value $\tau(1-\tau)$. Moreover, we expect matrix A_i to be also affected by the dataset and the parameter estimates, which becomes crucial when we use an iteration method to estimate parameters where the estimates $\hat{\beta}_\tau$ need to be updated within each iteration step. In this case, as long as the sample size is large enough, we can estimate the diagonal elements of A_i by the following

$$\begin{aligned} \hat{\sigma}_{ijj} &= \Pr(y_{ij} < x_{ij}^T \beta_\tau)(1 - \Pr(y_{ij} < x_{ij}^T \beta_\tau)) \\ &= \frac{1}{m} \sum_{i=1}^{m} I(y_{ij} < x_{ij}^T \beta_\tau)(1 - \frac{1}{m} \sum_{i=1}^{m} I(y_{ij} < x_{ij}^T \beta_\tau)), \end{aligned} \tag{3.17}$$

for all $j = 1, \ldots, n_i$ and $i = 1, \ldots, m$. By using $\hat{\sigma}_{ijj}$ in estimating Σ_i, the solution of estimating equations (3.10) gives an adjusted estimate of β_τ. We call this method as adjusted quantile regression (AQR).

3.2 Estimation of Parameter and Its Covariance Matrix

The difficulty of solving the estimating equation (3.10) is that the non-continuous and non-convex objective function $U(\beta_\tau)$ can not be differentiated. Though several methods can be applied to estimate β_τ from equation (3.10) without requiring any derivatives and continuity of the estimating function, they may become very complicated and cause a high burden of computation. Furthermore, if re-sampling and perturbing methods were used to estimate the covariance matrix of parameter estimators, numerical computational problems would be produced.

To overcome these difficulties, the induced smoothing method has been extended to the quantile regression for longitudinal data assuming a working correlation by Fu and Wang (2012). Here, let $\tilde{U}(\beta_\tau) = E_Z[U(\beta_\tau + \Omega^{1/2}Z)]$, with expectation taken over Z, where $Z \sim N(0, I_p)$, and Ω is updated as an estimate of the covariance matrix of parameter estimators. After some algebra calculations, a smoothed estimating function $\tilde{U}(\beta_\tau)$ is obtained as

$$\tilde{U}(\beta_\tau) = \sum_{i=1}^{m} X_i^T \Gamma_i \Sigma_i^{-1}(\rho) \tilde{\psi}_\tau (y_i - X_i \beta_\tau), \tag{3.18}$$

where

$$\tilde{\psi}_\tau = \begin{pmatrix} \tau - 1 + \Phi(\frac{y_{i1} - x_{i1}^T \beta_\tau}{r_{i1}}) \\ \vdots \\ \tau - 1 + \Phi(\frac{y_{in_i} - x_{in_i}^T \beta_\tau}{r_{in_i}}) \end{pmatrix}$$

and $r_{ij} = \sqrt{x_{ij}^T \Omega x_{ij}}$ for $j = 1, \ldots, n_i$. Thus the differentiation of $\tilde{U}(\beta_\tau)$ with respect to β_τ can be easily calculated, and we can use $\partial\tilde{U}(\beta_\tau)/\partial\beta_\tau$ as an approximation of $\partial U(\beta_\tau)/\partial\beta_\tau$ as

$$\frac{\partial\tilde{U}(\beta_\tau)}{\partial\beta_\tau} = -\sum_{i=1}^{m} X_i^T \Gamma_i \Sigma_i^{-1}(\rho)\tilde{\Lambda}_i X_i, \tag{3.19}$$

where $\tilde{\Lambda}_i$ is an $n_i \times n_i$ diagonal matrix with the jth diagonal element $\frac{1}{r_{ij}}\phi((y_{ij} - x_{ij}^T\beta_\tau)/r_{ij})$.

Generally, let $\hat{\beta}_{WI\tau}$ be the estimate under the working independence assumption and I_p be a identity matrix of size p, smoothed estimators of β_τ and its covariance matrix Ω can be obtained by using the following Newton-Raphson iteration:

Step 1. Given initial values of β_τ and the symmetric positive definite matrix Ω as $\tilde{\beta}_\tau(0) = \hat{\beta}_{WI\tau}$ and $\tilde{\Omega}(0) = \frac{1}{m}I_p$ respectively.

Step 2. Using $\tilde{\beta}_\tau(r)$ and $\tilde{\Omega}(r)$ given from the rth iteration, update $\tilde{\beta}_\tau(r+1)$ and $\tilde{\Omega}(r+1)$ by

$$\begin{aligned}
\tilde{\beta}_\tau(r+1) &= \tilde{\beta}_\tau(r) + \left[-\frac{\partial\tilde{U}(\beta_\tau)}{\partial\beta_\tau}\right]_r^{-1} \times \left[\tilde{U}(\beta_\tau)\right]_r \\
&= \tilde{\beta}_\tau(r) + \left[\sum_{i=1}^{m} X_i^T \Gamma_i \Sigma_i^{-1}(\rho)\tilde{\Lambda}_i X_i\right]_r^{-1} \\
&\quad \times \left[\sum_{i=1}^{m} X_i^T \Gamma_i \Sigma_i^{-1}(\rho)\tilde{\psi}_\tau(y_i - X_i\beta_\tau)\right]_r,
\end{aligned}$$

and

$$\begin{aligned}
\tilde{\Omega}(r+1) &= \left[-\frac{\partial\tilde{U}(\beta_\tau)}{\partial\beta_\tau}\right]_r^{-1} \times \left[\mathrm{cov}(\tilde{U}(\beta_\tau))\right]_r \times \left[-\frac{\partial\tilde{U}(\beta_\tau)}{\partial\beta_\tau}\right]_r^{-1} \\
&= \left[\sum_{i=1}^{m} X_i^T \Gamma_i \Sigma_i^{-1}(\rho)\tilde{\Lambda}_i X_i\right]_r^{-1} \\
&\quad \times \left[\sum_{i=1}^{m} X_i^T \Gamma_i \Sigma_i^{-1}(\rho)\tilde{\psi}_\tau(y_i - X_i\beta_\tau)\tilde{\psi}_\tau^T(y_i - X_i\beta_\tau)\Sigma_i^{-1}(\rho)\Gamma_i X_i\right]_r \\
&\quad \times \left[\sum_{i=1}^{m} X_i^T \Gamma_i \Sigma_i^{-1}(\rho)\tilde{\Lambda}_i X_i\right]_r^{-1},
\end{aligned}$$

where $[]_r$ denotes that the expression between the square brackets is evaluated at $\beta_\tau = \tilde{\beta}_\tau(r)$.

Step 3. Repeat step 2 until convergence.

This method provides accurate estimates of β_τ and its covariance matrix Ω. Furthermore, compared with other techniques, Newton-Raphson method is much faster.

3.3 Asymptotic Properties

In this section, we derive asymptotic distributions of the proposed estimates obtained by solving both the non-smoothed and smoothed estimate equations (3.10), $U(\hat{\beta}_\tau) = 0$, and $\tilde{U}(\tilde{\beta}_\tau) = 0$, where $\tilde{U}(\tilde{\beta}_\tau)$ is defined in formula (3.18). The following regularity conditions are required:

A1. For each i, the number of repeated measures n_i (≥ 2) is bounded and the dimension p of covariates x_{ij} is fixed. The cumulative distribution functions $F_{ij}(z) = P(y_{ij} - x_{ij}^T\beta_\tau \leq z | x_{ij})$ are absolutely continuous, with continuous densities f_{ij} and its first derivative being uniformly bounded away from 0 and ∞ at the point 0, $i = 1, \ldots, m$; $j = 1, \ldots, n_i$.

A2. The true value β_τ is an interior point of a bounded convex region $\mathfrak{B}$.

A3. Each x_i satisfies the following conditions

(a) For any positive definite matrix W_i, $\frac{1}{m}\sum_{i=1}^m X_i^T W_i \Gamma_i X_i$ converges to a positive definite matrix; where Γ_i is an $n_i \times n_i$ diagonal matrix with the jth diagonal element $f_{ij}(0)$.

(b) $\sup_i \|x_i\| < +\infty$.

A4. Matrix Ω is positive definite and $\Omega = O(\frac{1}{m})$.

A5. The differentiation of negative $\tilde{U}(\beta_\tau)$, $-\partial\tilde{U}(\beta_\tau)/\partial\beta_\tau$, is positive definite with probability 1.

3.3.1 Consistency and Asymptotic Normality of Estimators without Smoothing

In this section, we will show the consistency and asymptotic normality of estimators without smoothing.

Theorem 3.3.1. *Under regularity conditions A1-A5 listed above, the non-smoothed estimator $\hat{\beta}_\tau$ is $\sqrt{m}$-consistent and asymptotically normal,*

$$\sqrt{m}(\hat{\beta}_\tau - \beta_\tau) \to N(0, G^{-1}(\beta_\tau)V\{G^{-1}(\beta_\tau)\}^T),$$

where in the covariance matrix, $G(\beta_\tau) = \lim_{m\to\infty} \frac{1}{m}\sum_{i=1}^m X_i^T \mathit{\Gamma}_i \Sigma_i^{-1}(\rho)\mathit{\Gamma}_i X_i$ *and* $V = \lim_{m\to\infty} \frac{1}{m}\sum_{i=1}^m X_i^T \mathit{\Gamma}_i \Sigma_i^{-1}(\rho)\operatorname{cov}\{\psi_\tau(y_i - X_i\beta_\tau)\}\Sigma_i^{-1}(\rho)\mathit{\Gamma}_i X_i$.

Proof of Theorem 3.3.1. Let $H_i^T = X_i^T \mathit{\Gamma}_i \Sigma_i^{-1}(\rho)$ and $\psi_i = \psi_\tau(y_i - X_i\hat{\beta}_\tau)$, therefore $U(\hat{\beta}_\tau) = \sum_{i=1}^m H_i^T \psi_i$. Let $\bar{U}(\hat{\beta}_\tau) = \sum_{i=1}^m H_i^T \varphi_i$, where $\varphi_i = (\tau - P(y_{i1} - x_{i1}^T\hat{\beta}_\tau \leq 0), \ldots, \tau - P(y_{in_i} - x_{in_i}^T\hat{\beta}_\tau \leq 0))^T$. We can obtain

$$\begin{aligned}
\frac{1}{m}(U(\hat{\beta}_\tau) - \bar{U}(\hat{\beta}_\tau)) &= \frac{1}{m}\sum_{i=1}^m H_i^T(\psi_i - \varphi_i) \\
&= \frac{1}{m}\sum_{i=1}^m H_i^T \begin{pmatrix} P(y_{i1} - x_{i1}^T\hat{\beta}_\tau \leq 0) - I(y_{i1} - x_{i1}^T\hat{\beta}_\tau \leq 0) \\ \vdots \\ P(y_{in_i} - x_{in_i}^T\hat{\beta}_\tau \leq 0) - I(y_{in_i} - x_{in_i}^T\hat{\beta}_\tau \leq 0) \end{pmatrix} \\
&= \frac{1}{m}\sum_{i=1}^m \sum_{j=1}^{n_i} h_{ij}[P(y_{ij} - x_{ij}^T\hat{\beta}_\tau \leq 0) - I(y_{ij} - x_{ij}^T\hat{\beta}_\tau \leq 0)],
\end{aligned}$$

where h_{ij} is a $p \times 1$ vector and $(h_{i1}, \ldots, h_{in_i}) = H_i^T$. According to the uniform strong law of large numbers (Pollard 1990), under condition A3 we have

$$\sup_{\hat{\beta}_\tau \in \mathfrak{B}} \left| \frac{1}{m} \sum_{i=1}^{m} \sum_{j=1}^{n_i} h_{ij} [P(y_{ij} - x_{ij}^T \hat{\beta}_\tau \leq 0) - I(y_{ij} - x_{ij}^T \hat{\beta}_\tau \leq 0)] \right| = o(m^{-1/2}) \quad \text{a.s..}$$

Therefore,

$$\sup_{\hat{\beta}_\tau \in \mathfrak{B}} \| \frac{1}{m} (U(\hat{\beta}_\tau) - \bar{U}(\hat{\beta}_\tau)) \| = o(m^{-1/2}) \quad \text{a.s..}$$

Now,

$$G_m(\beta_\tau) = -\frac{1}{m} \frac{\partial \bar{U}(\hat{\beta}_\tau)}{\partial \hat{\beta}_\tau} \bigg|_{\hat{\beta}_\tau = \beta_\tau} = \frac{1}{m} \sum_{i=1}^{m} H_i^T \mathit{\Gamma}_i X_i$$

is positive definite and, with probability 1, $G_m(\beta_\tau) \to G(\beta_\tau)$ when $m \to +\infty$. Because $P(y_{ij} - x_{ij}^T \beta_\tau \leq 0) = \tau$, β_τ is the unique solution of the equation $\bar{U}(\hat{\beta}_\tau)$. Together with $U(\hat{\beta}_\tau) = 0$ and condition A3, implies that $\hat{\beta}_\tau \to \beta_\tau$ as $m \to \infty$.

Because ψ_i are independent random variables with mean zero, and $\text{var}\{U(\beta_\tau)/m\} = \frac{1}{m} \sum_{i=1}^{m} X_i^T \mathit{\Gamma}_i \Sigma_i^{-1}(\rho) \operatorname{cov}(\psi_i) \Sigma_i^{-1}(\rho) \mathit{\Gamma}_i X_i$, the multivariate central limit theorem implies that $\frac{1}{\sqrt{m}} U(\beta_\tau) \to N(0, V)$.

For any $\hat{\beta}_\tau$ satisfying $\|\hat{\beta}_\tau - \beta_\tau\| < cm^{-1/3}$,

$$\begin{aligned} U(\hat{\beta}_\tau) - U(\beta_\tau) &= \sum_{i=1}^{m} H_i^T(\hat{\beta}_\tau) \psi_i(\hat{\beta}_\tau) - \sum_{i=1}^{m} H_i^T(\beta_\tau) \psi_i(\beta_\tau) \\ &= \sum_{i=1}^{m} H_i^T(\hat{\beta}_\tau) \{\psi_i(\hat{\beta}_\tau) - \psi_i(\beta_\tau)\} + \sum_{i=1}^{m} \{H_i^T(\hat{\beta}_\tau) - H_i^T(\beta_\tau)\}^T \psi_i(\beta_\tau). \end{aligned}$$

The first term can be written as

$$\begin{aligned} &\sum_{i=1}^{m} H_i^T(\hat{\beta}_\tau) \{\psi_i(\hat{\beta}_\tau) - \psi_i(\beta_\tau)\} \\ &\quad = \sum_{i=1}^{m} H_i^T(\hat{\beta}_\tau) \varphi_i(\hat{\beta}_\tau) + \sum_{i=1}^{m} H_i^T(\hat{\beta}_\tau) \{\psi_i(\hat{\beta}_\tau) - \psi_i(\beta_\tau) - \varphi_i(\hat{\beta}_\tau)\} \\ &\quad = \sum_{i=1}^{m} H_i^T(\hat{\beta}_\tau) \varphi_i(\hat{\beta}_\tau) + \sum_{i=1}^{m} H_i^T(\hat{\beta}_\tau) \{P(y_{ij} - x_{ij}^T \hat{\beta}_\tau \leq 0) - I(y_{ij} - x_{ij}^T \hat{\beta}_\tau \leq 0) \\ &\qquad + I(y_{ij} - x_{ij}^T \beta_\tau \leq 0) - \tau\} \end{aligned}$$

The Lemma in Jung (1996) tells us that

$$\sup\left|\sum_{i=1}^{m} H_i^T(\hat{\beta}_\tau)\{P(y_{ij} - x_{ij}^T\hat{\beta}_\tau \leq 0) - I(y_{ij} - x_{ij}^T\hat{\beta}_\tau \leq 0) + I(y_{ij} - x_{ij}^T\beta_\tau \leq 0) - \tau\}\right|$$
$$= o_p(\sqrt{m}).$$

Therefore,

$$\begin{aligned}\sum_{i=1}^{m} H_i^T(\hat{\beta}_\tau)\{\psi_i(\hat{\beta}_\tau) - \psi_i(\beta_\tau)\} &= \sum_{i=1}^{m} H_i^T(\hat{\beta}_\tau)\varphi_i(\hat{\beta}_\tau) + o_p(\sqrt{m}) \\ &= \bar{U}(\hat{\beta}_\tau) + o_p(\sqrt{m})\end{aligned}$$

From the law of large numbers, the second term

$$\begin{aligned}\sum_{i=1}^{m}\{H_i^T(\hat{\beta}_\tau) - H_i^T(\beta_\tau)\}^T\psi_i(\beta_\tau) &= \sum_{i=1}^{m}\sum_{j=1}^{n_i}(h_{ij}(\hat{\beta}_\tau) - h_{ij}(\beta_\tau))[P(y_{ij} - x_{ij}^T\beta_\tau \leq 0) \\ &\quad - I(y_{ij} - x_{ij}^T\beta_\tau \leq 0)] \\ &= o_p(\sqrt{m}).\end{aligned}$$

Hence, $U(\hat{\beta}_\tau) - U(\beta_\tau) = \bar{U}(\hat{\beta}_\tau) + o_p(\sqrt{m})$. Using Taylor's expansion of $\bar{U}(\hat{\beta}_\tau)$ around β_τ, we have

$$\frac{1}{\sqrt{m}}\{U(\hat{\beta}_\tau) - U(\beta_\tau)\} = \frac{1}{m}\frac{\partial \bar{U}(\hat{\beta}_\tau)}{\partial \hat{\beta}_\tau}\bigg|_{\hat{\beta}_\tau = \beta_\tau}\sqrt{m}(\hat{\beta}_\tau - \beta_\tau) + o_p(1).$$

Because $\hat{\beta}_\tau$ is in the $m^{-1/3}$ neighborhood of β_τ and $U(\hat{\beta}_\tau) = 0$, we have

$$\sqrt{m}(\hat{\beta}_\tau - \beta_\tau) = G_m^{-1}(\beta_\tau)\frac{1}{\sqrt{m}}U(\beta_\tau) + o_p(1).$$

Therefore, we showed that $\sqrt{m}(\hat{\beta}_\tau - \beta_\tau) \rightarrow N(0, G^{-1}(\beta_\tau)V\{G^{-1}(\beta_\tau)\}^T)$ as $m \rightarrow +\infty$. □

3.3.2 Consistency and Asymptotic Normality of Estimators with Smoothing

Lemma 3.3.1. *Under some regularity conditions, the smoothed estimating functions* $\tilde{U}(\beta_\tau)$ *are equivalent to the non-smoothed estimating functions* $U(\beta_\tau)$*,*

$$\frac{1}{\sqrt{m}}\{\tilde{U}(\beta_\tau) - U(\beta_\tau)\} = o_p(1).$$

Proof of Lemma 3.3.1. Let $\psi_{ij} = \psi_\tau(y_{ij} - x_{ij}^T\beta_\tau)$, $\tilde{\psi}_{ij} = \tilde{\psi}_\tau(y_{ij} - x_{ij}^T\beta_\tau)$ and $d_{ij} = \varepsilon_{ij}/r_{ij}$, where $\varepsilon_{ij} = y_{ij} - x_{ij}^T\beta_\tau$, $r_{ij} = \sqrt{x_{ij}^T\Omega x_{ij}}$. Since $\tilde{\psi}_{ij} - \psi_{ij} = \mathrm{sgn}(-d_{ij})\Phi(-|d_{ij}|)$, where $\mathrm{sgn}(\cdot)$ is the sign function, we have

$$\begin{aligned}\frac{1}{\sqrt{m}}\{\tilde{U}(\beta_\tau) - U(\beta_\tau)\} &= \frac{1}{\sqrt{m}}\sum_{i=1}^{m} X_i^T\Gamma_i\Sigma_i^{-1}(\rho)\begin{pmatrix}\mathrm{sgn}(-d_{i1})\Phi(-|d_{i1}|)\\ \vdots \\ \mathrm{sgn}(-d_{in_i})\Phi(-|d_{in_i}|)\end{pmatrix}\\ &= \frac{1}{\sqrt{m}}\sum_{i=1}^{m}\sum_{j=1}^{n_i} z_{ij}\,\mathrm{sgn}(-d_{ij})\Phi(-|d_{ij}|),\end{aligned}$$

where z_{ij} is the jth column of $X_i^T\Gamma_i\Sigma_i^{-1}(\rho)$. Because

$$\begin{aligned}E(\tilde{\psi}_{ij} - \psi_{ij}) &= \int_{-\infty}^{+\infty}\mathrm{sgn}(-d_{ij})\Phi(-|d_{ij}|)f_{ij}(\varepsilon)d\varepsilon\\ &= \int_{-\infty}^{+\infty}\Phi(-|\varepsilon|/r_{ij})\{2I(\varepsilon \le 0) - 1\}f_{ij}(\varepsilon)d\varepsilon\\ &= r_{ij}\int_{-\infty}^{+\infty}\Phi(-|t|)\{2I(t \le 0) - 1\}[f_{ij}(0) + f_{ij}'(\zeta(t))r_{ij}t]dt,\end{aligned}$$

where $\zeta(t)$ is between 0 and $r_{ij}t$. Because $\int_{-\infty}^{+\infty}\Phi(-|t|)\{2I(t \le 0) - 1\}dt = 0$, we have $r_{ij}\int_{-\infty}^{+\infty}\Phi(-|t|)\{2I(t \le 0) - 1\}f_{ij}(0)dt = 0$. Since $\int_{-\infty}^{+\infty}|t|\Phi(-|t|)dt = 1/2$, and by condition A1, there exists a constant M such that $\sup_{ij}|f_{ij}^{'}(\zeta(t))|\leq M$. Therefore,

$$\begin{aligned}|E(\tilde{\psi}_{ij}-\psi_{ij})| &\leq r_{ij}^2\int_{-\infty}^{+\infty}|t|\Phi(-|t|)|f_{ij}^{'}(\zeta(t))|dt\\ &\leq Mr_{ij}^2/2.\end{aligned}$$

Under regularity conditions A3 and A4, when $m\to+\infty$,

$$\left\|\frac{1}{\sqrt{m}}E\{\tilde{U}(\beta_\tau)-U(\beta_\tau)\}\right\|\leq\frac{1}{\sqrt{m}}\sup_{i,j}|z_{ij}|\sum_{i=1}^{m}Mr_{ij}^2/2=o(1).$$

Moreover,

$$\frac{1}{m}\operatorname{var}\{\tilde{U}(\beta_\tau)-U(\beta_\tau)\}=\frac{1}{m}\sum_{i=1}^{m}\operatorname{var}\left\{\sum_{j=1}^{n_i}z_{ij}\operatorname{sgn}(-d_{ij})\Phi(-|d_{ij}|)\right\}.$$

By Cauchy-Schwartz inequality,

$$\begin{aligned}\frac{1}{m}\operatorname{var}\{\tilde{U}(\beta_\tau)-U(\beta_\tau)\}\leq{}&\frac{1}{m}\sum_{i=1}^{m}\sum_{j=1}^{n_i}z_{ij}z_{ij}^T\operatorname{var}(\tilde{\psi}_{ij}-\psi_{ij})\\ &+\frac{1}{m}\sum_{i=1}^{m}\sum_{j=1}^{n_i}\sum_{k\neq j}^{n_i}z_{ij}z_{ik}^T\sqrt{\operatorname{var}(\tilde{\psi}_{ij}-\psi_{ij})\operatorname{var}(\tilde{\psi}_{ik}-\psi_{ik})}.\end{aligned}$$

Hence for each $j=1,\ldots,n_i$,

$$\begin{aligned}\operatorname{var}(\tilde{\psi}_{ij}-\psi_{ij})\leq E(\tilde{\psi}_{ij}-\psi_{ij})^2&=\int_{-\infty}^{+\infty}\{\operatorname{sgn}(-d_{ij})\Phi(-|d_{ij}|)\}^2f_{ij}(\varepsilon)d\varepsilon\\ &=r_{ij}\int_{-\infty}^{+\infty}\Phi^2(-|t|)f_{ij}(r_{ij}t)dt\\ &=r_{ij}\int_{|t|>\Delta}\Phi^2(-|t|)f_{ij}(r_{ij}t)dt+r_{ij}\int_{|t|\leq\Delta}\Phi^2(-|t|)f_{ij}(r_{ij}t)dt\\ &\leq\Phi^2(-\Delta)+r_{ij}\Delta f_{ij}(\zeta),\end{aligned}$$

where Δ is a positive value, and ζ is in the interval $(-r_{ij}\Delta,r_{ij}\Delta)$. Let $\Delta=m^{1/3}$. Under condition A4, because $r_{ij}=O(m^{-1/2})$, we have $r_{ij}\Delta=O(m^{-1/6})$. Moreover, both $\Phi^2(-\Delta)$ and $r_{ij}\Delta f_{ij}(\zeta)$ converges to 0 as $m\to+\infty$. By conditions A2 and A3, it can be easily obtained that $\frac{1}{m}\,\mathrm{var}\{\tilde{U}(\beta_\tau) - U(\beta_\tau)\} = o(1)$. Therefore, for any β_τ, we have $\frac{1}{\sqrt{m}}\{\tilde{U}(\beta_\tau) - U(\beta_\tau)\} \to 0$ as $m \to +\infty$. □

Theorem 3.3.2. *Under regularity conditions A1-A5, the smoothed estimator $\tilde{\beta}_\tau$ is $\sqrt{m}$-consistent and asymptotically normal,*

$$\sqrt{m}(\tilde{\beta}_\tau - \beta_\tau) \to N(0, G^{-1}(\beta_\tau)V\{G^{-1}(\beta_\tau)\}^T),$$

where $G(\beta_\tau)$ and V have the same expressions as in Theorem 3.3.1.

Proof of Theorem 3.3.2. From the results in Theorem 3.3.1 along with the fact that $\sup_{\hat{\beta}_\tau \in \mathfrak{B}} \| m^{-1}\{U(\hat{\beta}_\tau) - \bar{U}(\hat{\beta}_\tau)\}\| = o(m^{-1/2})$ a.s., and by the triangle inequality, we have $\sup_{\hat{\beta}_\tau \in \mathfrak{B}} \| m^{-1}\{\tilde{U}(\hat{\beta}_\tau) - \bar{U}(\hat{\beta}_\tau)\}\| = o(m^{-1/2})$. If we denote β_τ as the unique solution of equation $\bar{U}(\hat{\beta}_\tau) = 0$ and $\tilde{\beta}_\tau$ solving $\tilde{U}(\hat{\beta}_\tau) = 0$, we can obtain that $\tilde{\beta}_\tau \to \beta_\tau$ as $m \to +\infty$.

Before proving the asymptotic normality of $\tilde{\beta}_\tau$, we first prove that $m^{-1}\{\tilde{G}(\beta_\tau) - G(\beta_\tau)\} \xrightarrow{p} 0$, where $\tilde{G}(\beta_\tau) = -\partial \tilde{U}(\beta_\tau)/\partial \beta_\tau = \sum_{i=1}^m X_i^T \mathit{\Gamma}_i \Sigma_i^{-1}(\rho)\tilde{\mathit{\Lambda}}_i X_i$. Let $H_i^T = X_i^T \mathit{\Gamma}_i \Sigma_i^{-1}(\rho) = (h_{i1}, \ldots, h_{in_i})$, where h_{ij} is a $p \times 1$ vector. We can obtain that

$$E\{\tilde{G}(\beta_\tau)\} - G(\beta_\tau) = \sum_{i=1}^{m}\sum_{j=1}^{n_i} h_{ij}\Big\{\frac{1}{r_{ij}} E\phi\Big(\frac{\varepsilon_{ij}}{r_{ij}}\Big) - f_{ij}(0)\Big\} x_{ij}.$$

Because

$$\begin{aligned}
\Big|\frac{1}{r_{ij}} E\phi\Big(\frac{\varepsilon_{ij}}{r_{ij}}\Big) - f_{ij}(0)\Big| &= \Big|\frac{1}{r_{ij}}\int_{-\infty}^{+\infty} \phi\Big(\frac{\varepsilon}{r_{ij}}\Big) f_{ij}(\varepsilon)d\varepsilon - f_{ij}(0)\Big| \\
&= \Big|\int_{-\infty}^{+\infty} \phi(t)\{f_{ij}(0) + r_{ij} t f_{ij}(\xi_t)\}dt - f_{ij}(0)\Big| \\
&= \Big| r_{ij}\int_{-\infty}^{+\infty} \phi(t) t f_{ij}(\xi_t) dt\Big| \\
&\le r_{ij}\int_{-\infty}^{+\infty} |\phi(t) t f_{ij}(\xi_t)| dt,
\end{aligned}$$

where ξ_t lies between 0 and $r_{ij}t$. By condition A1, there exists a constant M such that $f_{ij}(\xi_t) \leq M$. Furthermore, according to condition A4, we have

$$\left|\frac{1}{r_{ij}} E\phi\Big(\frac{\varepsilon_{ij}}{r_{ij}}\Big) - f_{ij}(0)\right| \leq \sqrt{\frac{2}{\pi}} r_{ij} M \to 0.$$

By the strong law of large numbers, we know that $m^{-1}\tilde{G}(\beta_\tau) \to E\{m^{-1}\tilde{G}(\beta_\tau)\}$. Using the triangle inequality, we have

$$|m^{-1}\{\tilde{G}(\beta_\tau) - G(\beta_\tau)\}| \leq |m^{-1}\{\tilde{G}(\beta_\tau) - E\tilde{G}(\beta_\tau)\}| + |m^{-1}\{E\tilde{G}(\beta_\tau) - G(\beta_\tau)\}| \to o(1),$$

which is equivalent to $m^{-1}\{\tilde{G}(\beta_\tau) - G(\beta_\tau)\} \xrightarrow{p} 0$.

By Taylor series expansion of $\tilde{U}(\hat{\beta}_\tau)$ around β_τ gives us

$$\tilde{U}(\hat{\beta}_\tau) = \tilde{U}(\beta_\tau) - \tilde{G}(\hat{\beta}^*_\tau)(\hat{\beta}_\tau - \beta_\tau),$$

where $\hat{\beta}^*_\tau$ lies between $\hat{\beta}_\tau$ and β_τ. Let $\hat{\beta}_\tau = \tilde{\beta}_\tau$. Because $\tilde{U}(\tilde{\beta}_\tau) = 0$ and $\tilde{\beta}_\tau \to \beta_\tau$, we therefore obtain $\hat{\beta}^*_\tau \to \beta_\tau$ and $\tilde{G}(\hat{\beta}^*_\tau) \to \tilde{G}(\beta_\tau)$. By Lemma 3.3.1 and $m^{-1}\{\tilde{G}(\beta_\tau) - G(\beta_\tau)\} \xrightarrow{p} 0$, we thus have

$$\sqrt{m}(\tilde{\beta}_\tau - \beta_\tau) = G_m^{-1}(\beta_\tau)\frac{1}{\sqrt{m}}U(\beta_\tau) + o_p(1).$$

Therefore $\sqrt{m}(\tilde{\beta}_\tau - \beta_\tau) \to N(0, G^{-1}(\beta_\tau)V\{G^{-1}(\beta_\tau)\}^T)$ as $m \to +\infty$. □

From Theorem 3.3.2, we can obtain a natural sandwich form estimator of the variance-covariance matrix of $\sqrt{m}(\tilde{\beta}_\tau - \beta_\tau)$ as

$$\widehat{\text{cov}}(\sqrt{m}(\tilde{\beta}_\tau - \beta_\tau)) = \hat{G}^{-1}(\tilde{\beta}_\tau)\tilde{V}\{\hat{G}^{-1}(\tilde{\beta}_\tau)\}^T, \qquad (3.20)$$

where $\tilde{G}(\tilde{\beta}_\tau) = \frac{1}{m}\sum_{i=1}^m X_i^T \mathit{\Gamma}_i \Sigma_i^{-1}(\rho)\mathit{\Gamma}_i X_i$ and $\tilde{V} = \frac{1}{m}\sum_{i=1}^m X_i^T \mathit{\Gamma}_i \Sigma_i^{-1}(\rho)\,\mathrm{cov}\{\tilde{\psi}_\tau(y_i - X_i\tilde{\beta}_\tau)\}\Sigma_i^{-1}(\rho)\mathit{\Gamma}_i X_i$. Based on this formula, we update matrix $\tilde{\mathit{\Omega}}$ in the Newton-Raphson iteration on page 32.

Chapter 4

Numerical Study

In order to examine the small sample performance of the proposed method, we conducted extensive simulation studies. A part of the simulation results are reported in this section. The method is then applied to the labor pain data (Davis 1991) as an illustration.

4.1 Simulation

In this section, we report the results of a simulation study to investigate the performance of the proposed methods. This section consist of two parts. In the first part, we compare the different quantile regression estimators. In the second part, we make a comparison between parameter estimates from median regressions using the proposed method and parameter estimates from mean regressions using a linear mixed effect method.

4.1.1 Simulation Setup

We generate datasets from the model

$$y_{ij} = \beta_0 + x_{ij1}\beta_1 + x_{ij2}\beta_2 + \varepsilon_{ij} \tag{4.1}$$

for $i = 1, \ldots, m$ and $j = 1, \ldots, n_i$, where x_{ij1} are sampled from the Bernoulli distribution with probability 0.5, $Bernoulli(0.5)$, and x_{ij2} are generated from a standard normal distribution. To carry out the simulation study, we must specify the values of ε_{ij}, m, n_i, τ, and the correlations structure for the error term $\varepsilon_i = (\varepsilon_{i1}, \ldots, \varepsilon_{in_i})^T$ as well. In this simulation study, we set the sample size $m = 500$ and a balanced design $n_i = 4$ for all $i = 1, \ldots, 500$. For the correlation of the error terms within each individual, we let the variance-covariance matrix of ε_i to be an AR(1) structure expressed as

$$\Sigma_\varepsilon(\rho) = \begin{pmatrix} 1 & \rho & \rho^2 & \ldots & \rho^{n_i-1} \\ \rho & 1 & \rho & \ldots & \rho^{n_i-2} \\ & & & \vdots & \\ \rho^{n_i-1} & \rho^{n_i-2} & & \ldots & 1 \end{pmatrix},$$

where ρ is set to be 0.1, 0.5, or 0.9 respectively, which will generate errors with small, medium and large correlation, respectively. Three different distributions are considered for the random error ε_i:

***Case* 1.** Normal distribution, assume that ε_i follows a multivariate normal distribution with mean $-q_\tau$ and covariance $\Sigma_\varepsilon(\rho)$, $N_p(-q_\tau, \Sigma_\varepsilon(\rho))$, where q_τ is the τth quantile of the standard normal distribution.

***Case* 2.** Chi-squared distribution, assume that $\varepsilon_i - q_\tau$ follows a multivariate Chi-squared distribution with two degrees of freedom (χ^2_2), where q_τ is the τth quantile of the χ^2_2 distribution.

***Case* 3.** Student's T distribution, suppose that $\varepsilon_i - q_\tau$ follows a multivariate T distribution with three degrees of freedom (T_3), where q_τ is the τth quantile of the T_3 distribution.

The values of parameters used in the simulation are $\beta_0 = -0.5$, $\beta_1 = 0.5$ and $\beta_2 = 1$. Quantiles of $\tau = 0.25$, 0.5 and 0.95 are chosen to study the performance of the quantile regression estimators for the response distribution.

4.1.2 Analysis Methods

We set the number of replications used in the simulation to $N = 1000$. Several statistics are used to analyze the results of the simulation. Within each simulation, the *Bias* is defined as the average of differences between estimated parameter values and the true parameter values: for $k = 0, 1, 2$

$$Bias(\hat{\beta}_k) = \frac{1}{N}\sum_{i=1}^{N} \hat{\beta}_k^{(i)} - \beta_k, \tag{4.2}$$

where $\hat{\beta}_k^{(i)}$ is the estimated component of each estimator for the ith replication and β_k is the corresponding true parameter.

To see how stable each estimator is, the sample standard deviation (SD) of 1000 estimates of β_k, $k = 0, 1, 2$, is examined. That is, we calculate

$$SD(\hat{\beta}_k) = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(\hat{\beta}_k^{(i)} - \bar{\beta}_k)^2}, \tag{4.3}$$

where $\bar{\beta}_k$ is the sample mean of 1000 estimated β_k. Also, the average of 1000 estimated asymptotic standard errors (SE) is reported, which can be obtained using formula (3.20). Furthermore, we calculate the percentage of simulation runs when the true parameter falls into the 95% confidence interval constructed based on $\hat{\beta}_k^{(i)}$ and the sandwich estimate of the covariance matrix of $\hat{\beta}_k^{(i)}$ ($sd_k^{(i)}$), $CI_{0.95}^{(i)}(k) = \hat{\beta}_k^{(i)} \pm Z_{0.025} \cdot sd_k^{(i)}$. We denote this percentage of simulation runs as $P_{0.95}$.

We also compare the performances of the estimators using different methods, e.g., the estimated statistic of relative efficiency (EFF). EFF can be calculated by using the following formula

$$EFF(k) = \frac{MSE_2(\hat{\beta}_k)}{MSE_1(\hat{\beta}_k)} \tag{4.4}$$

where $MSE_r(\hat{\beta}_k)$, $r = 1, 2$, is the mean square error of $\hat{\beta}_k$ using the rth method, which is expressed as

$$MSE_r(\hat{\beta}_k) = Bias_r(\hat{\beta}_k)^2 + SD_r(\hat{\beta}_k)^2, \tag{4.5}$$

where $Bias_r$ and SD_r are averaged estimated bias and estimated sample standard deviation using the rth method.

4.1.3 Comparison of Quantile Regressions Using Different Methods

In this section the results of 1000 simulation runs of quantile regression using different parameter estimators are analyzed. We report the averaged bias ($Bias$) and relative efficiency (EFF) of the estimates of β_0, β_1 and β_2 using different quantile regression methods (ordinary quantile regression method assuming working independence (WI), quasi-likelihood working independence method (QLWI), proposed quantile regression method (PQR), and adjusted quantile regression method (AQR)). Moreover, the results from the simulations of three cases of random effect distributions are also reported.

Table 4.1: Biases($Bias$) and relative efficiencies(EFF) to the estimators of β_0, β_1 and β_2 using different methods (AQR, PQR, QLWI and WI) at five quantiles 0.05, 0.25, 0.5, 0.75, 0.95. Where ρ is specified as 0.1, 0.5, and 0.9 respectively, and the error follows a multivariate normal distribution($case$ 1).

			β_0		β_1		β_2	
τ	ρ	Method	Bias	EFF	Bias	EFF	Bias	EFF
0.05	0.1	AQR	0.0081	1.023	0.0001	1.083	-0.0032	1.102
		PQR	-0.0047	1.058	0.0002	1.081	-0.0031	1.128
		QLWI	-0.0032	1.056	-0.0053	1.068	-0.0085	1.076
		WI	-0.0017	1.000	0.0031	1.000	-0.0037	1.000
	0.5	AQR	0.0116	1.010	-0.0022	1.122	0.0007	1.225
		PQR	-0.0023	1.085	-0.0019	1.157	0.0007	1.248
		QLWI	-0.0005	1.044	-0.0064	1.079	-0.0052	1.116
		WI	0.0041	1.000	-0.0044	1.000	0.0002	1.000
	0.9	AQR	0.0213	1.150	-0.0039	2.427	0.0002	2.341
		PQR	0.0033	1.264	-0.0035	2.378	0.0001	2.353
		QLWI	0.0063	1.047	-0.0141	1.066	-0.0026	1.090
		WI	0.0101	1.000	-0.0099	1.000	0.0026	1.000
0.25	0.1	AQR	0.0019	1.046	0.0015	1.041	-0.0001	1.069
		PQR	-0.0005	1.040	0.0014	1.044	-0.0001	1.069
		QLWI	-0.0011	1.052	-0.0004	1.054	-0.0021	1.072
		WI	0.0000	1.000	0.0020	1.000	-0.0002	1.000
	0.5	AQR	0.0043	1.098	-0.0017	1.191	0.0015	1.236
		PQR	0.0017	1.111	-0.0017	1.194	0.0015	1.242
		QLWI	0.0014	1.039	-0.0039	1.057	-0.0005	1.081

Continued on next page

τ	ρ	Method	Bias	EFF	Bias	EFF	Bias	EFF
		WI	0.0025	1.000	-0.0019	1.000	0.0016	1.000
	0.9	AQR	0.0048	1.186	-0.0006	2.811	0.0000	2.707
		PQR	0.0013	1.195	-0.0008	2.816	0.0000	2.706
		QLWI	0.0014	1.016	-0.0043	1.057	-0.0014	1.081
		WI	0.0028	1.000	-0.0023	1.000	0.0006	1.000
0.5	0.1	AQR	0.0022	1.050	-0.0020	1.054	0.0006	1.049
		PQR	0.0022	1.050	-0.0020	1.053	0.0006	1.049
		QLWI	0.0021	1.051	-0.0020	1.065	0.0007	1.063
		WI	0.0019	1.000	-0.0017	1.000	0.0008	1.000
	0.5	AQR	0.0002	1.059	0.0018	1.260	0.0010	1.247
		PQR	0.0002	1.059	0.0018	1.260	0.0010	1.247
		QLWI	0.0001	1.024	0.0027	1.076	0.0008	1.061
		WI	-0.0000	1.000	0.0025	1.000	0.0006	1.000
	0.9	AQR	-0.0003	1.256	-0.0005	3.135	0.0001	3.026
		PQR	-0.0003	1.256	-0.0005	3.136	0.0001	3.026
		QLWI	-0.0000	1.020	-0.0010	1.049	0.0012	1.069
		WI	-0.0004	1.000	-0.0013	1.000	0.0015	1.000
0.75	0.1	AQR	-0.0039	1.043	-0.0001	1.057	-0.0006	1.051
		PQR	0.0020	1.044	-0.0001	1.061	-0.0006	1.053
		QLWI	0.0024	1.042	0.0020	1.072	0.0014	1.068
		WI	0.0019	1.000	-0.0013	1.000	-0.0008	1.000
	0.5	AQR	-0.0020	1.110	0.0025	1.245	-0.0007	1.259
		PQR	0.0006	1.113	0.0026	1.239	-0.0007	1.258
		QLWI	0.0014	1.028	0.0043	1.059	0.0011	1.093

Continued on next page

τ	ρ	Method	Bias	EFF	Bias	EFF	Bias	EFF
		WI	-0.0005	1.000	0.0033	1.000	-0.0010	1.000
	0.9	AQR	-0.0027	1.238	-0.0005	2.712	-0.0010	2.695
		PQR	0.0008	1.239	-0.0005	2.699	-0.0009	2.687
		QLWI	0.0019	1.023	0.0011	1.060	-0.0000	1.093
		WI	0.0008	1.000	-0.0015	1.000	-0.0021	1.000
0.95	0.1	AQR	-0.0094	1.066	0.0047	1.069	0.0029	1.099
		PQR	0.0033	1.092	0.0046	1.071	0.0028	1.118
		QLWI	0.0022	1.073	0.0088	1.071	0.0077	1.089
		WI	-0.0004	1.000	0.0032	1.000	0.0031	1.000
	0.5	AQR	-0.0142	0.988	-0.0008	1.136	-0.0018	1.257
		PQR	0.0001	1.092	-0.0016	1.144	-0.0015	1.248
		QLWI	-0.0009	1.051	0.0021	1.088	0.0038	1.103
		WI	-0.0039	1.000	-0.0039	1.000	-0.0017	1.000
	0.9	AQR	-0.0136	1.212	-0.0003	2.152	-0.0015	2.187
		PQR	0.0037	1.244	-0.0001	2.155	-0.0017	2.129
		QLWI	0.0018	1.023	0.0088	1.053	0.0048	1.079
		WI	-0.0021	1.000	0.0052	1.000	-0.0002	1.000

Table 4.1 shows the results when ε_i follows a multivariate normal distribution (*case* 1) with an AR(1) correlation structure where the value of ρ is specified as 0.1, 0.5 and 0.9 respectively. As we can see, when the correlation is low ($\rho = 0.1$, repeated observations are almost independently distributed), the averaged biases and relative efficiencies of quantile regression estimators $\hat{\beta}_{QLWI\tau}$, $\hat{\beta}_{PQR\tau}$ and $\hat{\beta}_{AQR\tau}$ are comparable, and these three estimators perform slightly better than the standard quantile regression estimator assuming working independence ($\hat{\beta}_{WI\tau}$). When the correlation is

high ($\rho = 0.5$, $\rho = 0.9$), the proposed estimators $\hat{\beta}_{PQR\tau}$ and $\hat{\beta}_{AQR\tau}$ are equally efficient with small biases and much smaller variances than the other two working independent estimators. Moreover, the estimators $\hat{\beta}_{PQR\tau}$ and $\hat{\beta}_{AQR\tau}$ become more efficient as the correlation (ρ) increases. In general, these two proposed methods provide more efficient estimates of $\beta_{1\tau}$ and $\beta_{2\tau}$ than the intercept parameter $\beta_{0\tau}$. Table A.1 and Table A.2 in Appendix show the biases and relative efficiencies of different quantile regression estimators when ε_i is χ^2_2 (*case* 2) and T_3 (*case* 3) distributed respectively. Similar performance are observed from Table A.1 and Table A.2 except that the proposed estimators are more efficient at higher quantiles when the random effect follows a χ^2_2 distribution (*case* 2).

4.1.4 Evaluation of Asymptotic Properties

To evaluate the asymptotic properties, the sample standard deviation (SD) and the averaged asymptotic standard errors (SE) of the proposed and adjusted proposed estimators are reported in Table 4.2, and Tables A.3 and A.4 in Appendix, respectively according to three cases of the ε_i's distribution. For our proposed estimators, $P_{0.95}$ denotes the percentage of simulation runs when the true parameter falls into the 95% confidence intervals constructed based on the sandwich estimate of the covariance matrix of $\hat{\beta}_\tau$. We can see that whatever the distribution of random effects follows, each value of SD is very small and close to the corresponding SE value. This means the proposed estimators and the estimate of the standard deviation of $\hat{\beta}_\tau$ perform well. Furthermore, since all values of $P_{0.95}$ in these three tables are close to 0.95, the proposed estimators are asymptotically normally distributed and inferences based on it are reliable.

Table 4.2: Standard deviation (SD) of 1000 estimates of three β-parameters, the sample average of 1000 estimated standard errors (SE) and the probability of estimates falling into the 95% confidence interval are reported for different methods (AQR, PQR, QLWI and WI) at five quantiles 0.05, 0.25, 0.5, 0.75, 0.95. Where ρ is specified as 0.1, 0.5, and 0.9 respectively, and the error follows a multivariate normal distribution (*case* 1).

			β_0			β_1			β_2		
τ	ρ	Method	SD	SE	$P_{0.95}$	SD	SE	$P_{0.95}$	SD	SE	$P_{0.95}$
0.05	0.1	AQR	0.0664	0.0617	0.926	0.0901	0.0869	0.935	0.0439	0.0416	0.936
		PQR	0.0655	0.0623	0.936	0.0902	0.0876	0.933	0.0433	0.0416	0.940
	0.5	AQR	0.0681	0.0667	0.939	0.0860	0.0842	0.948	0.0444	0.0412	0.927
		PQR	0.0666	0.0675	0.948	0.0847	0.0853	0.952	0.0440	0.0418	0.937
	0.9	AQR	0.0780	0.0769	0.951	0.0596	0.0576	0.939	0.0308	0.0282	0.929
		PQR	0.0770	0.0778	0.956	0.0603	0.0586	0.944	0.0307	0.0287	0.929
0.25	0.1	AQR	0.0427	0.0424	0.951	0.0598	0.0589	0.949	0.0294	0.0296	0.950
		PQR	0.0429	0.0424	0.953	0.0597	0.0590	0.951	0.0294	0.0296	0.949
	0.5	AQR	0.0472	0.0463	0.947	0.0545	0.0548	0.947	0.0275	0.0274	0.952
		PQR	0.0471	0.0463	0.950	0.0544	0.0549	0.946	0.0274	0.0275	0.952
	0.9	AQR	0.0536	0.0531	0.944	0.0376	0.0363	0.939	0.0185	0.0180	0.949

Continued on next page

τ	ρ	Method	SD	SE	$P_{0.95}$	SD	SE	$P_{0.95}$	SD	SE	$P_{0.95}$
		PQR	0.0536	0.0532	0.945	0.0375	0.0364	0.939	0.0185	0.0181	0.947
0.5	0.1	AQR	0.0395	0.0395	0.943	0.0554	0.0546	0.945	0.0270	0.0273	0.956
		PQR	0.0395	0.0395	0.943	0.0554	0.0546	0.945	0.0270	0.0273	0.956
	0.5	AQR	0.0443	0.0431	0.948	0.0501	0.0494	0.949	0.0247	0.0244	0.955
		PQR	0.0443	0.0431	0.948	0.0501	0.0494	0.949	0.0247	0.0244	0.955
	0.9	AQR	0.0450	0.0499	0.947	0.0323	0.0329	0.948	0.0165	0.0164	0.950
		PQR	0.0450	0.0499	0.947	0.0323	0.0329	0.948	0.0165	0.0164	0.950
0.75	0.1	AQR	0.0429	0.0424	0.950	0.0607	0.0590	0.949	0.0299	0.0294	0.956
		PQR	0.0428	0.0424	0.948	0.0606	0.0590	0.948	0.0296	0.0294	0.954
	0.5	AQR	0.0469	0.0455	0.937	0.0561	0.0541	0.943	0.0286	0.0278	0.942
		PQR	0.0469	0.0455	0.943	0.0563	0.0541	0.945	0.0287	0.0278	0.948
	0.9	AQR	0.0536	0.0535	0.948	0.0379	0.0366	0.942	0.0186	0.0176	0.934
		PQR	0.0536	0.0536	0.950	0.0379	0.0366	0.946	0.0186	0.0176	0.929
0.95	0.1	AQR	0.0660	0.0624	0.926	0.0897	0.0880	0.945	0.0457	0.0433	0.934
		PQR	0.0658	0.0630	0.933	0.0896	0.0889	0.944	0.0453	0.0437	0.939
	0.5	AQR	0.0700	0.0653	0.924	0.0897	0.0839	0.923	0.0435	0.0411	0.935

Continued on next page

τ	ρ	Method	SD	SE	$P_{0.95}$	SD	SE	$P_{0.95}$	SD	SE	$P_{0.95}$
		PQR	0.0680	0.0658	0.939	0.0894	0.0848	0.931	0.0436	0.0417	0.943
	0.9	AQR	0.0795	0.0789	0.926	0.0633	0.0576	0.919	0.0309	0.0278	0.928
		PQR	0.0796	0.0790	0.920	0.0633	0.0585	0.925	0.0312	0.0283	0.929

4.1.5 Comparison of Median and Mean Regression

As discussed in Chapter 2, quantile regressions outperform mean regression when the random error distribution is skewed or heavy-tailed. This fact has been verified in our simulation studies. The results from median regression ($\tau = 0.5$) using methods of WI, PQR and AQR, and the mean regression using a linear mixed effect model (LME), are reported in Table 4.3. We analyze the results mainly by comparing the LME with our proposed methods of PQR and AQR, WI seen as a control. As expected, when the error terms follow a normal distribution, the LME and proposed quantile methods have comparable bias, but the LME is more efficient than the median regressions according to the average of the estimated efficiencies of the three β_τ parameters. However, when the error terms follow chi-square distribution (χ_2^2) or student's t distribution (T_3), the LME performs worse than our proposed median regression methods, particularly in estimating the intercept parameter $\beta_{0\tau}$. The median regression model is more robust to the model misspecification, while LME can only provide misleading results in those cases.

Table 4.3: Simulation results comparing the linear mixed effect model and the proposed median regression models. Biases ($Bias$) and relative efficiencies (EFF) to each estimator are reported for three different error distributions ($case$ 1, 2 and 3).

			β_0		β_1		β_2	
Err	ρ	Method	Bias	EFF	Bias	EFF	Bias	EFF
Nor	0.1	LME	0.0015	1.542	-0.0020	1.546	0.0004	1.543
		AQR	0.0022	1.050	-0.0020	1.054	0.0006	1.049
		PQR	0.0022	1.050	-0.0020	1.053	0.0006	1.049
		QLWI	0.0021	1.051	-0.0020	1.065	0.0007	1.063
		WI	0.0019	1.000	-0.0017	1.000	0.0008	1.000
	0.5	LME	0.0008	1.517	0.0022	2.088	0.0006	2.048
		AQR	0.0002	1.059	0.0018	1.260	0.0010	1.247
		PQR	0.0002	1.059	0.0018	1.260	0.0010	1.247
		QLWI	0.0001	1.024	0.0027	1.076	0.0008	1.061
		WI	-0.0000	1.000	0.0025	1.000	0.0006	1.000
	0.9	LME	-0.0021	1.729	0.0007	7.979	0.0001	7.605
		AQR	-0.0003	1.256	-0.0005	3.135	0.0001	3.026
		PQR	-0.0003	1.256	-0.0005	3.136	0.0001	3.026
		QLWI	-0.0000	1.020	-0.0010	1.049	0.0012	1.069
		WI	-0.0004	1.000	-0.0013	1.000	0.0015	1.000
Chi	0.1	LME	0.6193	0.011	-0.0024	1.008	0.0005	1.002
		AQR	0.0084	1.030	-0.0059	1.037	0.0015	1.042
		PQR	0.0084	1.030	-0.0059	1.037	0.0015	1.042
		QLWI	0.0078	1.022	-0.0045	1.043	0.0034	1.037
		WI	0.0069	1.000	-0.0060	1.000	0.0018	1.000

Continued on next page

Err	ρ	Method	Bias	EFF	Bias	EFF	Bias	EFF
	0.5	LME	0.6159	0.011	0.0014	1.148	-0.0005	1.020
		AQR	0.0064	1.033	0.0008	1.061	-0.0004	1.084
		PQR	0.0064	1.033	0.0008	1.061	-0.0004	1.084
		QLWI	0.0057	1.019	0.0025	1.038	0.0013	1.047
		WI	0.0051	1.000	0.0006	1.000	0.0001	1.000
	0.9	LME	0.6149	0.018	0.0003	2.802	-0.0002	2.406
		AQR	0.0053	1.188	-0.0019	1.982	0.0004	1.864
		PQR	0.0053	1.188	-0.0019	1.982	0.0004	1.864
		QLWI	0.0035	1.016	0.0027	1.045	0.0022	1.056
		WI	0.0024	1.000	0.0010	1.000	0.0008	1.000
T	0.1	LME	0.0016	0.613	0.0001	0.650	-0.0007	0.631
		AQR	0.0004	1.031	-0.0007	1.052	-0.0001	1.046
		PQR	0.0004	1.031	-0.0007	1.052	-0.0001	1.046
		QLWI	0.0003	1.033	-0.0001	1.057	-0.0001	1.052
		WI	0.0005	1.000	-0.0009	1.000	0.0001	1.000
	0.5	LME	0.0008	0.583	-0.0002	0.849	-0.0009	0.831
		AQR	0.0006	1.122	0.0003	1.234	0.0004	1.301
		PQR	0.0006	1.122	0.0003	1.234	0.0004	1.301
		QLWI	0.0003	1.023	0.0012	1.061	0.0001	1.051
		WI	0.0005	1.000	0.0009	1.000	0.0001	1.000
	0.9	LME	0.0034	0.591	-0.0012	2.863	0.0006	3.338
		AQR	0.0021	1.223	-0.0003	2.916	0.0002	2.740
		PQR	0.0020	1.223	-0.0003	2.916	0.0002	2.740
		QLWI	0.0029	1.032	-0.0011	1.062	0.0008	1.047

Continued on next page

Err	ρ	Method	Bias	EFF	Bias	EFF	Bias	EFF
		WI	0.0031	1.000	-0.0015	1.000	0.0010	1.000

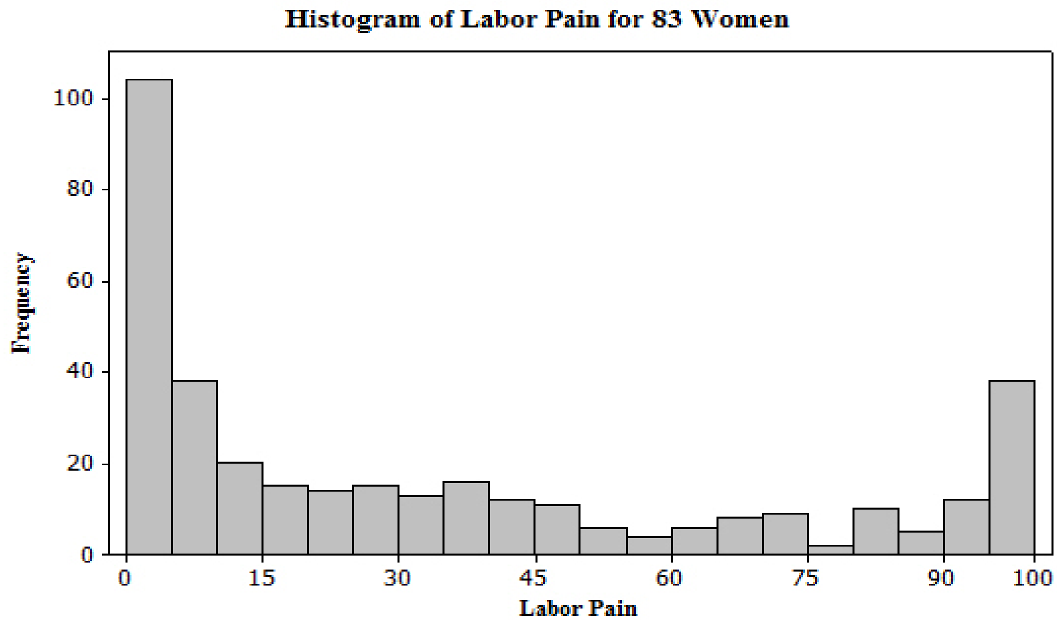

Figure 4.1: Histogram of measured labor pain for all 83 women.

4.2 An Example: Labor Pain Study

In this section, we illustrated the proposed method for quantile regression by analyzing the labor pain data, reported by Davis (1991) and analyzed by Jung (1996). The data set arose from a randomized clinical trial on the effectiveness of a medication for relieving labor pain. A total of $m = 83$ women were randomly assigned to either a pain medication group (43 women) or a placebo group (40 women). The response is a self-reported measure of pain measured every 30 minutes on a 100-mm line, where $0 =$ no pain and $100 =$ extreme unbearable pain. The maximum number of measurements for each women was 6, but at later measurement times there are numerous values missing with a nearly monotone pattern. Figure 4.1, a histogram of all the pains, shows that the data is severely skewed. Therefore, the mean regression may not be appropriate. In Figure 4.2, a box-plot shows the mean and median of the pain over time for all

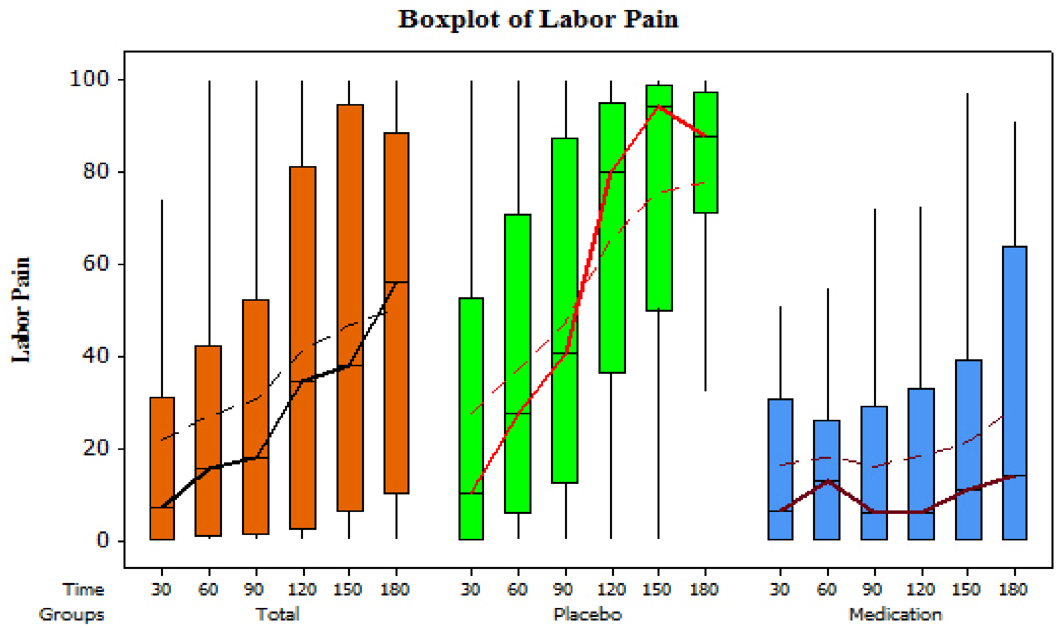

Figure 4.2: Box-plot of measured labor pain for all women in placebo and medication groups. The thick solid lines represent the median, while the means are connected with thin dashed lines.

83 women and those in two different groups. Statistical dependence on the temporal course of the quartiles of the response is evident, especially for the placebo group.

Let y_{ij} be the amount of pain for the ith patient at time j, R_i be the treatment indicator taking 0 for placebo and 1 for medication, and T_{ij} be the measurement time divided by 30 minutes. Jung (1996) considered the median regression model

$$y_{ij} = \beta_0 + \beta_1 R_i + \beta_2 T_{ij} + \beta_3 R_i T_{ij} + \varepsilon_{ij}, \tag{4.6}$$

where ε_{ij} is a zero-median error term. Note that $(\beta_0+\beta_1)+(\beta_2+\beta_3)T_{ij}$ is the median for the treatment group and the median for the placebo group is $\beta_0 + \beta_2 T_{ij}$.

Our proposed quantile regression model was fit for three quartiles, $\tau = 0.25$, 0.5 and 0.75, respectively. We report the estimated parameters (EP), their asymptotic

Table 4.4: Estimated parameters (EP), their standard errors (SE) and corresponding 95% confidence intervals (CI) from fitting both the proposed quantile regression model (PQR) and usual quantile regression assuming working independence (WI) at three quartiles, $\tau = 0.25$, 0.5, and 0.75.

		Proposed Method			WI		
τ	β	EP	SE	CI	ES	SE	CI
0.25	β_0	-10.32	0.42	(-11.13, -9.50)	-10.83	2.20	(-15.14, -6.52)
	β_1	9.08	0.42	(8.27, 9.90)	10.83	2.20	(6.51, 15.15)
	β_2	17.72	0.41	(16.92, 18.51)	10.83	2.20	(6.52, 15.14)
	β_3	-15.58	0.41	(-16.38, -14.79)	-10.83	2.20	(-15.15, -6.51)
0.5	β_0	-10.44	1.54	(-13.45, -7.43)	-6.20	7.95	(-21.77, 9.37)
	β_1	8.96	1.54	(5.95, 11.97)	12.20	8.88	(-5.21, 29.61)
	β_2	21.05	1.27	(18.56, 23.53)	17.20	2.35	(12.60, 21.80)
	β_3	-12.25	1.27	(-14.74, -9.77)	-16.20	2.72	(-21.53, -10.87)
0.75	β_0	1.02	4.08	(-6.97, 9.02)	58.67	14.83	(29.60, 87.74)
	β_1	20.42	4.08	(12.43, 28.42)	-42.67	16.30	(-74.61, -10.72)
	β_2	22.84	0.68	(21.51, 24.17)	7.67	3.44	(0.93, 14.40)
	β_3	-10.46	0.68	(-11.79, -9.13)	-2.67	4.02	(-10.54, 5.21)

standard errors (SE) and the 95% confidence intervals (CI) in Table 4.4. Here we also list the results of the usual quantile regression method assuming working independence for comparison. At the 0.25th quantile, we see that our proposed method gives smaller standard errors, although these two methods produce comparable estimates of parameters. Note that all parameter estimates are significant at 5% level, meaning that each covariate has effect on the 25% quantile labor pain. Parameter estimates to the median regression methods have similar properties, except that the

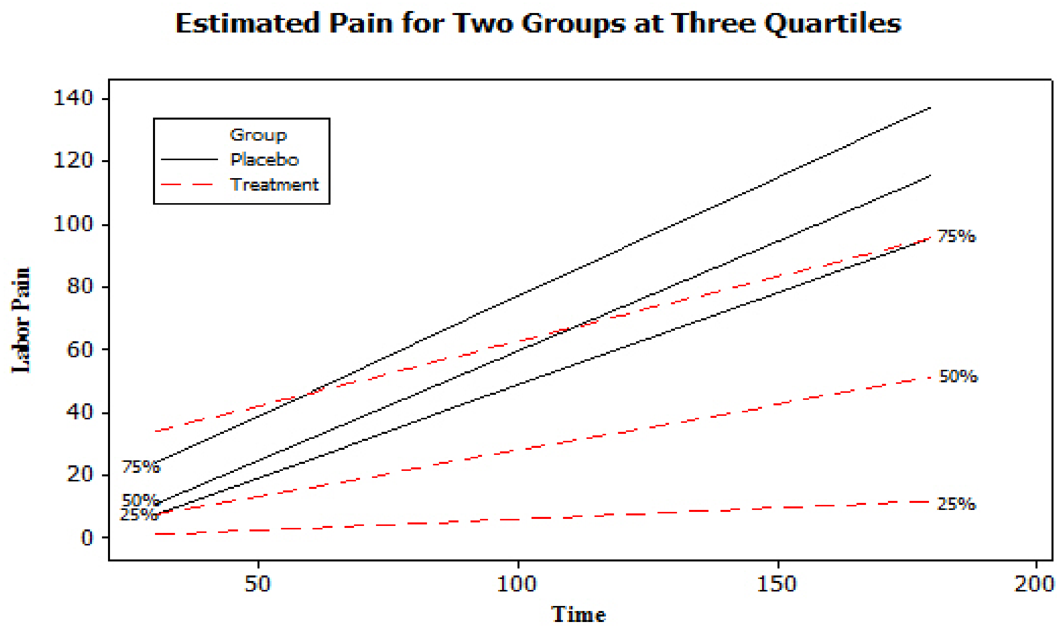

Figure 4.3: Labor pain obtained by using proposed quantile regression method at three quartiles 25%, 50% and 75%.

usual quantile regression method assuming working independence gives insignificant estimates of β_0 and β_1, indicating similar baseline pain among two groups. While, for the third quartile (0.75th quantile), our proposed method and the WI method have very different parameter estimates with the proposed method giving much smaller standard errors of the estimates. The insignificant β_3 in WI method indicates similar time effects on the amount of pain in groups of placebo and medication, which contradicts our medical knowledge, while the significance of β_3 in our proposed method provides a perfect interpretation.

To investigate how treatment and time affect the amount of labor pain at three quartiles (0.25, 0.5, 0.75), we use our proposed method to compare the estimated values of β_0 with $\beta_0 + \beta_1$ and β_2 with $\beta_2 + \beta_3$ at each quartile, respectively. The result is visualized in Figure 4.3, where we can easily see that medication treatment do help women relieve their labor pain, and the pain of women in the placebo group grows faster with time than that in the treatment group. Moreover, the amount of pain tends to grow slightly faster at higher quantiles than that at lower quantiles. These conclusions are consistent with the box plots shown in Figure 4.2 and those given by Jung (1996) and Leng and Zhang (2012).

Chapter 5

Conclusion

In this thesis, we have proposed a new quantile regression model for longitudinal data, incorporating the correlations between repeated measures. We applied a general stationary auto-correlation structure to the estimating equations. To reduce the computational burden caused by the non-continuous estimating functions, we have employed the induced smoothing method of Fu and Wang (2012) for quantile regression. The estimates of the regression parameters and their covariance matrix are then obtained using Newton-Raphson iteration technique. It can be seen that our proposed method is a simple and efficient way to account for within-subject correlations in quantile regression for longitudinal data. This approach drew the inferential methods of quantile regression and the classical mean regression much closer. It reveals that the techniques in GEE's are applicable in quantile regression modeling. The simulation studies in Chapter 4 indicate that the proposed method performs better than the methods assuming working independence especially when the within correlation is high. Furthermore, a comparison is also made between the proposed median regression estimator and the corresponding mean regression estimator, where the former is found to be better in analyzing heavy-tailed or skewed data. Finally, the proposed quantile regression estimator is applied to the labor pain dataset where pain of two groups of women are reported, which reveals how treatment and time affect the amount of labor pain at three quartiles.

We were trying to take the within-subject correlation into consideration of quantile regression modeling, while the effects of unobserved covariates which may be different from individual to individual have not been captured. For instance, in our real data application, the personal perception of labor pain may vary from one to another. Therefore, following Koenker (2004), we may extend our proposed model to a penalized version allowing individual specific effects by adding subject specific parameters and a penalty term. Further developments of our proposed method include extending quantile regression to well studied research areas in mean regression for longitudinal data such as mixed models for count and binary data (Sutradhar, 2011), nonlinear models (He et al., 2003), semi-parametric models (Lin and Carroll, 2006), and nonparametric models (Wu and Zhang, 2006; Qu and Li, 2006).

Bibliography

Bassett Jr, G., Koenker, R., 1978. Asymptotic theory of least absolute error regression. Journal of the American Statistics and Association 73, 618–622.

Brown, B.M., Wang, Y.G., 2005. Standard errors and covariance matrices for smoothed rank estimators. Biometrika 92, 149–158.

Brown, B.M., Wang, Y.G., 2007. Induced smoothing for rank regression with censored survival times. Statistics in Medicine 92, 828–836.

Cade, B.S., Guo, Q., 2000. Estimating effects of constraints on plant performance with regression quantiles. Oikos 91, 245–254.

Cade, B.S., Noon, B.R., 2003. A gentle introduction to quantile regression for ecologists. Frontiers in Ecology and the Environment 1, 412–420.

Chen, L., Wei, L.J., Parzen, M.I., 2004. Quantile regression for correlated observations, in: Proceedings of the Second Seattle Symposium in Biostatistics: Analysis of Correlated Data, pp. 51–70.

Davis, S., 1991. Semi-parametric and non-parametric methods for the analysis of repeated measurements with applications to clinical trials. Statistics in Medicine 10, 1959–1980.

Farcomeni, A., 2012. Quantile regression for longitudinal data based on latent markov subject-specific parameters. Statistics and Computing 22, 141–152.

Fu, L., Wang, Y.G., 2012. Quantile regression for longitudinal data with a working correlation model. Computational Statistics and Data Analysis 56, 2526–2538.

Geraci, M., Bottai, M., 2007. Quantile regression for longitudinal data using the asymmetric laplace distribution. Biostatistics 8, 140–154.

Hao, L., Naiman, D.Q., 2007. Quantile Regression. volume 149 of *Quantitative Application in the Social Sciences*. SAGE Publications, Inc.

He, X., Fu, B., Fung, W.K., 2003. Median regression of longitudinal data. Statistics in Medicine 22, 3655–3669.

Hendricks, W., Koenker, R., 1992. Hierarchical spline models for conditional quantiles and the demand for electricity. Journal of the American Statistical Association 87, 58–68.

Johnson, L.M., Strawderman, R.L., 2009. Induced smoothing for the semiparametric accelerated failure time model: asymptotics and extensions to clustered data. Biometrika 96, 557–590.

Jung, S.H., 1996. Quasi-likelihood for median regression models. Journal of the American Statistical Association 91, 251–257.

Jung, S.H., Ying, Z., 2003. Rank-based regression with repeated measurements data. Biometrika 90, 732–740.

Karlsson, A., 2008. Nonlinear quantile regression estimation of longitudinal data. Communications in Statistics-Simulation and Computation 37, 114–131.

Koenker, R., 2004. Quantile regression for longitudinal data. Journal of Multivariate Analysis 91, 74–89.

Koenker, R., 2005. Quantile Regression. Number 38 in Econometric Society monographs, New York: Combridge University Press.

Koenker, R., Bassett Jr, G., 1978. Regression quantiles. Econometrica 46, 33–55.

Koenker, R., D'Orey, V., 1987. Computing regression quantiles. Applied Statistics 36, 383–393.

Koenker, R., Hallock, K.F., 2001. Quantile regression. Journal of Economeic Perspectives 15, 143–156.

Koenker, R., Machado, J., 1999. Goodness of fit and related inference processes for quantile regression. Journal of the American Statistical Association 94, 1296–1309.

Leng, C., Zhang, W., 2012. Smoothing combined estimating equations in quantile regression for longitudinal data. Statistics and Computing 17 October 2012, 1–14.

Liang, K.Y., Zeger, S.L., 1986. Longitudinal data analysis using generalized linear models. Biometrika 73, 13–22.

Lin, X., Carroll, R.J., 2006. Semiparametric estimation in general repeated measures problems. Journal of the Royal Statistical Society: Series B 68, 69–88.

Lipsitz, S.R., Fitzmaurice, G.M., Molenberghs, G., Zhao, L.P., 1997. Quantile regression methods for longitudinal data with drop-outs: application to cd4 cell counts of patients infected with the human immunodeficiency virus. Journal of the Royal Statistical Society: Series C 46, 463–476.

Liu, Y., Bottai, M., 2009. Mixed-effects models for conditional quantiles with longitudinal data. International Journal of Biostatistics 5, Article 28.

McCullagh, P., 1983. Quasilikelihood functions. Annals of Statistics 11, 59–67.

Mu, Y., Wei, Y., 2009. A dynamic quantile regression transformation model for longitudinal data. Statistica Sinica 19, 1137–1153.

Pang, L., Lu, W., Wang, H.J., 2010. Variance estimation in censored quantile regression via induced smoothing. Computational Statistics and Data Analysis doi:10.1016/j.csda.2010.10.018.

Pollard, D., 1990. Empirical processes: theory and applications, in: NSF-CBMS Regional Conference Series in Probability and Statistics, Institute of Mathematical Statistics, Hayward, California.

Qu, A., Li, R., 2006. Nonparametric modeling and inference function for longitudinal data. Biometrics 62, 379–391.

Qu, A., Lindsay, B.G., Li, B., 2000. Improving generalized estimating equations using quadratic inference functions. Biometrika 87, 823–836.

Reich, B.J., Bondell, H.D., Wang, H.J., 2010. Flexible bayesian quantile regression for independent and clusterd data. Biostatistics 11, 337–352.

Sutradhar, B.C., 2003. An overview on regression models for descrete longitudinal responses. Statistical Science 18, 377–393.

Sutradhar, B.C., 2011. Dynamic Mixed Models for Familial Longitudinal Data. Springer Series in Statistics, New York: Springer. doi:10.1007/978-1-4419-8342-8.

Sutradhar, B.C., Kovacevic, M., 2000. Analyzing ordinal longitudinal survey data: generalized estimating equations approach. Biometrika 87, 837–848.

Tang, C.Y., Leng, C., 2011. Empirical likelihood and quantile regression in longitudinal data analysis. Biometrika 98, 1001–1006.

Walker, A.M., 1968. A note on the asymptotic distribution of sample quantiles. Journal of the Royal Statistical Society: Series B 30, 570–575.

Wang, H., 2009. Inference on quantile regression for heteroscedastic mixed models. Statistica Sinica 19, 1247–1261.

Wang, H., Fygenson, M., 2009. Inference for censored quantile regression models in longitudinal studies. Annals of Statistics 37, 756–781.

Wang, H., He, X., 2007. Detecting differential expressions in genechip microarray studies: a quantile approach. Journal of the American Statistical Association 102, 104–112.

Wang, H., Zhu, Z., Zhou, J., 2009a. Quantile regression in partially linear varying coefficient models. Annals of Statistics 37, 3841–3866.

Wang, Y.G., Shao, Q., Zhu, M., 2009b. Quantile regression withour the curse of unsmoothness. Computational Statistics and Data Analysis 53, 3696–3705.

Wei, Y., He, X., 2006. Conditional growth charts (with discussion). Annals of Statistics 34, 2069–2131.

Wu, H., Zhang, J.T., 2006. Nonparametric regression methods for longitudinal data analysis: mixed-effects modeling approaches. Wiley Series in Probability and Statistics, Wiley-Interscience.

Yu, K., Lu, Z., Stander, J., 2003. Quantile regression: applications and current research areas. Journal of the Royal Statistical Society: Series D 52, 331–350.

Appendix

Table A.1: Biases (*Bias*) and relative efficiencies (*EFF*) to the estimators of β_0, β_1 and β_2 using different methods (AQR, PQR, QLWI and WI) at five quantiles 0.05, 0.25, 0.5, 0.75, 0.95. Where ρ is specified as 0.1, 0.5, and 0.9 respectively, and the error follows a multivariate Chi-squared distribution with two degrees of freedom (*case* 2).

			β_0		β_1		β_2	
τ	ρ	Method	Bias	EFF	Bias	EFF	Bias	EFF
0.05	0.1	AQR	0.0046	0.975	-0.0007	1.028	0.0001	1.083
		PQR	0.0019	1.130	-0.0007	1.086	0.0001	1.144
		QLWI	-0.0007	1.260	-0.0069	1.049	-0.0065	0.747
		WI	0.0023	1.000	-0.0016	1.000	0.0001	1.000
	0.5	AQR	0.0044	0.962	-0.0003	1.043	-0.0009	1.053
		PQR	0.0018	1.106	-0.0004	1.106	-0.0009	1.133
		QLWI	-0.0007	1.227	-0.0067	1.111	-0.0073	0.698
		WI	0.0014	1.000	0.0005	1.000	-0.0010	1.000
	0.9	AQR	0.0053	0.946	-0.0005	1.095	0.0005	1.144
		PQR	0.0024	1.078	-0.0007	1.138	0.0005	1.206
		QLWI	-0.0004	1.147	-0.0068	1.064	-0.0061	0.791
		WI	0.0022	1.000	-0.0006	1.000	0.0005	1.000

Continued on next page

τ	ρ	Method	Bias	EFF	Bias	EFF	Bias	EFF
0.25	0.1	AQR	0.0033	1.048	0.0003	1.037	-0.0002	1.070
		PQR	0.0015	1.054	0.0001	1.036	-0.0003	1.083
		QLWI	0.0021	1.073	0.0016	1.096	0.0011	1.108
		WI	0.0014	1.000	-0.0006	1.000	-0.0004	1.000
	0.5	AQR	0.0058	1.035	-0.0020	1.045	-0.0002	1.097
		PQR	0.0038	1.058	-0.0021	1.053	-0.0002	1.107
		QLWI	0.0044	1.078	-0.0005	1.087	0.0014	1.124
		WI	0.0034	1.000	-0.0023	1.000	-0.0002	1.000
	0.9	AQR	0.0032	1.123	0.0006	1.513	-0.0002	1.513
		PQR	0.0007	1.140	0.0005	1.523	-0.0002	1.522
		QLWI	0.0018	1.045	0.0017	1.085	0.0013	1.086
		WI	0.0003	1.000	0.0008	1.000	-0.0004	1.000
0.5	0.1	AQR	0.0084	1.030	-0.0059	1.037	0.0015	1.042
		PQR	0.0084	1.030	-0.0059	1.037	0.0015	1.042
		QLWI	0.0078	1.022	-0.0045	1.043	0.0034	1.037
		WI	0.0069	1.000	-0.0060	1.000	0.0018	1.000
	0.5	AQR	0.0064	1.033	0.0008	1.061	-0.0004	1.084
		PQR	0.0064	1.033	0.0008	1.061	-0.0004	1.084
		QLWI	0.0057	1.019	0.0025	1.038	0.0013	1.047
		WI	0.0051	1.000	0.0006	1.000	0.0001	1.000
	0.9	AQR	0.0053	1.188	-0.0019	1.982	0.0004	1.864
		PQR	0.0053	1.188	-0.0019	1.982	0.0004	1.864
		QLWI	0.0035	1.016	0.0027	1.045	0.0022	1.056
		WI	0.0024	1.000	0.0010	1.000	0.0008	1.000

Continued on next page

τ	ρ	Method	Bias	EFF	Bias	EFF	Bias	EFF
0.75	0.1	AQR	-0.0000	1.025	-0.0060	1.044	-0.0004	1.073
		PQR	0.0058	1.023	-0.0061	1.046	-0.0002	1.066
		QLWI	0.0025	1.015	-0.0044	1.033	0.0012	1.045
		WI	0.0014	1.000	-0.0056	1.000	-0.0006	1.000
	0.5	AQR	-0.0009	1.064	0.0055	1.083	-0.0002	1.128
		PQR	0.0057	1.059	0.0051	1.083	-0.0003	1.124
		QLWI	0.0024	1.022	0.0062	1.031	0.0004	1.042
		WI	0.0026	1.000	0.0028	1.000	-0.0008	1.000
	0.9	AQR	-0.0061	1.226	0.0047	2.040	-0.0012	2.059
		PQR	0.0022	1.227	0.0047	2.029	-0.0012	2.049
		QLWI	0.0019	1.007	0.0013	1.019	-0.0007	1.024
		WI	0.0010	1.000	-0.0005	1.000	-0.0025	1.000
0.95	0.1	AQR	-0.0047	1.077	-0.0029	1.067	0.0013	1.110
		PQR	0.0455	1.030	-0.0035	1.053	0.0005	1.098
		QLWI	0.0246	1.006	-0.0059	1.012	0.0031	1.017
		WI	0.0168	1.000	0.0045	1.000	0.0014	1.000
	0.5	AQR	-0.0093	1.033	-0.0214	1.083	0.0024	1.111
		PQR	0.0403	1.046	-0.0194	1.107	0.0023	1.136
		QLWI	0.0125	1.013	-0.0117	1.023	0.0042	1.018
		WI	0.0096	1.000	-0.0101	1.000	0.0021	1.000
	0.9	AQR	-0.0142	1.287	-0.0283	1.919	-0.0055	1.866
		PQR	0.0470	1.243	-0.0301	1.857	-0.0059	1.846
		QLWI	0.0257	1.009	-0.0271	1.027	-0.0017	1.032
		WI	0.0251	1.000	-0.0298	1.000	-0.0031	1.000

Table A.2: Biases ($Bias$) and relative efficiencies (EFF) to the estimators of β_0, β_1 and β_2 using different methods (AQR, PQR, QLWI and WI) at five quantiles 0.05, 0.25, 0.5, 0.75, 0.95. Where ρ is specified as 0.1, 0.5, and 0.9 respectively, and the error follows a multivariate T-distribution with three degrees of freedom (*case* 3).

			β_0		β_1		β_2	
τ	ρ	Method	Bias	EFF	Bias	EFF	Bias	EFF
0.05	0.1	AQR	0.0070	1.051	0.0034	1.134	0.0011	1.161
		PQR	-0.0227	1.028	0.0017	1.098	0.0014	1.147
		QLWI	-0.0107	1.019	0.0005	1.025	-0.0031	1.062
		WI	-0.0107	1.000	0.0076	1.000	-0.0002	1.000
	0.5	AQR	0.0071	1.094	-0.0041	1.267	0.0027	1.396
		PQR	-0.0252	1.060	-0.0040	1.244	0.0025	1.341
		QLWI	-0.0103	1.010	-0.0119	1.048	-0.0024	1.042
		WI	-0.0053	1.000	-0.0140	1.000	0.0017	1.000
	0.9	AQR	0.0324	1.289	0.0018	3.024	-0.0011	2.854
		PQR	-0.0033	1.284	-0.0063	3.012	-0.0016	2.742
		QLWI	-0.0006	1.021	0.0066	1.053	-0.0029	1.061
		WI	0.0002	1.000	0.0111	1.000	0.0011	1.000
0.25	0.1	AQR	-0.0021	1.070	-0.0010	1.071	-0.0017	1.096
		PQR	-0.0050	1.058	-0.0011	1.067	-0.0017	1.100
		QLWI	-0.0054	1.043	-0.0029	1.054	-0.0039	1.052
		WI	-0.0042	1.000	-0.0002	1.000	-0.0017	1.000
	0.5	AQR	0.0024	1.112	-0.0032	1.288	-0.0007	1.264
		PQR	-0.0008	1.114	-0.0033	1.276	-0.0007	1.257
		QLWI	-0.0011	1.033	-0.0065	1.038	-0.0034	1.039

Continued on next page

τ	ρ	Method	Bias	EFF	Bias	EFF	Bias	EFF
		WI	0.0002	1.000	-0.0026	1.000	-0.0005	1.000
	0.9	AQR	0.0028	1.236	0.0010	2.835	-0.0000	2.979
		PQR	-0.0017	1.232	0.0011	2.814	-0.0001	2.975
		QLWI	-0.0028	1.026	-0.0011	1.044	-0.0019	1.058
		WI	-0.0018	1.000	0.0025	1.000	0.0008	1.000
0.5	0.1	AQR	0.0004	1.031	-0.0007	1.052	-0.0001	1.046
		PQR	0.0004	1.031	-0.0007	1.052	-0.0001	1.046
		QLWI	0.0003	1.033	-0.0001	1.057	-0.0001	1.052
		WI	0.0005	1.000	-0.0009	1.000	0.0001	1.000
	0.5	AQR	0.0006	1.122	0.0003	1.234	0.0004	1.301
		PQR	0.0006	1.122	0.0003	1.234	0.0004	1.301
		QLWI	0.0003	1.023	0.0012	1.061	0.0001	1.051
		WI	0.0005	1.000	0.0009	1.000	0.0001	1.000
	0.9	AQR	0.0021	1.223	-0.0003	2.916	0.0002	2.740
		PQR	0.0020	1.223	-0.0003	2.916	0.0002	2.740
		QLWI	0.0029	1.032	-0.0011	1.062	0.0008	1.047
		WI	0.0031	1.000	-0.0015	1.000	0.0010	1.000
0.75	0.1	AQR	-0.0010	1.037	-0.0006	1.059	-0.0009	1.077
		PQR	0.0018	1.036	-0.0005	1.052	-0.0009	1.074
		QLWI	0.0021	1.026	0.0015	1.052	0.0015	1.068
		WI	-0.0001	1.000	0.0003	1.000	-0.0010	1.000
	0.5	AQR	-0.0022	1.071	-0.0004	1.185	-0.0003	1.290
		PQR	0.0006	1.078	-0.0003	1.182	-0.0003	1.293
		QLWI	0.0014	1.035	0.0010	1.056	0.0027	1.077

Continued on next page

τ	ρ	Method	Bias	EFF	Bias	EFF	Bias	EFF
		WI	0.0001	1.000	-0.0006	1.000	0.0008	1.000
	0.9	AQR	-0.0014	1.261	-0.0008	2.868	0.0006	2.925
		PQR	0.0030	1.251	-0.0008	2.844	0.0007	2.912
		QLWI	0.0036	1.010	0.0013	1.038	0.0037	1.053
		WI	0.0024	1.000	-0.0025	1.000	0.0011	1.000
0.95	0.1	AQR	-0.0042	1.102	-0.0085	1.144	-0.0035	1.182
		PQR	0.0264	1.060	-0.0083	1.121	-0.0031	1.161
		QLWI	0.0151	1.012	-0.0073	1.039	-0.0009	1.042
		WI	0.0112	1.000	-0.0066	1.000	-0.0049	1.000
	0.5	AQR	-0.0121	1.120	-0.0014	1.365	-0.0038	1.468
		PQR	0.0207	1.121	-0.0004	1.339	-0.0040	1.449
		QLWI	0.0062	1.033	0.0075	1.033	-0.0035	1.082
		WI	0.0070	1.000	0.0001	1.000	-0.0063	1.000
	0.9	AQR	-0.0216	1.149	0.0002	2.978	0.0003	2.825
		PQR	0.0166	1.117	0.0006	2.878	0.0003	2.766
		QLWI	0.0062	0.997	0.0056	1.049	0.0074	1.051
		WI	0.0003	1.000	0.0109	1.000	0.0043	1.000

Table A.3: Standard deviation (SD) of 1000 estimates of three β-parameters, the sample average of 1000 estimated standard errors (SE) and the probability of estimates falling into the 95% confidence interval are reported for different methods (AQR, PQR, QLWI and WI) at five quantiles 0.05, 0.25, 0.5, 0.75, 0.95. Where ρ is specified as 0.1, 0.5, and 0.9 respectively, and the error follows a multivariate chi-squared distribution (*case* 2).

			β_0			β_1			β_2		
τ	ρ	Method	SD	SE	$P_{0.95}$	SD	SE	$P_{0.95}$	SD	SE	$P_{0.95}$
0.05	0.1	AQR	0.0139	0.0139	0.948	0.0198	0.0196	0.938	0.0096	0.0095	0.952
		PQR	0.0135	0.0137	0.953	0.0193	0.0193	0.943	0.0093	0.0094	0.961
	0.5	AQR	0.0142	0.0140	0.941	0.0201	0.0196	0.938	0.0100	0.0093	0.935
		PQR	0.0138	0.0138	0.952	0.0195	0.0193	0.943	0.0097	0.0092	0.938
	0.9	AQR	0.0160	0.0150	0.936	0.0196	0.0187	0.932	0.0096	0.0090	0.930
		PQR	0.0156	0.0148	0.932	0.0192	0.0184	0.929	0.0094	0.0089	0.927
0.25	0.1	AQR	0.0356	0.0356	0.948	0.0490	0.0498	0.953	0.0252	0.0246	0.939
		PQR	0.0356	0.0356	0.949	0.0490	0.0497	0.955	0.0251	0.0246	0.940
	0.5	AQR	0.0385	0.0367	0.948	0.0510	0.0496	0.949	0.0245	0.0246	0.946
		PQR	0.0383	0.0367	0.945	0.0509	0.0495	0.948	0.0244	0.0246	0.948
	0.9	AQR	0.0429	0.0419	0.944	0.0412	0.0411	0.950	0.0213	0.0206	0.939

Continued on next page

τ	ρ	Method	SD	SE	$P_{0.95}$	SD	SE	$P_{0.95}$	SD	SE	$P_{0.95}$
		PQR	0.0427	0.0418	0.946	0.0411	0.0410	0.951	0.0212	0.0206	0.937
0.5	0.1	AQR	0.0646	0.0622	0.938	0.0875	0.0868	0.948	0.0455	0.0437	0.948
		PQR	0.0646	0.0622	0.938	0.0875	0.0868	0.948	0.0455	0.0437	0.948
	0.5	AQR	0.0651	0.0647	0.946	0.0880	0.0855	0.953	0.0438	0.0427	0.945
		PQR	0.0651	0.0647	0.947	0.0880	0.0855	0.953	0.0438	0.0427	0.945
	0.9	AQR	0.0762	0.0754	0.946	0.0628	0.0628	0.944	0.0316	0.0318	0.955
		PQR	0.0762	0.0754	0.946	0.0628	0.0628	0.944	0.0316	0.0318	0.955
0.75	0.1	AQR	0.1018	0.1053	0.959	0.1419	0.1484	0.963	0.0767	0.0753	0.946
		PQR	0.1017	0.1055	0.958	0.1218	0.1486	0.959	0.0769	0.0754	0.946
	0.5	AQR	0.1113	0.1108	0.952	0.1471	0.1468	0.953	0.0751	0.0754	0.958
		PQR	0.1114	0.1110	0.954	0.1471	0.1471	0.954	0.0752	0.0756	0.959
	0.9	AQR	0.1329	0.1313	0.943	0.1083	0.1063	0.953	0.0558	0.0524	0.931
		PQR	0.1329	0.1317	0.942	0.1086	0.1066	0.952	0.0559	0.0526	0.930
0.95	0.1	AQR	0.2401	0.2504	0.961	0.3638	0.3571	0.943	0.1756	0.1752	0.946
		PQR	0.2412	0.2535	0.964	0.3661	0.3620	0.942	0.1765	0.1773	0.953
	0.5	AQR	0.2678	0.2621	0.948	0.3652	0.3483	0.934	0.1803	0.1749	0.945

Continued on next page

τ	ρ	Method	SD	SE	$P_{0.95}$	SD	SE	$P_{0.95}$	SD	SE	$P_{0.95}$
		PQR	0.2632	0.2654	0.952	0.3613	0.3533	0.942	0.1783	0.1776	0.951
	0.9	AQR	0.2936	0.3039	0.956	0.2773	0.2578	0.935	0.1341	0.1281	0.943
		PQR	0.2954	0.3080	0.958	0.2817	0.2619	0.931	0.1349	0.1304	0.943

Table A.4: Standard deviation (SD) of 1000 estimates of three β-parameters, the sample average of 1000 estimated standard errors (SE) and the probability of estimates falling into the 95% confidence interval are reported for different methods (AQR, PQR, QLWI and WI) at five quantiles 0.05, 0.25, 0.5, 0.75, 0.95. Where ρ is specified as 0.1, 0.5, and 0.9 respectively, and the error follows a multivariate T-distribution (*case* 3).

			β_0			β_1			β_2		
τ	ρ	Method	SD	SE	$P_{0.95}$	SD	SE	$P_{0.95}$	SD	SE	$P_{0.95}$
0.05	0.1	AQR	0.1638	0.1509	0.934	0.2023	0.1886	0.934	0.0993	0.0940	0.926
		PQR	0.1643	0.1536	0.941	0.2056	0.1922	0.931	0.0999	0.0940	0.934
	0.5	AQR	0.1755	0.1591	0.935	0.2001	0.1761	0.913	0.0904	0.0834	0.929
		PQR	0.1766	0.1622	0.936	0.0202	0.1792	0.918	0.0923	0.0850	0.937
	0.9	AQR	0.1798	0.1710	0.938	0.1222	0.1115	0.932	0.0588	0.0539	0.932
		PQR	0.1831	0.1736	0.938	0.2543	0.1138	0.925	0.0599	0.0551	0.941
0.25	0.1	AQR	0.0557	0.0541	0.938	0.0732	0.0722	0.952	0.0368	0.0367	0.957
		PQR	0.0559	0.0542	0.938	0.0733	0.0723	0.950	0.0367	0.0368	0.956
	0.5	AQR	0.0587	0.0583	0.953	0.0679	0.0666	0.938	0.0324	0.0322	0.945
		PQR	0.0587	0.0584	0.954	0.0682	0.0667	0.937	0.0325	0.0322	0.946
	0.9	AQR	0.0659	0.0667	0.944	0.0437	0.0435	0.955	0.0218	0.0217	0.950

Continued on next page

τ	ρ	Method	SD	SE	$P_{0.95}$	SD	SE	$P_{0.95}$	SD	SE	$P_{0.95}$
		PQR	0.0661	0.0669	0.949	0.0438	0.0436	0.958	0.0219	0.0217	0.948
0.5	0.1	AQR	0.0430	0.0421	0.944	0.0596	0.0588	0.946	0.0292	0.0297	0.953
		PQR	0.0430	0.0421	0.944	0.0588	0.0588	0.946	0.0292	0.0297	0.953
	0.5	AQR	0.0474	0.0467	0.948	0.0556	0.0543	0.946	0.0270	0.0264	0.942
		PQR	0.0474	0.0467	0.948	0.0556	0.0543	0.946	0.0270	0.0264	0.942
	0.9	AQR	0.0531	0.0542	0.953	0.0358	0.0360	0.958	0.0185	0.0180	0.946
		PQR	0.0531	0.0542	0.953	0.0358	0.0360	0.958	0.0185	0.0180	0.946
0.75	0.1	AQR	0.0541	0.0534	0.945	0.0736	0.0721	0.949	0.0362	0.0358	0.959
		PQR	0.0541	0.0535	0.944	0.0739	0.0723	0.947	0.0363	0.0358	0.955
	0.5	AQR	0.0602	0.0575	0.941	0.0730	0.0663	0.927	0.0332	0.0322	0.950
		PQR	0.0604	0.0576	0.941	0.0731	0.0665	0.927	0.0333	0.0322	0.948
	0.9	AQR	0.0666	0.0657	0.954	0.0440	0.0439	0.952	0.0221	0.0216	0.941
		PQR	0.0668	0.0659	0.955	0.0442	0.0440	0.952	0.0222	0.0216	0.939
0.95	0.1	AQR	0.1562	0.1484	0.944	0.1994	0.1906	0.944	0.1011	0.0929	0.934
		PQR	0.1571	0.1513	0.957	0.2015	0.1947	0.944	0.1020	0.0948	0.945
	0.5	AQR	0.1663	0.1584	0.943	0.1870	0.1767	0.937	0.0888	0.0869	0.948

Continued on next page

τ	ρ	Method	SD	SE	$P_{0.95}$	SD	SE	$P_{0.95}$	SD	SE	$P_{0.95}$
		PQR	0.1654	0.1619	0.946	0.1888	0.1804	0.935	0.0893	0.0886	0.957
	0.9	AQR	0.1803	0.1707	0.940	0.1185	0.1120	0.941	0.0611	0.0560	0.923
		PQR	0.1835	0.1750	0.944	0.1206	0.1146	0.939	0.0617	0.0572	0.926