UPPSALA
UNIVERSITET

Quantile Regression Estimation of Nonlinear Longitudinal Data

Andreas Karlsson

Division of Statistics
Research Report 2004:5

Research Report Series
ISSN 1403-7572

Department of Information Science
P.O. Box 513
SE-751 20 Uppsala, Sweden

Quantile Regression Estimation of Nonlinear Longitudinal Data

Andreas Karlsson
Uppsala University, Sweden

December 9, 2004

Abstract

This paper examines a weighted version of the quantile regression estimator defined by Koenker and Bassett (1978), adjusted to the case of nonlinear longitudinal data. Different weights are used and compared by computer simulation using a four-parameter logistic growth function and error terms following an AR(1) model. It is found that the estimator is performing quite well, especially for the median regression case, that the differences between the weights are small, and that increasing the correlation in the AR(1) model leads to better behaviour of the estimator. A comparison is made with the corresponding mean regression estimator, which is found to be less robust. Finally the estimator is applied to a data set with growth patterns of two genotypes of soybean.

AMS (2000) *subject classification.* 62M10, 62G30.

Key words and phrases. Quantile regression, longitudinal data, nonlinear model, Monte Carlo simulation.

1 Introduction

In many applied statistical studies, the observed data are from subjects measured repeatedly over time. This kind of data appear in such disparate disciplines as economics, sociology, agricultural science, psychology and medicine. It is known as repeated measures, longitudinal data or, especially in economics and sociology, panel data. Some authors distinguish between these terms and give them different meanings. In this paper the term longitudinal data will be used, and it is defined as data collected at multiple occasions from different subjects over some period of time.

Longitudinal data can be seen as a time series of cross-sectional data, so that one can follow the development of the cross-sectional data set over time. Compared to ordinary time series, which usually consists of a single long series, longitudinal data usually consists of many but shorter time series. This gives the data a special structure, that need to be dealt with when the data is analyzed. Specifically, for longitudinal data are observations from different subjects assumed to be independent of each other, while observations from the same subject are assumed to be correlated. The main advantages of longitudinal data is that it is possible to follow the individual patterns of change for each subject, and that in the inference process one can borrow strength across subjects. For details about the merits of longitudinal data, see Zeger and Liang (1992), Davis (2002) and Diggle *et al.* (2002).

There are a number of different methods available for analyzing longitudinal data, like MANOVA, repeated measures ANOVA, mixed models, Generalized Linear Models (with its extension Generalized Estimation Equations) and different nonlinear models. For details about the different approaches, see Davidian and Giltinan (1995), Davis (2002) and Diggle *et al.* (2002). A common feature of these methods is that they usually use the mean as the measure of centrality. As is well known, the mean is not a good measure of centrality for skewed data. In such cases it often has low efficiency, compared to other estimators. But what is more important in

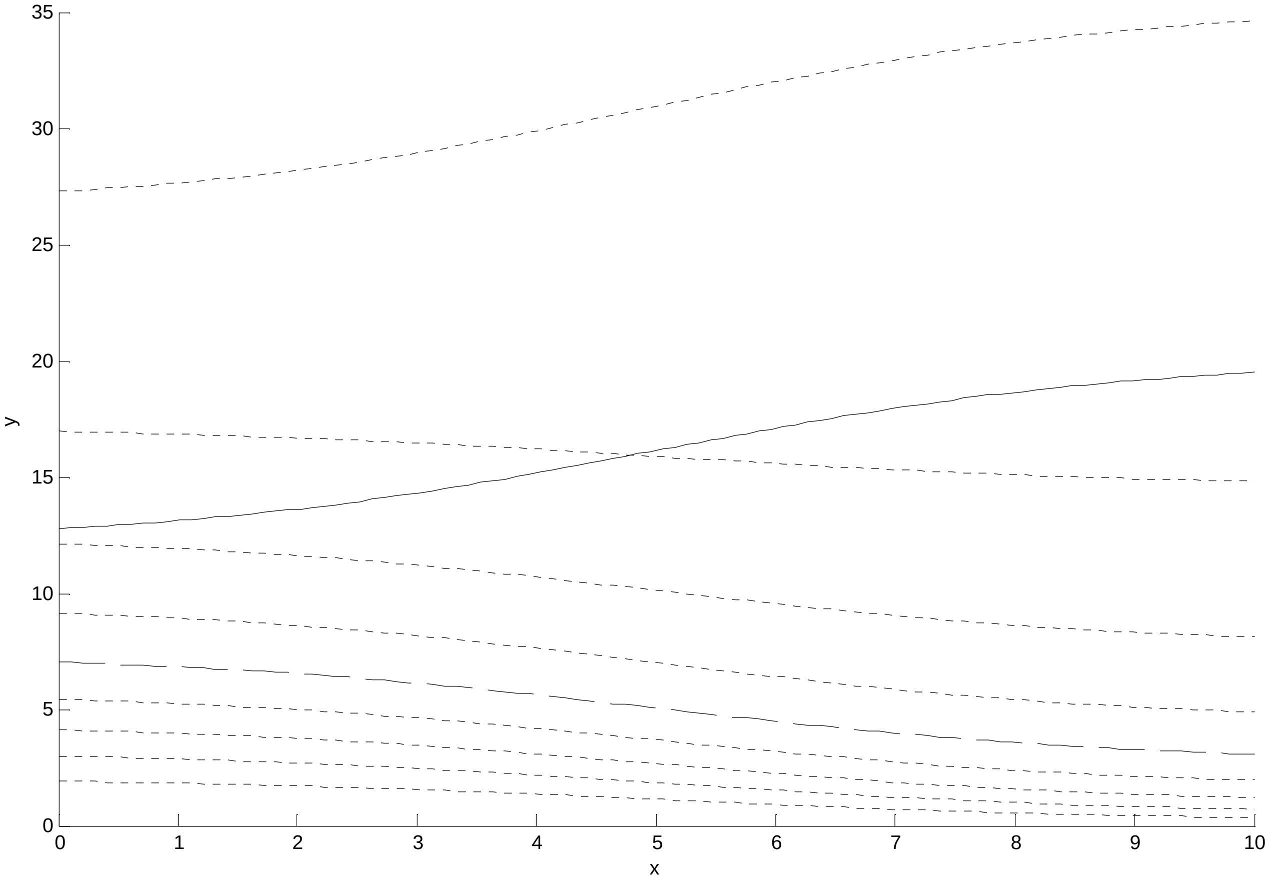

Figure 1: Regression lines for quantile regressions (dotted; median regression dashed) and mean regression (solid) for a logistic growth function. The data are distributed as lognormal(2,1) at minus infinity and lognormal(1,2) at plus infinity.

a practical situation is that it is difficult, especially for non-statisticians, to interpret what the mean measures when the data is skewed. In contrast to this the median, besides often having higher efficiency than the mean for skewed data (see e.g. Koenker and Bassett, 1978), always has an easy interpretation. Since it is common with skewed data in practical situations, and one don't know the distribution of the underlying population, this makes the median a preferable measure to use, and motivates the use of median regression for longitudinal data, instead of the mean regression.

Median regression, in turn, is a special case of quantile regression (Koenker and Bassett, 1978), which makes it possible to characterize any arbitrary quantile regression line for a data set. Thus, by calculating quantile regression lines for different quantiles for the same data set, one gets a distribution of regression lines that estimates the shape of the underlying distribution of the data for the dependent variable conditional on every single dependent variable. This makes it possible to study the change over time in the shape of the entire conditional distribution of longitudinal data and not only the change in the conditional mean or median. Thus, one can study the development over time for different parts of the distribution and compare these to each other. For example, the slope of a curve need not be the same for the 10th, 50th and 90th quantiles, and a quantile regression estimator could estimate these different slopes. This gives a more complete description of the data than the mean or the median can give.

Figure 1 illustrates some of the advantages of quantile regression compared with median and mean regression. It gives the nine regression lines for quantile regressions with quantiles $0.10, ..., 0.90$, (dotted; median regression dashed) and the mean regression (solid) for the nonlinear four-parameter logistic growth function

$$f(\mathbf{x}_{it}, \boldsymbol{\beta}) = \beta_1 + \frac{\beta_2 - \beta_1}{1 + \exp(\beta_4 (x_{it} - \beta_3))}, \tag{1}$$

where β_1 and β_2 are distributed as lognormal(2,1) and lognormal(1,2), respectively, while $\beta_3 = 5$ and $\beta_4 = -0.5$ (see Section 2.1). The mean regression shows a marked positive slope, although 80 percent of the observations have a negative slope. This is due to the heavily skewed distributions used, and that the distribution is not the same for the x values. The negative slope of the median regression line tells us that at least half of the observations have a negative slope, but only the distribution of quantile regression lines show where the slopes are positive and where they are negative. Also, note that the quantile regression lines are the contour curves for the three dimensional distribution of the data, so that the spacing of the curves show the relative steepness at different parts of the distribution.

Of course, there are some drawbacks with using quantile or median regression instead of mean regression. There exist no closed form formulas with explicit solutions for the estimators, which means that it is harder to compute e.g. the asymptotic properties, at least when using linear functions, where the mean regression has explicit solutions. The estimates must be calculated with numerical methods, which historically has been a great drawback, because it was time consuming. Nowadays, with high-speed computers and optimization algorithms, this problem is almost negligible. What still can be a problem, though, is the possibility that the numerical optimization methods converge to a local optimum instead of a global optimum. But this is a problem also for nonlinear mean regression.

Very little has been written about the use of quantile or median regression for longitudinal data, and what has been written has mostly been only about median regression and linear models. Koenker (2004) examines quantile regression of longitudinal data, especially penalized quantile regression, for a linear model with fixed effects. Jung (1996) develops a quasi-likelihood method for median regression models for longitudinal data, Lipsitz *et al.* (1997) look at quantile regression for longitudinal data when there are missing data, and He *et al.* (2003) compare three different estimators of the parameters for median regression of longitudinal data in linear models. He and Kim (2002) and He *et al.* (2002) examine the use of median regression in semiparametric models for longitudinal data, while Yu (2004) look at nonparametric quantile regression for longitudinal data analysis.

There is thus a lack of a coherent examination of quantile regression estimation for longitudinal data. This paper will try to fill parts of this gap by examining a general method for quantile regression estimation of longitudinal data based on the method of Koenker and Bassett (1978), applicable for both linear and nonlinear models. This examination will be done by investigating the performance of the estimation method for small data sets by Monte Carlo simulations of a nonlinear model when different weights are used for the estimation method. The special case of median regression will also be compared with the mean regression case. The nonlinear model is chosen because it is more flexible than a linear model, and has many interesting applications in biostatistics. For the use of quantile regression in nonlinear models, see Jurečková and Procházka (1994), and for the use of nonlinear models for longitudinal data, see Davidian and Giltinan (1995) and Davidian and Giltinan (2003) .

The outline of this paper is as follows: Section 2 will present the longitudinal data model used in this paper, and introduce the general quantile regression estimation method and the different weights used. Section 3 presents the results from Monte Carlo simulations where the performance of the different weights are compared with each other and with a mean regression estimation method. In Section 4 the methods are applied to a data set with growth patterns of two genotypes of soybean, while Section 5 contains a discussion of the conclusions that can be drawn from this paper.

2 Model and methodology

The notation is as follows (cf. Diggle *et al.*, 2002): for $i = 1, ..., n$, $t = 1, ..., T_i$ and $k = 1, ..., p$, let y_{it} denote the observed value of the response variable for observation t of subject i at time τ_{it} and x_{itk} the corresponding observed value of explanatory variable k, so that $\mathbf{x}_{it} = (x_{it1}, ..., x_{itp})'$ is the column vector of length p that contains the values of all the explanatory variables for observation t of subject i at time τ_{it}. If one then has error terms ε_{it}, the general model can be formulated as

$$y_{it} = f(\mathbf{x}_{it}, \boldsymbol{\beta}) + \varepsilon_{it}, \quad i = 1, ..., n, \; t = 1, ..., T_i, \tag{2}$$

where $f(\mathbf{x}_{it}, \boldsymbol{\beta})$ is a known response function, possible nonlinear, and $\boldsymbol{\beta} = (\beta_1, .., \beta_h)'$ is a column vector of length h with unknown parameters. This is thus a form of the marginal analysis approach for longitudinal data discussed in Diggle *et al.* (2002).

To write this in matrix terms, for subject i measured at times $\tau_{i1}, ..., \tau_{iT_i}$, let $\mathbf{y}_i = (y_{i1}, ..., y_{iT_i})'$ be the column vector of length T_i with response variables, $\mathbf{X}_i = (\mathbf{x}_{i1}, ..., \mathbf{x}_{iT_i})'$ the $T_i \times p$ matrix of explanatory variables with $\mathbf{x}'_{it}$ in row t, and $\boldsymbol{\varepsilon}_i = (\varepsilon_{i1}, ..., \varepsilon_{iT_i})'$ the corresponding column vector of length T_i with error terms. Then Model (2) can be written in matrix terms as

$$\mathbf{y}_i = f(\mathbf{X}_i, \boldsymbol{\beta}) + \boldsymbol{\varepsilon}_i, \quad i = 1, ..., n. \tag{3}$$

In this model, the error terms $\boldsymbol{\varepsilon}_1, ..., \boldsymbol{\varepsilon}_n$ from different subjects are assumed to be independent of each other, but the error terms $\varepsilon_{i1}, ..., \varepsilon_{iT_i}$ from a single subject are assumed to be correlated to an unknown degree and with an unknown correlation structure.

Having formulated the model to use, the next step is to make a quantile regression estimation of the parameter vector $\boldsymbol{\beta}$. This will be done by first formulating the special case of a least absolute deviation (LAD) median regression estimator and then extend the idea behind this to the general quantile regression case based on the method used in Koenker and Bassett (1978).

Since the data are correlated, a natural idea is to consider a weighted LAD estimator (see Zhao (2001) for the non-longitudinal case) that could take this correlation into account. Such an estimator can be formulated as

$$\widehat{\boldsymbol{\beta}}_{\min} = \min_{\boldsymbol{\beta}} \sum_{i=1}^{n} \sum_{t=1}^{T_i} w_{it} \left| y_{it} - f(\mathbf{x}_{it}, \boldsymbol{\beta}) \right|, \tag{4}$$

where w_{it} is the weight for observation t of subject i at time τ_{ij}. This estimates the parameter vector $\boldsymbol{\beta}$ for a median regression. This idea can then be extended to the general case of a weighted quantile regression estimator, similar to the unweighted one used in Koenker and Bassett (1978). For $0 < \theta < 1$, let

$$\rho_\theta(y_{it} - f(\mathbf{x}_{it}, \boldsymbol{\beta})) = (y_{it} - f(\mathbf{x}_{it}, \boldsymbol{\beta}))(\theta - I\{y_{it} - f(\mathbf{x}_{it}, \boldsymbol{\beta}) < 0\}), \tag{5}$$

where $I\{\bullet\}$ is the indicator function taking the value 1 if $\{\bullet\}$ is true and 0 otherwise, and let

$$\boldsymbol{\rho}_\theta(\mathbf{y}_i - f(\mathbf{X}_i, \boldsymbol{\beta})) = (\rho_\theta(y_{i1} - f(\mathbf{x}_{i1}, \boldsymbol{\beta})), ..., \rho_\theta(y_{iT_i} - f(\mathbf{x}_{iT_i}, \boldsymbol{\beta})))' \tag{6}$$

be the corresponding column vector of length T_i. Then, the general θ-th quantile regression estimator can be formulated as

$$\widehat{\boldsymbol{\beta}}_{\min}(\theta) = \min_{\boldsymbol{\beta}} \sum_{i=1}^{n} \sum_{t=1}^{T_i} w_{it} \rho_\theta(y_{it} - f(\mathbf{x}_{it}, \boldsymbol{\beta})), \tag{7}$$

where the special case $\theta = 0.5$ estimates the median regression, i.e., the same thing as (4), although these two estimators are not exactly the same. Estimator (7) is thus a generalization to the case of correlated longitudinal data of the simple quantile regression estimator in Koenker and Bassett (1978), making it possible to give different weights to different observations t and subjects i when estimating the quantile regressions.

To generalize the results above to the matrix case, let $\mathbf{1}_{T_i} = (1, ..., 1)'$ be a $T_i \times 1$ vector of 1's and $\mathbf{W}_i$ a symmetric $T_i \times T_i$ weight matrix with elements $w_{itt'}$, $t, t' = 1, ..., T_i$, for observation t of subject i. Also, for a matrix $\mathbf{A}$, let $|\mathbf{A}|$ denote a matrix consisting of the absolute values of the elements of $\mathbf{A}$, so that, e.g.,

$$|\mathbf{y}_i - f(\mathbf{X}_i, \boldsymbol{\beta})| = (|y_{i1} - f(\mathbf{x}_{i1}, \boldsymbol{\beta})|, ..., |y_{iT_i} - f(\mathbf{x}_{iT_i}, \boldsymbol{\beta})|)'. \tag{8}$$

Now, the generalized least squares (GLS) estimator of the conditional mean regression is given by

$$\widetilde{\boldsymbol{\beta}}_{\min} = \min_{\boldsymbol{\beta}} \sum_{i=1}^{n} (\mathbf{y}_i - f(\mathbf{X}_i, \boldsymbol{\beta}))' \mathbf{W}_i (\mathbf{y}_i - f(\mathbf{X}_i, \boldsymbol{\beta})). \tag{9}$$

Based on this a generalized least absolute deviation (GLAD) estimator can be constructed by replacing the least squares element $(\mathbf{y}_i - f(\mathbf{X}_i, \boldsymbol{\beta}))' \mathbf{W}_i (\mathbf{y}_i - f(\mathbf{X}_i, \boldsymbol{\beta}))$ in (9) with the corresponding LAD element $\mathbf{1}'_{T_i} \mathbf{W}_i |\mathbf{y}_i - f(\mathbf{X}_i, \boldsymbol{\beta})|$, so that the generalized form of Estimator (4), the conditional median regression, is given by

$$\begin{aligned}
\widehat{\boldsymbol{\beta}}_{\min} &= \min_{\boldsymbol{\beta}} \sum_{i=1}^{n} \mathbf{1}'_{T_i} \mathbf{W}_i |\mathbf{y}_i - f(\mathbf{X}_i, \boldsymbol{\beta})| \qquad (10) \\
&= \min_{\boldsymbol{\beta}} \sum_{i=1}^{n} \begin{pmatrix} 1_1 & \cdots & 1_{T_i} \end{pmatrix} \begin{pmatrix} w_{i11} & w_{i12} & \cdots & w_{i1T_i} \\ w_{i21} & w_{i22} & & \\ \vdots & & \ddots & \\ w_{iT_i1} & & & w_{iT_iT_i} \end{pmatrix} \begin{pmatrix} |y_{i1} - f(\mathbf{x}_{i1}, \boldsymbol{\beta})| \\ \vdots \\ |y_{iT_i} - f(\mathbf{x}_{iT_i}, \boldsymbol{\beta})| \end{pmatrix} \\
&= \min_{\boldsymbol{\beta}} \sum_{i=1}^{n} \begin{pmatrix} \sum_{t=1}^{T_i} w_{it1} & \cdots & \sum_{t=1}^{T_i} w_{itT_i} \end{pmatrix} \begin{pmatrix} |y_{i1} - f(\mathbf{x}_{i1}, \boldsymbol{\beta})| \\ \vdots \\ |y_{iT_i} - f(\mathbf{x}_{iT_i}, \boldsymbol{\beta})| \end{pmatrix} \\
&= \min_{\boldsymbol{\beta}} \sum_{i=1}^{n} \left(\sum_{t=1}^{T_i} w_{it1} |y_{i1} - f(\mathbf{x}_{i1}, \boldsymbol{\beta})| + \cdots + \sum_{t=1}^{T_i} w_{itT_i} |y_{iT_i} - f(\mathbf{x}_{iT_i}, \boldsymbol{\beta})| \right) \\
&= \min_{\boldsymbol{\beta}} \sum_{i=1}^{n} \sum_{t'=1}^{T_i} \sum_{t=1}^{T_i} w_{itt'} |y_{it'} - f(\mathbf{x}_{it'}, \boldsymbol{\beta})| \\
&= \min_{\boldsymbol{\beta}} \sum_{i=1}^{n} \sum_{t=1}^{T_i} \sum_{t'=1}^{T_i} w_{it't} |y_{it} - f(\mathbf{x}_{it}, \boldsymbol{\beta})| \\
&= \min_{\boldsymbol{\beta}} \sum_{i=1}^{n} \sum_{t=1}^{T_i} \sum_{t'=1}^{T_i} w_{itt'} |y_{it} - f(\mathbf{x}_{it}, \boldsymbol{\beta})|,
\end{aligned}$$

since $\mathbf{W}_i$ is symmetric. This bear an obvious resemblance to (4). When all $\mathbf{W}_i$ matrices are diagonal with the weights w_{it}, $t = 1, ..., T_i$, on the diagonals, then Estimator (10) is the same as Estimator (4). The corresponding matrix formula for the general case of a quantile regression is obviously given by

$$\begin{aligned}
\widehat{\boldsymbol{\beta}}_{\min}(\theta) &= \min_{\boldsymbol{\beta}} \sum_{i=1}^{n} \mathbf{1}'_{T_i} \mathbf{W}_i \boldsymbol{\rho}_\theta (\mathbf{y}_i - f(\mathbf{X}_i, \boldsymbol{\beta})) \qquad (11) \\
&= \min_{\boldsymbol{\beta}} \sum_{i=1}^{n} \sum_{t=1}^{T_i} \sum_{t'=1}^{T_i} w_{itt'} \rho_\theta (y_{it} - f(\mathbf{x}_{it}, \boldsymbol{\beta})),
\end{aligned}$$

with the special case $\theta = 0.5$ estimating the median regression.

Note that it is not assumed in the models and estimators above that the subjects are measured equally many times, i.e., that there are equally many observations t for every subject i, nor that all subjects are measured at the same time points. However, to make it simple, throughout this paper it will be assumed that all subjects are measured equally many times and at the same time points, and that there are no missing values. When, in fact, equally many observations t for every subject i is used, the number of observations will be denoted with T, i.e., $T = T_1 = \cdots = T_n$.

2.1 Model specification

Model (2) is very general, with lots of possible specifications for the response function $f(\mathbf{x}_{it}, \boldsymbol{\beta})$. This paper will be restricted to use the nonlinear four-parameter logistic growth function

$$f(\mathbf{x}_{it}, \boldsymbol{\beta}) = \beta_1 + \frac{\beta_2 - \beta_1}{1 + \exp(\beta_4 (x_{it} - \beta_3))}, \tag{12}$$

where x_{it} is the time. This is a common function used for modelling growth in biological data. The parameters β_1 and β_2 give the values for the dependent variable at minus infinity and plus infinity, respectively, so that $|\beta_2 - \beta_1|$ gives the distance between these two asymptotic values. The parameter β_3 gives the EC_{50} value, the value of x for which the value of the dependent variable is half of the distance between β_1 and β_2, i.e., half the value of $|\beta_2 - \beta_1|$, while β_4 is a slope parameter that governs the steepness of the growth curve. Note that it is not assumed that the ε_{it}'s are normally distributed.

2.2 Estimator specifications

Just like Model (2), the quantile regression estimator (7) is also very general, since there are many possible specifications for the weight w_{it}. If no correlation existed in the data set, so that all ε_{it}'s were independent, the weight $w_{it} = 1$, i.e., all subjects i and observations t have the same weight, would be natural to use. This weight is thus a natural benchmark to compare other weights with, to see if weights that take correlation into account are worthwhile. Also, it has been shown (He *et al.*, 2002; He *et al.*, 2003) that when using this weight to estimate the median regression for longitudinal data one consistently estimates $\boldsymbol{\beta}$ for a linear response function, and that it in that case also performs well compared to other estimators. Since the estimators that use weights that take the correlation among the errors ε_{it}'s into account somehow most have residuals e_{it} to compute the weights from, it is a natural choice to compute these residuals based on the weight $w_{it} = 1$, which can be calculated without knowing the residuals e_{it}. This is done by first estimating

$$\widehat{\boldsymbol{\beta}}_{\min}(0.5) = \min_{\boldsymbol{\beta}} \sum_{i=1}^{n} \sum_{t=1}^{T_i} \rho_{0.5}(y_{it} - f(\mathbf{x}_{it}, \boldsymbol{\beta})), \tag{13}$$

i.e., the $\widehat{\boldsymbol{\beta}}_{\min}$-values are calculated when $w_{it} = 1$ and $\theta = 0.5$, and then with the help of these values calculate the estimated errors as

$$e_{it} = y_{it} - f\left(\mathbf{x}_{it}, \widehat{\boldsymbol{\beta}}_{\min}(0.5)\right), \quad i = 1, ..., n, \ t = 1, ..., T_i. \tag{14}$$

The weights to be used will then be based on these e_{it}-values. It would of course be possible to instead of using $\widehat{\boldsymbol{\beta}}_{\min}(0.5)$ in (13) use $\widehat{\boldsymbol{\beta}}_{\min}(\theta)$, i.e., the same quantile θ that one will estimate in (7), and maybe this seems to be a more natural choice. But after having compared those two choices in a preliminary simulation study, it was found that the choice $\widehat{\boldsymbol{\beta}}_{\min}(0.5)$ performed somewhat better than $\widehat{\boldsymbol{\beta}}_{\min}(\theta)$, in that weights based on this generally had lower MAPE values for the quantiles $\theta = 0.75$ and $\theta = 0.9$. For details, see Appendix. A possible explanation of this result could be that the median regression estimate $\widehat{\boldsymbol{\beta}}_{\min}(0.5)$ has a lower bias and is more robust than the estimate $\widehat{\boldsymbol{\beta}}_{\min}(\theta)$, since the latter estimates a quantile regression line that is located in the tail of a distribution, giving it lower probability mass, and thus should be harder to estimate correctly than the median regression line.

The weights will take into account the dispersion of the e_{it}'s for either the time points or the subjects. Somehow these dispersions have to be measured. In this paper two different measures will be used. The first natural choice is to measure the dispersions with the variances and covariances of the residual vectors $\mathbf{e}_i = (e_{i1}, ..., e_{iT_i})$. Let $\mathbf{E} = (\mathbf{e}_1, ..., \mathbf{e}_n)$ be the $n \times T$ matrix of residuals and $\mathbf{1}_n = (1, ..., 1)'$ an $n \times 1$ vector of 1's, so that

$$\mathbf{S} = \frac{1}{n}\left(\mathbf{E}'\mathbf{E} - \frac{1}{n}\mathbf{E}'\mathbf{1}_n\mathbf{1}_n'\mathbf{E}\right) \tag{15}$$

is the estimated $T \times T$ covariance matrix for the residuals with elements

$$s_{tt'} = \frac{1}{n}\sum_{i=1}^{n}\left(e_{it} - \overline{e}_t\right)\left(e_{it'} - \overline{e}_{t'}\right), \quad t, t' = 1, ..., T, \tag{16}$$

where

$$\overline{e}_t = \frac{1}{n}\sum_{i=1}^{n} e_{it}, \quad t = 1, ..., T. \tag{17}$$

Weights based on $\mathbf{S}$ use both the variances and covariances of the e_{it}'s for the time points. Those weights that only take the variances into account are based on the s_t^2's on the diagonal of $\mathbf{S}$, i.e.,

$$s_t^2 = \frac{1}{n}\sum_{i=1}^{n}\left(e_{it} - \overline{e}_t\right)^2, \quad t = 1, ..., T. \tag{18}$$

For weights that use the variances of the e_{it}'s for the individual subjects the formula

$$s_i^2 = \frac{1}{T_i}\sum_{t=1}^{T_i}\left(e_{it} - \overline{e}_i\right)^2, \quad i = 1, ..., n \tag{19}$$

is used, where the $\overline{e}_i$'s are given by

$$\overline{e}_i = \frac{1}{T_i}\sum_{t=1}^{T_i} e_{it}, \quad i = 1, ..., n. \tag{20}$$

An alternative measure of the dispersion of the e_{it}'s is to use the mean absolute deviation instead of the squared deviation in (18) and (19), which gives the formulas

$$u_t = \frac{1}{n}\sum_{i=1}^{n}\left|e_{it} - \overline{e}_t\right|, \quad t = 1, ..., T \tag{21}$$

and

$$u_i = \frac{1}{T_i}\sum_{t=1}^{T_i}\left|e_{it} - \overline{e}_i\right|, \quad i = 1, ..., n, \tag{22}$$

for weights taking into account the dispersion of the e_{it}'s for the time points and the subjects, respectively.

The weights w_{it} should give more weight to observations with small dispersion values than to those with large dispersion values, and thus give them a greater influence in the estimation of the parameter values. This is achieved by using the inverse of the dispersion measures. For (18) and (19) one can use either s_t^2 and

s_i^2 or their corresponding square roots $\sqrt{s_t^2}$ and $\sqrt{s_i^2}$. After extensive preliminary Monte Carlo-simulations comparing the mean absolute percent errors (MAPE) of these it was found that the square roots were to be preferred. For the weight matrix $\mathbf{W}_i$ it would be natural to use $\mathbf{W}_i = \mathbf{S}^{-1}$. However, this is not possible, since this leads to great convergence problems, with convergence to extremely large (almost infinite) values of MAPE. Instead $\mathbf{W}_i = \left|\mathbf{S}^{-1}\right|$ has to be used, with $\left|\mathbf{S}^{-1}\right|$ denoting a matrix consisting of the absolute values of the elements of $\mathbf{S}^{-1}$. Another possible choice for $\mathbf{W}_i$ is to let each element $w_{itt'}$ be the inverse of the corresponding element $s_{tt'}$ in (15), but again this leads to great convergence problems, so $|s_{tt'}|$ has to be used instead, which gives $w_{itt'} = 1/|s_{tt'}|$. This weight will be denoted $\mathbf{W}_i = (1/|s_{tt'}|)$. The following weights will thus be used:

i. $w_{it} = 1$,

ii. $w_{it} = \frac{1}{u_t}$,

iii. $w_{it} = \frac{1}{u_i}$,

iv. $w_{it} = \frac{1}{\sqrt{s_t^2}}$,

v. $w_{it} = \frac{1}{\sqrt{s_i^2}}$,

vi. $\mathbf{W}_i = \left(\frac{1}{|s_{tt'}|}\right)$, i.e., $w_{itt'} = \frac{1}{|s_{tt'}|}$,

vii. $\mathbf{W}_i = \left|\mathbf{S}^{-1}\right|$.

There are of course many other weights that could be used, but after some preliminary Monte Carlo-simulations different weights that use different combinations of dispersion measures, these seven weights were found to qualify for further examination. For these weights, in general, weights i, iii and v can be used even if the subjects are not measured equally many times or at the same time points. For weights ii, iv, vi and vii all subjects must be measured both at the same time points and equally many times, although weights ii and iv can be modified to handle even these cases.

3 Monte Carlo simulation

This section will present the results from the Monte Carlo simulations where the different weights i-vii that are used with the quantile regression estimators (7) and (11) are compared for the quantiles $\theta = 0.5$, $\theta = 0.75$ and $\theta = 0.9$. It will also present the results from a comparison of the parameter estimates from median regressions using $\theta = 0.5$ in estimators (7) and (11) and mean regressions using a weighted least squares (WLSQ) regression method. The latter is given by

$$\widetilde{\boldsymbol{\beta}}_{\min} = \min_{\boldsymbol{\beta}} \sum_{i=1}^{n} \sum_{t=1}^{T_i} w_{it} \left(y_{it} - f\left(\mathbf{x}_{it}, \boldsymbol{\beta}\right)\right)^2, \tag{23}$$

which in the matrix case corresponds to

$$\widetilde{\boldsymbol{\beta}}_{\min} = \min_{\boldsymbol{\beta}} \sum_{i=1}^{n} \left(\mathbf{y}_i - f\left(\mathbf{X}_i, \boldsymbol{\beta}\right)\right)' \mathbf{W}_i \left(\mathbf{y}_i - f\left(\mathbf{X}_i, \boldsymbol{\beta}\right)\right), \tag{24}$$

where $\mathbf{W}_i$ for the weight specifications i-v are diagonal matrices with the weights w_{it}, $t = 1, ..., T_i$, on the diagonals. However, weight vi, $\mathbf{W}_i = (1/|s_{tt'}|)$, can not be used by the WLSQ regression, since it leads to great convergence problems for this case (however, it does not lead to convergence problems for the median regression). Also, the same problem remains when modifying $\mathbf{W}_i = (1/|s_{tt'}|)$ to $\mathbf{W}_i = (1/s_{tt'})$. Instead, while the median regression uses $\mathbf{W}_i = (1/|s_{tt'}|)$ as weight vi, the WLSQ regression uses $\mathbf{W}_i = \mathbf{S}^{-1}$.

To carry out the Monte Carlo simulations, the values of ε_{it}, n, T and θ must be specified, as well as the correlation structure for the error terms $\varepsilon_{i1}, ..., \varepsilon_{iT}$. In this simulation study, the correlations to be used will be the AR(1)-model

$$\varepsilon_{it} = \rho\varepsilon_{i,t-1} + u_{it}, \quad i = 1, ..., n, \ t = 1, ..., T, \tag{25}$$

where the value of ρ will be set to 0, 0.5 and 0.95, respectively, which will give both observations that are independent and observations with medium and large correlation.

The distributions used for u_{it} when comparing the median regression method with the WLSQ regression method are the standard normal distribution, the uniform distribution on $(-1, 1)$, the Laplace$(0, 1)$ distribution and Student's t distribution with 3 degrees of freedom. The latter two are chosen since they are known to have lower variance for the sample median than for the sample mean. Since these four distributions are symmetric the median and mean regression estimators are estimating the same thing, which they are not for the case of a nonsymmetric distribution. For the comparison of the performance of the different weights i-vii for estimators (7) and (11) using the quantiles $\theta = 0.5$, $\theta = 0.75$, and $\theta = 0.9$ two symmetric and two right skewed distributions are used for the u_{it}'s. These are the standard normal distribution, the uniform distribution on $(-1, 1)$, the standard lognormal distribution and the gamma(α, β) distribution with $\alpha = 2$ and $\beta = 1$. However, for the resulting ε_{it}'s a location shift and scale transformation are made such that the median of ε_{it} is 0 and the variance is 1.

The β-values in (12) also need to be specified. These will be random, so that they are different for different subjects. Let β_{ig}, $i = 1, ..., n$, $g = 1, ..., h$, denote parameter g for subject i. Then, the distributions to be used for (12) are $\beta_{i1} \sim N(10, 1)$, $\beta_{i2} \sim N(20, 1)$, $\beta_{i3} \sim N(5, 0.25)$ and $\beta_{i4} \sim N(-0.5, 0.05)$.

The values for the times x_{it} will be of two different types:

i. Nonrandom with $x_{it} = 10 \times \frac{t}{T}$, so that $0 < x_{it} \leq 10$, and

ii. Random with $x_{it} \sim 10 \times Beta(2, 2)$, i.e., a Beta distribution with $0 < x_{it} < 10$.

Thus, for the logistic model, increasing the number of time points will lead to a more precise measure of the curve in the interval $0 \leq x_{it} \leq 10$, where the main part of the growth occur. The values of x used are the same for all distributions and values of ρ used.

The number of subjects, n, to be used will be 20 and 50, respectively, while the number of time points, T, to be used will be 5 and 10. For $n = 50$ are the values of x and y for the first 20 subjects the same as those used for $n = 20$. In the same manner are the x and y values used for $T = 10$ the same as those used for $T = 5$. There is thus quite a strong dependence between results for different values of n and T.

Finally, the number of Monte Carlo-replications R used is $R = 1,000$ for the comparison of the median regression method with the WLSQ regression method, and $R = 10,000$ for the comparison of the different weights i-vii for estimators (7) and (11) using the quantiles $\theta = 0.5$, $\theta = 0.75$ and $\theta = 0.9$. All estimators are estimated together in the same Monte Carlo-replication.

To analyze the results of the Monte Carlo-simulations, the loss function to be used is the mean absolute percent error (MAPE) of the estimated parameter values from the true parameter values β_g, $g = 1, ..., h$, i.e., the formula

$$MAPE\left(\widehat{\beta}_{\min,gr}\right) = \frac{1}{R}\sum_{r=1}^{R}\frac{\left|\widehat{\beta}_{\min,gr} - \beta_g\right|}{\left|\beta_g\right|} \times 100, \quad g = 1, ..., h, \tag{26}$$

will be used. Also, the bias, in percent, will be examined, estimated by

$$\widehat{bias}\left(\widehat{\beta}_{\min,gr}\right) = \frac{\frac{1}{R}\sum_{r=1}^{R}\widehat{\beta}_{\min,gr} - \beta_g}{\left|\beta_g\right|} \times 100, \quad g = 1, ..., h. \tag{27}$$

The true values of β_1 and β_2 used in (26)-(27) are based on the quantiles 0.5, 0.75 and 0.9 for the distribution of the ε_{it}'s in the AR(1)-model (25). Since these quantiles are known only for the normal distribution and the mean and median of the other symmetric distributions the values of β_1 and β_2 for the quantiles $\theta = 0.5$, $\theta = 0.75$ and $\theta = 0.9$ for the nonsymmetric distributions and for $\theta = 0.75$ and $\theta = 0.9$ for the symmetric distributions other than the normal distribution have to be estimated by simulation. This is done by simulating the distributions of the ε_{it}'s in the AR(1)-model (25) for these cases and calculating the corresponding quantile. For these simulations 5,000,000 replicates are used.

3.1 Comparison of median and mean regression

In this section the results from the Monte Carlo simulations of median regression using $\theta = 0.5$ in the quantile regression estimators (7) and (11), and the mean regression using the weighted least squares regression estimators (23) and (24) are analyzed, mainly by looking at the MAPE values. Overall has weight vi, $\mathbf{W}_i = (1/\left|s_{tt'}\right|)$, i.e., $w_{itt'} = 1/\left|s_{tt'}\right|$, the lowest MAPE values for the median regression case, while weight v, $w_{it} = 1/\sqrt{s_i^2}$, has the lowest MAPE values for the mean regression case. The general comparison of the results for the median and mean regression estimators are made by comparing the MAPE values of the best mean and median regression estimators for each of the four β-parameters. Table 8 in the Appendix give the median of the MAPE over all distributions and values of ρ used, separately for the two different types of time points x_{it} used. As is seen are the mean regression estimators overall performing somewhat better than the median regression estimators when the time points x_{it} are nonrandom, while it is the other way round when the time points are random. For some cases are the mean regression estimators performing very bad when the time points are random, especially for β_1 and β_4 when $T = 5$, with median MAPE values of up to over 200 per cent.

Table 1 give, aggregated over all distributions and values of ρ used, the maximal differences of MAPE between the best performing mean and median regression estimators. The differences are calculated as value of mean regression estimate minus value of median regression estimate and value of median regression estimate minus value of mean regression estimate, respectively. As is seen are the overall maximal differences in favour of the mean regression ranging from 2.03 for β_2 to 6.91 for β_4. This is to be compared with the cases when the median regression estimators outperform the mean regression estimators, in which cases the best of the median regression estimators have overall maximal differences compared with the best mean regression estimators ranging from 6.56 for β_2 to 318.12 for β_4. Thus are the maximal differences much larger in favour of the median regression estimators than they are in favour of the mean regression estimators, so that the mean regression estimators are performing much worse when it performs bad compared to the median regression estimators than the median regression estimators are performing when they perform bad compared to the mean regression estimators.

Table 2 gives the lowest and highest values of MAPE for β_1-β_4 for the best performing mean and median regression estimators. As is shown in the table, even though the mean regression estimators are performing somewhat better than the median regression estimators when they are working well, they are performing much worse when they are working bad. Thus the MAPE values for the median regression are ranging from $\beta_2 = 1.31$ to $\beta_4 = 6.56$ as the lowest and from $\beta_2 = 8.77$ to $\beta_4 = 37.58$ as the highest, which should be compared to the MAPE values for the mean regression, which are ranging from $\beta_2 = 0.87$ to $\beta_4 = 2.88$ and from $\beta_2 = 13.86$ to $\beta_4 = 355.70$, as the highest and lowest. Especially bad are the mean regression estimators performing for β_4.

Table 1: Maximal differences of MAPE for weights with lowest MAPE calculated as value of mean regression estimate minus value of median regression estimate and value of median regression estimate minus values of mean regression estimate, respectively

		Mean vs median				Median vs mean			
Time points	n	β_1	β_2	β_3	β_4	β_1	β_2	β_3	β_4
Nonrandom	20	5.11	1.66	3.87	6.79	1.48	-0.21	-0.33	1.63
	50	**5.45**	1.16	**4.17**	5.47	-0.94	-0.31	-0.78	-1.82
Random	20	1.82	1.83	3.92	5.25	**15.80**	**6.56**	**6.66**	**318.12**
	50	3.12	**2.03**	3.24	**6.91**	13.88	5.69	6.51	112.44

Table 2: Lowest and highest values of MAPE for the mean and median regression estimators with lowest parameter estimates

			Lowest MAPE				Highest MAPE			
estim.	time	n	β_1	β_2	β_3	β_4	β_1	β_2	β_3	β_4
med.	non-	20	4.09	1.88	4.03	6.97	**16.01**	4.36	13.09	24.68
	rand.	50	**3.28**	**1.31**	**2.84**	**6.56**	13.75	3.13	10.02	19.96
	rand.	20	5.83	4.04	6.19	11.78	14.20	**8.77**	**16.45**	**37.58**
		50	5.40	3.05	4.49	10.95	12.88	8.06	13.97	35.08
mean	non-	20	3.12	1.37	1.92	4.52	14.16	3.25	9.90	22.18
	rand.	50	**1.98**	**0.87**	**1.21**	**2.88**	8.30	1.99	5.84	14.50
	rand.	20	5.59	2.75	3.24	8.73	**26.38**	**13.86**	**19.06**	**355.70**
		50	3.02	1.49	1.85	5.26	24.75	12.11	17.52	143.83

Even though the mean regression estimators in general are somewhat better than the median regression estimators, the latter seem to be more robust, so that one is not risking as large deviations from the true β-values when using median regression as one are when using mean regression. This is obvious especially for the case when $T = 5$ and the time points are random. For this case the median regression estimators clearly outperforms the mean regression estimators.

Something should also be said about the estimated biases of the mean and median regression estimators. These are generally quite low, usually below 5 per cent for both estimators. Like in the MAPE case are weights v and vi performing best for the mean and median regression cases, respectively, having the lowest values (in an absolute sense) of the estimated bias. Table 9 in the Appendix give the median of the absolute values of the estimated bias aggregated over all distributions and values of ρ used, separately for the two different types of time points x_{it} used. As is seen are the mean regression estimators overall performing somewhat better for both random and nonrandom time points, although the differences usually are quite small. For β_4, however, are the median regression estimators generally performing better when the time points are random, especially for $T = 5$, where the mean regression estimators perform very bad, with a bias of 188.58 per cent, compared with a bias of only 11.81 per cent for the median regression estimators.

As an example, Figure 2 is comparing the mean and median regression estimators showing the true mean and median regression line (solid) together with the estimated mean and median regression lines (dashed and dotted, respectively) using the average β-values of the overall best mean and median regression estimators, weights v and vi, respectively, for the case $n = 20$, $T = 5$, $\rho = 0.5$ and random time points, with the u_{it}'s

Table 3: Parameter estimates for mean and median regression using weight v for n=20, T=5, ρ =0.5 and random time points with a normal distribution, together with the true values, used for figure 2.

estimator	β_1	β_2	β_3	β_4
mean regression	9.39	20.55	5.03	-1.73
median regression	9.84	20.33	5.06	-0.58
true values	10	20	5	-0.5

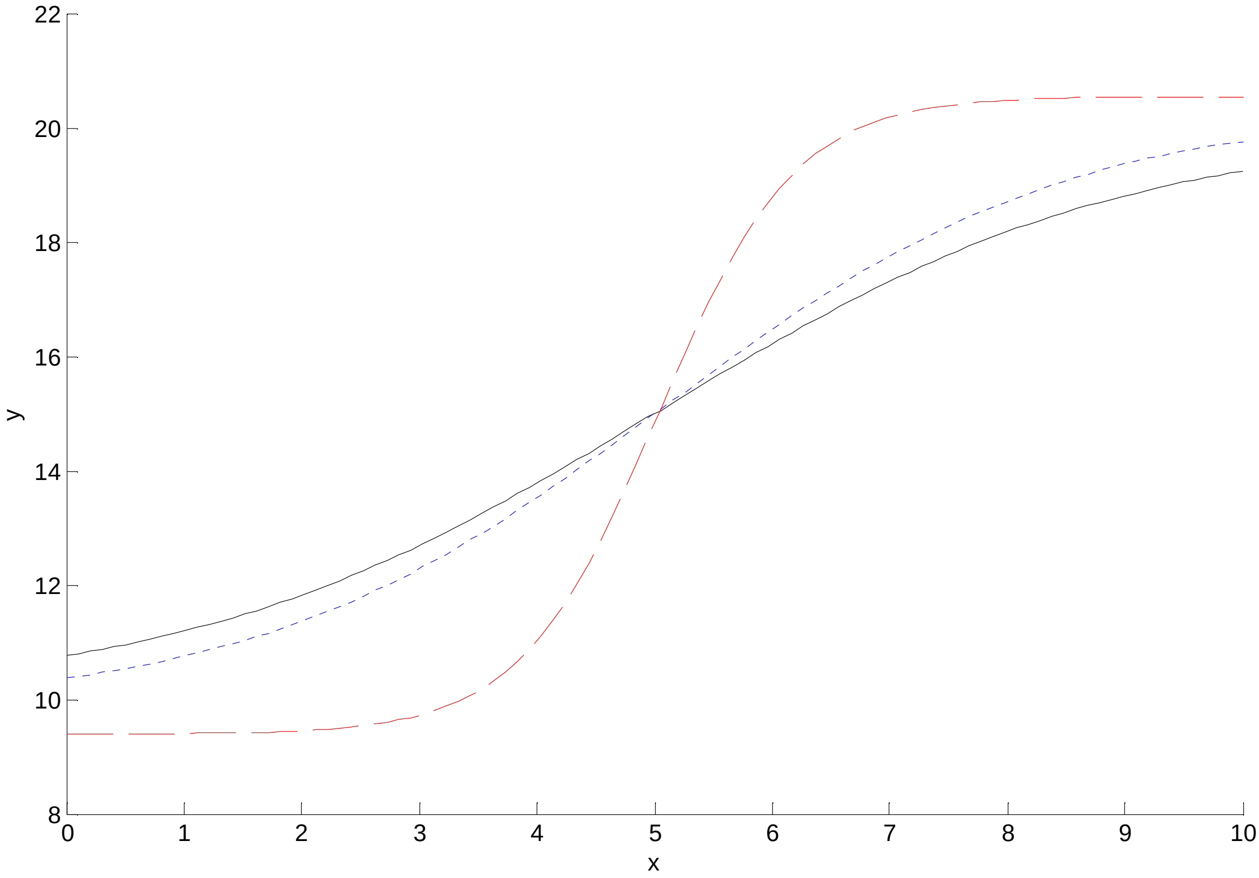

Figure 2: True regression line (solid), estimated mean regression line (dashed) and estimated median regression line (dotted) for the case $n = 20$, $T = 5$, $\rho = 0.5$ and random time points with a normal distribution, using weight v.

in (25) following a normal distribution. Table 3 give the β-values used for the mean and median regressions together with the true β-values. As is obvious from the figure the median regression estimator is performing considerably better than the mean regression estimator, with the latter having considerably larger bias, as can be seen from the large deviation from the true regression line that the mean regression line shows, while the median regression line follows the true regression line quite closely.

3.2 Comparison of quantile regressions with different weights

In this section the results from the Monte Carlo simulations of the quantile regression estimators (7) and (11) for $\theta = 0.5$, $\theta = 0.75$ and $\theta = 0.9$, using weights i-vii, are analyzed, mainly by looking at the MAPE values. Tables 10-12 in Appendix give the average MAPE values aggregated over all distributions used for the u_{it}'s in (25) and all the random and nonrandom time points x_{it}, separately for the different quantiles $\theta = 0.5$, $\theta = 0.75$ and $\theta = 0.9$, respectively. As the tables show there is, overall, small differences between the seven different weights that are used. Usually the average MAPE of the seven weights are within only one or two percentage points from each other, sometimes even within only 0.1 or 0.2 percentage points. The choice of weight to use thus does not seem to be important. However, if the best and worst weights should be singled out, i.e., the weights with lowest and highest average MAPE values, by pooling over all n, T, ρ and θ it is found that the best weight overall is weight v, $w_{it} = 1/\sqrt{s_i^2}$, followed by weight iii, $w_{it} = 1/u_i$. Note that these are the two weights that are based on the dispersion of the e_{it}'s over subjects i. Also weight vi, $\mathbf{W}_i = (1/|s_{tt'}|)$,

Table 4: Average MAPE for quantile regression

weight	Normal				Lognormal			
	β_1	β_2	β_3	β_4	β_1	β_2	β_3	β_4
i	10.32	5.73	10.35	19.37	9.54	5.77	9.99	18.02
ii	10.30	5.77	10.35	19.38	9.30	5.68	9.83	17.60
iii	10.30	5.52	10.21	19.35	7.49	4.27	7.75	14.75
iv	10.26	5.75	10.33	19.28	9.25	5.67	9.81	17.64
v	10.26	**5.48**	**10.16**	19.34	**7.41**	**4.22**	**7.66**	**14.67**
vi	**9.71**	5.60	10.16	**18.62**	8.97	5.54	9.70	17.12
vii	10.26	5.76	10.38	19.49	9.12	5.70	9.83	17.73
	Uniform				Gamma			
i	10.78	5.82	10.68	19.91	11.09	6.29	11.42	21.02
ii	10.67	5.83	10.68	19.85	10.96	6.29	11.38	20.94
iii	11.13	5.79	10.96	20.51	10.22	5.57	10.45	19.84
iv	10.68	5.80	10.66	19.84	10.95	6.27	11.37	20.88
v	11.14	5.78	10.94	20.50	**10.11**	**5.51**	**10.38**	**19.68**
vi	**10.07**	**5.61**	**10.43**	**19.02**	10.31	6.02	11.10	20.05
vii	10.78	5.82	10.74	20.05	10.95	6.27	11.42	21.08

Table 5: Average changes of MAPE from previous θ and ρ

weight	changes from previous θ				changes from previous ρ			
	β_1	β_2	β_3	β_4	β_1	β_2	β_3	β_4
i	0.81	1.18	2.12	2.28	**-3.69**	-1.77	**-3.34**	**-6.69**
ii	0.80	1.20	2.13	2.21	-3.60	-1.76	-3.30	-6.56
iii	**-0.07**	0.39	0.94	0.88	-3.27	-1.44	-2.98	-5.76
iv	0.74	1.19	2.09	2.16	-3.60	-1.75	-3.28	-6.55
v	-0.10	**0.36**	**0.89**	**0.82**	-3.24	-1.43	-2.95	-5.71
vi	0.68	1.15	2.06	2.04	-3.16	-1.60	-3.14	-6.02
vii	0.70	1.18	2.08	2.21	-3.54	**-1.79**	-3.30	-6.57

i.e., $w_{itt'} = 1/|s_{tt'}|$, is performing quite well. The worst weights overall are weight i, $w_{it} = 1$, i.e., with equal weights for all observations and time points, and weight vii, $\mathbf{W}_i = |\mathbf{S}^{-1}|$. These results are the same regardless of the value of ρ used in (25). Looking at the different quantiles θ used it is found that while weights iii and v are performing somewhat better than the other weights for $\theta = 0.75$ and $\theta = 0.90$, while they are performing somewhat worse for $\theta = 0.5$, where weight vi is the best weight overall. But it is worth repeating that the differences between the different weights are small.

Table 4 give the average MAPE value separately for the four distributions used for the u_{it}, aggregated over all values of n, T, ρ, θ and different types of time points. As is seen, there are overall quite small differences between the distributions and weights used, but it is evident that weights iii and v are outperforming the other weights when using nonsymmetric distributions. For the symmetric distributions weight vi shows the best overall results.

There are marked developments of the MAPE when n, T, ρ and θ are changed. These developments are, with a few exceptions, generally the same regardless of which weight is used. The most obvious tendency is that, for all values of θ, increasing the value of ρ in (25) leads to decreasing MAPE. Thus, the higher the correlation is among the error terms ε_{it}, the better is the quantile regression estimation of the parameter vector $\boldsymbol{\beta}$. This may seem surprising, but the results are clear. Comparing the values of MAPE for different values of θ it is seen that the values for $\theta = 0.75$ generally are higher than those for $\theta = 0.5$, and the values for $\theta = 0.9$ in turn generally are higher than those for $\theta = 0.75$. Thus, the further away from the median the quantiles come, the less good is the estimation of $\boldsymbol{\beta}$. The only clear exceptions from this pattern are the values for weights iii and v, which show some cases of lower values of MAPE for higher quantiles, especially for the

Table 6: Quantile regression estimates using weight v for quantiles 0.5, 0.75 and 0.9 when n=50, T=10, ρ =0.95 and nonrandom time points with a normal distribution, together with the true values, used for figure 3.

	θ	β_1	β_2	β_3	β_4
Estimated values	0.50	9.99	20.01	5.00	-0.50
	0.75	10.85	20.85	5.01	-0.50
	0.90	11.60	21.65	5.03	-0.50
True values	0.50	10.00	20.00	5.00	-0.50
	0.75	10.67	20.67	5.00	-0.50
	0.90	11.28	21.28	5.00	-0.50

case of β_1. Table 5 give average values for the changes of MAPE when the values of ρ or θ are increased.

Next question is what happens with the MAPE of the quantile regression estimators when n or T increase. The overall tendency is decreasing with increasing T. However, the tendency to decrease is weakening for larger values of ρ and θ, with for some cases giving increasing MAPE for increasing T, especially for β_1 and β_4. The tendency for the MAPE to decrease with increasing T is stronger for weights iii and v than for the other five weights. For increasing n is there an obvious overall tendency for MAPE to decrease. This tendency is especially obvious for weights iii and v. But for larger values of ρ and θ there are also some cases of increasing MAPE for increasing n. Overall the tendency for MAPE to decrease with increasing n is more marked than the tendency for it to decrease with increasing T.

When comparing the random and nonrandom types of time points x_{it} used it is found that the values of MAPE are in general higher for the random type, but that the differences in general decrease with increasing values of ρ and θ. That the random type of time points have worse estimates than the nonrandom time points should come as no surprise, since they have less observations close to the end points 0 and 10 used for the x_{it}'s than the nonrandom type of time points have. For both types of time points are weights iii and v overall showing the best results, followed by weight vi. For the four different β-parameters it is found that β_2 has, in general, the lowest values of MAPE for both types of time points used, followed by β_3, β_1 and β_4, respectively, when x_{it} is nonrandom, and by β_1, β_3 and β_4, respectively, when x_{it} is random.

It should also be said something about the estimated bias of the quantile regression estimator. Tables 13-15 in the Appendix give the average absolute values of the estimated bias aggregated over all distributions used for the u_{it}'s in (25) and all the random and nonrandom time points x_{it}, separately for quantiles $\theta = 0.5$, $\theta = 0.75$ and $\theta = 0.9$, respectively. There are no materially different conclusions from the corresponding MAPE case, and there are, overall, small differences between the weights used. The biases are usually below 10 per cent for all weights, ρ and θ, and they are decreasing for increasing ρ and increasing for increasing θ. Just like for the MAPE case are weights v and iii overall performing somewhat better than the other weights, with weight vi also performing quite well. The estimated bias is generally positive for β_2 and β_3, while it is negative for β_1 and β_4.

As an example of quantile regression for longitudinal data is Figure 3 showing the true quantile regression lines (solid) for quantiles $\theta = 0.5$, $\theta = 0.75$ and $\theta = 0.9$ for the case $n = 50$, $T = 20$ and $\rho = 0.95$ with the u_{it}'s following a normal distribution, using nonrandom time points, together with the estimated quantile regression lines (dotted) using the average β-values for weight v, which is the weight that overall has the lowest MAPE values, and also overall has the lowest estimated bias both for the normal distribution case and taken over all distributions used. The figure gives a visual estimate of the bias for weight v. Table 6 gives the β-values for weight v, together with the true β-values. As is seen in the figure, the estimated median regression line $\theta = 0.5$ is very close to the true regression line, in fact, they are so close that it is impossible to visually distinguish between them, which shows that any possible bias is negligible, although only β_4 is significantly unbiased on a 5 percent significance level. For the quantiles $\theta = 0.75$ and $\theta = 0.9$ the estimated quantile regression lines are following the true regression lines closely, showing that the possible bias of β_4 is small, but there is a

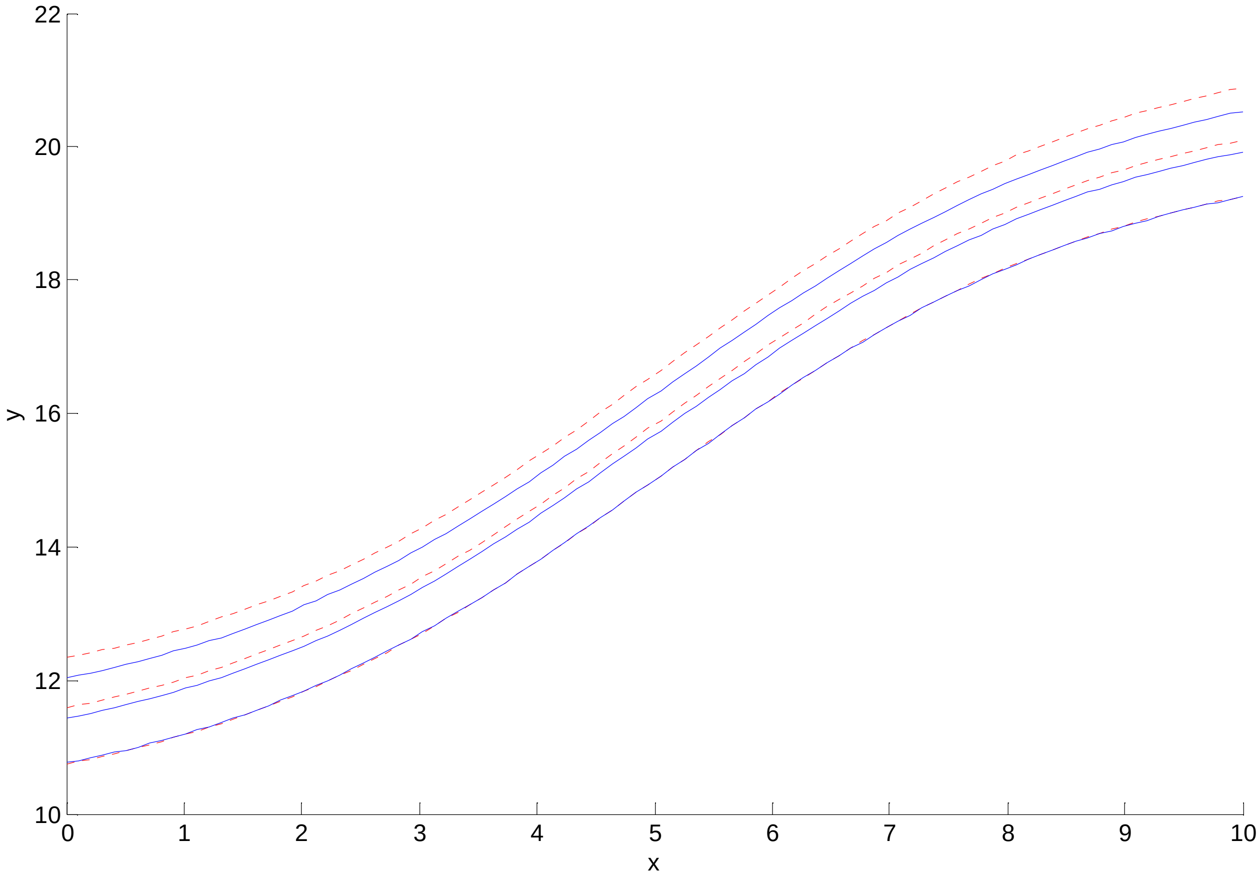

Figure 3: True (solid) and estimated (dotted) quantile regression lines for $\theta = 0.5$, $\theta = 0.75$ and $\theta = 0.9$ using normal distribution with $n = 50$, $T = 10$, $\rho = 0.95$ and nonrandom time points. The mean of the 10,000 parameter estimates for weight v is used for the quantile regression estimator.

positive bias for β_1 and β_2, somewhat larger for $\theta = 0.9$ than for $\theta = 0.75$, which is seen from the fact that the estimated regression lines are placed somewhat above the true regression lines.

4 An application

Davidian and Giltinan (1993) present data from an experiment where the growth patterns of two genotypes of soybean were compared. One was an experimental strain, Plant Introduction No. 416937 (P), and the other was a commercial variety, Forrest (F). Leaf weight was used to assess growth. The experiment was conducted as follows: In each of the three consecutive years 1988-1990 16 plots were planted with seeds, with eight plots used for each genotype, and different plots in different fields used in different years. Each plot was sampled 8-10 times with approximately weekly intervals, in 1988 from day 14 to day 77, in 1989 from day 14 to day 84 and in 1990 from day 15 to day 79. At each sampling time were six plants randomly selected from each plot, and leaves from these plants were weighted. The average leaf weight per plant, in grams, was then calculated for each plot. With the terminology used in this article, there is thus 16 subjects for each year, i.e., a total of 48 subjects, with half of them belonging to each of the two genotypes, and 8-10 observations for each subject.

This data set was analyzed both in Davidian and Giltinan (1993) (the same analyze was later used in Davidian and Giltinan, 1995) and Pinheiro and Bates (2000). Davidian and Giltinan used a three-parameter logistic growth function in a nonlinear mixed-effects model with year as a fixed effect and got six genotype-year

Table 7: Parameter estimates for mean and quantile regression on the experimental strain Plant Introduction No. 416937 (P) and the commercial variety Forrest (F), used in Figures 4 and 5

	Forrest (F)			Plant Introduction (P)		
quantile	β_1	β_2	β_3	β_1	β_2	β_3
0.1	11.21	52.14	-0.1423	17.74	56.38	-0.1185
0.2	14.81	54.48	-0.1264	17.81	54.34	-0.1208
0.3	18.04	56.45	-0.1191	18.44	54.34	-0.1188
0.4	19.23	56.32	-0.1201	18.36	52.95	-0.1226
0.5	19.25	56.10	-0.1184	18.59	52.24	-0.1251
0.6	19.52	55.73	-0.1147	20.33	53.11	-0.1209
0.7	20.55	56.03	-0.1113	22.09	54.25	-0.1142
0.8	20.72	54.48	-0.1158	22.84	53.93	-0.1122
0.9	21.01	53.56	-0.1149	23.51	52.75	-0.1138
mean	16.81	54.14	-0.1241	20.60	54.40	-0.1143

combinations with different parameter estimates, three for each genotype and two for each year. A motivation they used for this was that the weather condition was considerably different for these years, with 1988 being dry, 1989 being wet and 1990 being normal

The soybean data is here analyzed using the marginal analysis approach of Model (2),

$$y_{it} = f(\mathbf{x}_{it}, \boldsymbol{\beta}) + \varepsilon_{it}, \quad i = 1, ..., 24, \ t = 1, ..., T_i, \tag{28}$$

with T_i being 8, 9 or 10 and the function

$$f(\mathbf{x}_{it}, \boldsymbol{\beta}) = \frac{\beta_1}{1 + \exp(\beta_3 (x_{it} - \beta_2))}, \quad \beta_1, \beta_2 > 0, \ \beta_3 < 0, \tag{29}$$

estimated for all three years together, but separately for each of the two soybean types P and F. Here are β_1 giving the limiting growth value of the soybean plants, β_2 gives the EC_{50} value, i.e., the day at which half the limiting growth value is achieved, while β_3 is a growth rate constant governing the steepness of the growth curve. As weight for the mean and quantile regression estimators is weight v used, since this weight is performing best for both cases. Not that, for example, weight vi, $\mathbf{W}_i = (1/|s_{tt'}|)$, not can be used here, since all subjects are not measured equally many times.

The parameter estimates are presented in Table 7, while Figure 4 and 5 shows the mean (solid) and quantile (dotted, with median dashed) regression lines, using quantiles $\theta = 0.1, \ 0.2, \ \ldots \ 0.9$, for soybean types F and P, respectively, together with the original data points. Inspecting the table and figures give some interesting informations. Davidian and Giltinan (1993, 1995), using mean regression, found that for all three years had the experimental strain P higher limiting average leaf weight than the commercial variety F, which suggests that the experimental strain P is to be preferred to the commercial variety F. This agrees with the results found here for the mean regression case, were the limiting leaf weight (β_1) for P is clearly higher than for F, with 20.60 versus 16.81. But looking at the quantile regressions give a somewhat different picture. It is thus found that the limiting median leaf weight for the commercial variety F is in fact somewhat higher than for the experimental strain P, 19.25 versus 18.59, which thus instead suggests that the commercial variety F should be preferred to the experimental strain P. The quantile regression lines tells us that the commercial variety F is heavily skewed downwards, while the experimental strain P is instead skewed upwards, explaining why the mean and median regressions give such different pictures. It is also seen that the dispersion of the commercial variety F, measured as the distance between the quantile regression lines for quantiles 0.1 and 0.9, are considerably larger than for the experimental strain P, suggesting that it has larger variance. Even though the commercial variety F has somewhat larger limiting median leaf weight than the experimental strain

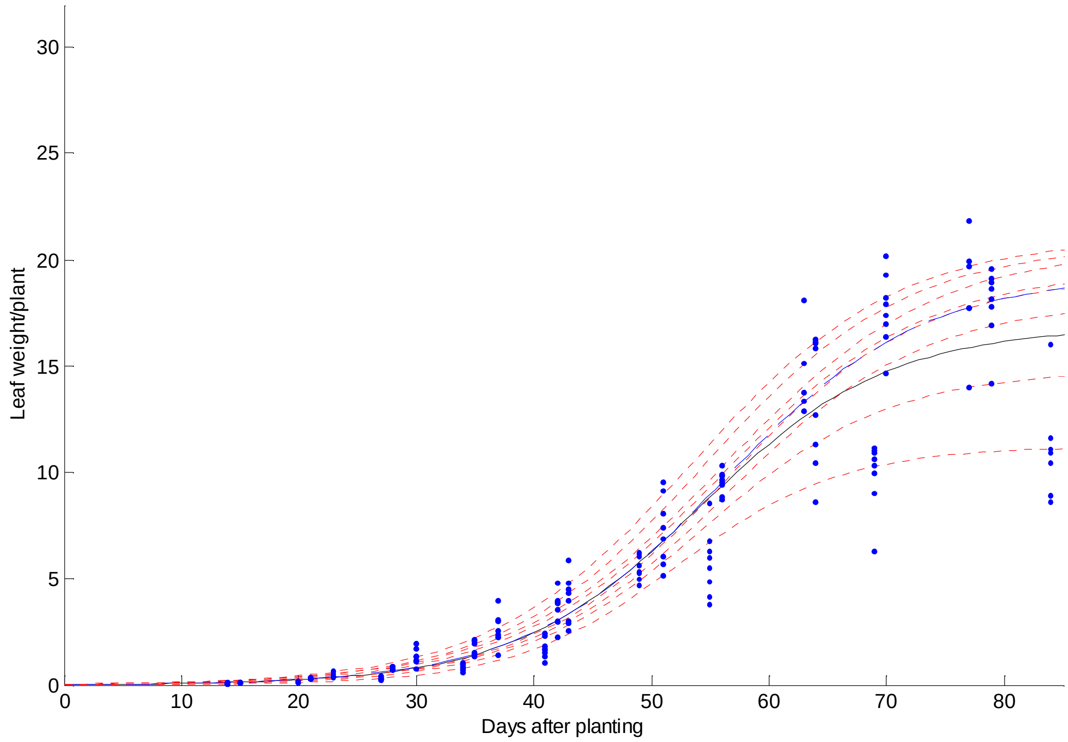

Figure 4: Mean and quantile regression lines for soybean Forrest (F) estimated using weight v. Mean regression line is solid, quantile regression lines ($\theta = 0.1,\ 0.2,\ \ldots\ 0.9$) are dotted, with median regression line dashed.

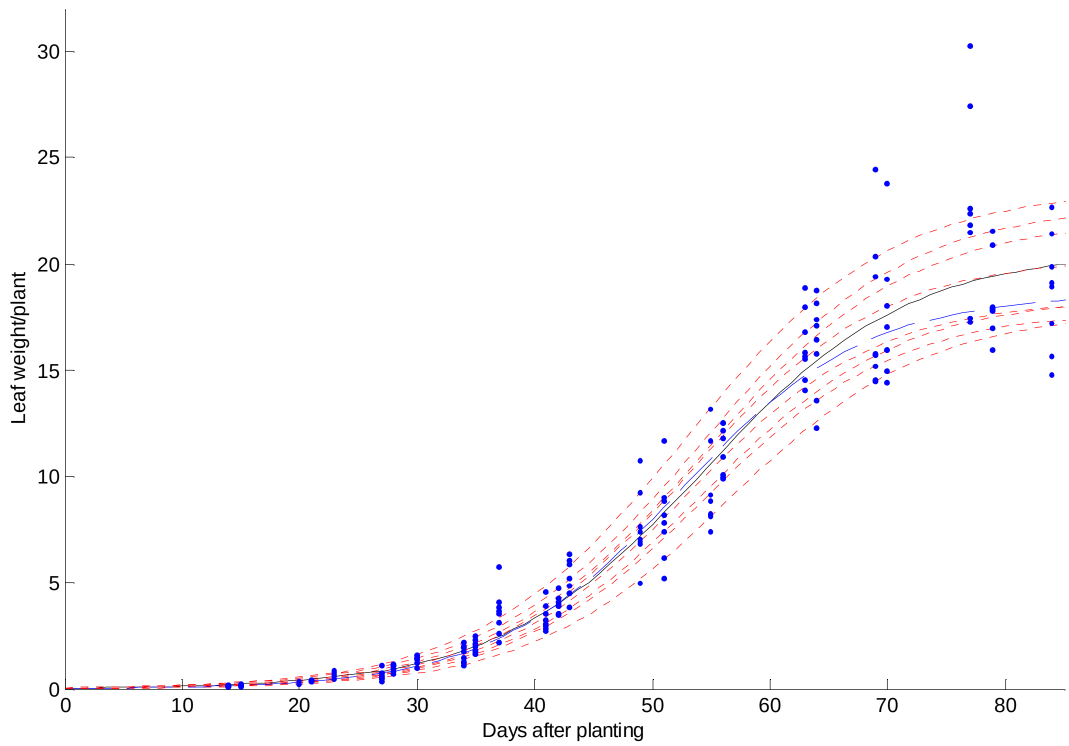

Figure 5: Mean and quantile regression lines for soybean Plant Introduction No. 416937 (P), estimated using weight v. Mean regression line is solid, quantile regression lines ($\theta = 0.1,\ 0.2,\ \ldots\ 0.9$) are dotted, with median regression line dashed.

P has the latter larger limiting values for quantiles 0.1, 0.2, 0.6, 0.7, 0.8 and 0.9, suggesting that both the soybean plants from the experimental strain P with the lowest limiting leaf weight and those with the highest limiting leaf weight has a higher weight than the soybean plants with lowest and highest limiting leaf weight for the commercial variety F. So there are many different aspect that need to be taken into consideration when deciding which soybean type is to be preferred for having the highest limiting leaf weight, and only using the mean regression could lead to erroneous conclusions. The quantile regression approach gives a more complete picture of the data material for the researcher to base his inference on than the mean regression gives, and is thus to be preferred.

Finally, comparing the mean, median and quantile regression lines also give some interesting examples of how the conditional distributions of the soybean plant are changing shape over time. For the experimental strain P is the conditional distribution somewhat skewed upwards until around day 40, then becomes slightly skewed downwards until around day 60, while it thereafter becomes more and more skewed upwards. For the commercial variety F the conditional distribution seem to be practically symmetric until around day 50, and after that it becomes more and more heavily skewed downwards. This implies that the underlying conditional distributions possibly could be normally distributed until around day 50 or 60, and then change to nonnormal distributions.

5 Summary and conclusions

In this paper the use of a quantile regression estimator for nonlinear longitudinal data has been examined. The estimator used is a weighted version of the quantile regression estimator defined by Koenker and Bassett (1978), adjusted to the case of nonlinear longitudinal data with a marginal model approach. A four-parameter logistic growth function is used, with error terms following an AR(1) model. Seven different weights, based on the estimated regression errors, are used and compared by Monte Carlo simulation for the quantiles 0.5, 0.75 and 0.9, mainly by looking at the MAPE. The comparison is made for different degrees of autocorrelation, distributions, number of observations, number of subjects and types of time points. The differences between the seven weights used are small, but the weights based on the dispersion of the estimated regression errors over subjects is overall performing somewhat better than the other weights.

It is found that the quantile regression estimator is performing quite well in terms of bias, especially for the median regression case, but the farther away from the median the quantiles are, the less well is it working. The latter is true also for the MAPE. When the autocorrelation ρ in the AR(1) model is increased all weights are performing better, which is somewhat surprising, but these results are clear. Increasing the number of subjects or observations is overall leading to better estimates. Overall are there quite small differences for the performances between the different distributions used, but the estimator is generally performing better for nonrandom than for random time points. A comparison is also made between the median regression case of the quantile regression estimator and the corresponding mean regression estimator, where the latter is found to be less robust. Finally the quantile regression estimator is, together with a mean regression estimator, applied to a real data set where the growth patterns of two genotypes of soybean are compared, which give some interesting insights into how the quantile regressions give a more complete picture of the data than the mean regression does.

6 Appendix

6.1 Tables

Table 8: Median of lowest MAPE values for each parameter and weight, for median and mean regressions

			Nonrandom time points				Random time points			
T	n	estimator	β_1	β_2	β_3	β_4	β_1	β_2	β_3	β_4
5	20	median	**9.89**	**3.00**	**8.53**	**15.98**	10.09	6.72	12.48	25.05
		mean	9.01	2.41	6.82	15.41	**23.54**	**12.15**	**16.86**	**202.69**
	50	median	7.38	2.04	5.74	13.00	10.27	6.18	10.43	23.88
		mean	5.02	1.42	3.85	9.61	19.09	9.69	13.73	97.37
10	20	median	7.64	2.90	6.04	15.76	10.63	6.02	9.34	22.56
		mean	5.91	2.22	4.51	13.05	11.99	6.05	8.66	23.66
	50	median	4.60	1.74	3.56	10.52	9.02	4.77	6.87	18.19
		mean	3.54	1.35	2.72	8.20	7.71	3.99	5.47	16.10

Table 9: Median of lowest absolute value of estimated bias for each parameter and weight, for median and mean regressions

			Nonrandom time points				Random time points			
T	n	estimator	β_1	β_2	β_3	β_4	β_1	β_2	β_3	β_4
5	20	median	**3.24**	**0.62**	1.48	**1.67**	1.28	1.28	**0.90**	11.81
		mean	2.71	0.40	**1.58**	0.27	1.05	0.79	0.05	**188.58**
	50	median	1.86	0.31	1.12	0.38	1.88	1.35	0.51	8.62
		mean	0.59	0.11	0.47	0.08	1.80	1.21	0.04	81.02
10	20	median	2.00	0.59	0.72	0.85	2.84	1.77	0.43	5.30
		mean	0.90	0.24	0.43	0.49	**2.86**	1.48	0.07	7.41
	50	median	0.50	0.14	0.25	0.22	2.75	**1.54**	0.06	1.75
		mean	0.21	0.03	0.11	0.24	1.55	0.96	0.02	1.96

Table 10: Average MAPE for quantile regression with θ=0.5

			T=5				T=10			
ρ	n	weight	β_1	β_2	β_3	β_4	β_1	β_2	β_3	β_4
0	20	i	16.44	7.44	15.40	27.08	12.19	6.11	10.01	21.64
		ii	15.94	7.38	15.17	26.72	12.12	6.07	10.02	21.70
		iii	16.69	7.72	16.04	27.44	12.66	6.35	10.58	22.44
		iv	16.20	7.34	15.20	26.74	12.13	6.09	10.00	21.72
		v	16.52	7.65	15.92	27.43	12.76	6.31	10.58	22.47
		vi	14.33	6.88	14.25	24.50	11.65	5.84	10.14	21.54
		vii	15.97	7.35	15.11	26.56	12.09	6.12	10.16	22.11
	50	i	13.60	6.06	11.74	23.78	8.66	4.20	6.49	16.03
		ii	13.35	6.08	11.68	23.28	8.61	4.14	6.50	16.12
		iii	14.05	6.26	12.57	23.80	9.09	4.40	6.92	16.81
		iv	13.48	6.10	11.74	23.31	8.65	4.17	6.51	16.10
		v	13.97	6.18	12.52	24.12	9.04	4.39	6.90	16.78
		vi	12.42	5.63	11.02	21.68	8.55	4.19	6.66	16.10
		vii	13.61	6.08	11.76	23.31	8.67	4.19	6.57	16.28
0.5	20	i	12.81	6.66	12.98	22.72	11.63	5.79	9.88	21.17
		ii	12.29	6.37	12.68	22.08	11.59	5.80	9.86	21.18
		iii	12.93	6.80	13.71	23.07	11.94	5.94	10.26	21.74
		iv	12.47	6.45	12.81	22.13	11.50	5.74	9.78	21.10
		v	12.94	6.70	13.66	22.91	11.92	5.95	10.27	21.61
		vi	11.41	6.06	12.22	20.90	11.09	5.69	9.99	20.64
		vii	12.49	6.40	12.90	22.29	11.48	5.79	9.96	21.44
	50	i	10.87	5.28	9.99	19.72	8.26	4.00	6.39	15.76
		ii	10.85	5.34	9.93	19.75	8.25	3.99	6.38	15.74
		iii	10.96	5.36	10.45	19.97	8.60	4.13	6.74	16.43
		iv	10.78	5.31	9.95	19.75	8.22	3.98	6.37	15.73
		v	10.93	5.40	10.45	20.12	8.48	4.09	6.65	16.40
		vi	10.32	5.21	9.70	18.94	8.12	3.93	6.42	15.64
		vii	10.82	5.29	9.95	19.78	8.29	4.01	6.42	15.90
0.95	20	i	5.72	3.54	6.62	10.60	6.43	3.47	5.94	12.85
		ii	5.73	3.52	6.58	10.49	6.44	3.46	5.94	12.78
		iii	6.57	4.04	7.30	11.91	6.49	3.54	5.86	12.60
		iv	5.74	3.49	6.56	10.49	6.45	3.47	5.95	12.83
		v	6.51	4.05	7.25	11.85	6.48	3.53	5.85	12.62
		vi	5.64	3.49	6.51	10.33	6.30	3.44	5.89	12.64
		vii	5.75	3.45	6.50	10.44	6.46	3.48	6.06	12.92
	50	i	5.37	2.95	5.37	10.34	4.84	2.41	3.99	9.91
		ii	5.25	2.90	5.37	10.25	4.84	2.41	3.99	9.91
		iii	5.48	3.13	5.56	10.51	4.89	2.52	4.02	9.93
		iv	5.31	2.91	5.40	10.34	4.82	2.42	4.00	9.92
		v	5.43	3.11	5.56	10.45	4.91	2.51	4.01	9.93
		vi	5.26	2.90	5.39	10.24	4.80	2.41	3.98	9.89
		vii	5.34	2.93	5.42	10.42	4.85	2.43	4.04	10.02

Table 11: Average MAPE for quantile regression with θ=0.75

			T=5				T=10			
ρ	n	weight	β_1	β_2	β_3	β_4	β_1	β_2	β_3	β_4
0	20	i	16.74	8.62	17.23	29.52	12.89	7.30	11.94	23.80
		ii	16.00	8.52	16.88	28.64	12.72	7.20	11.89	23.80
		iii	14.97	7.49	15.78	27.14	12.26	6.57	11.19	22.82
		iv	16.09	8.40	16.84	29.24	12.77	7.16	11.87	23.79
		v	14.93	7.57	15.76	26.82	12.21	6.55	11.19	22.77
		vi	14.70	7.94	15.94	26.46	12.12	6.98	12.11	23.53
		vii	16.02	8.45	16.75	28.77	12.63	7.24	12.05	24.17
	50	i	13.92	7.17	13.54	24.43	9.41	5.20	7.87	17.86
		ii	14.02	7.25	13.56	24.47	9.37	5.21	7.88	17.90
		iii	12.64	6.06	12.12	22.73	8.91	4.82	7.45	17.41
		iv	13.96	7.21	13.49	24.55	9.43	5.23	7.90	17.87
		v	12.43	6.03	12.01	22.55	8.79	4.77	7.40	17.27
		vi	12.69	6.79	12.81	22.87	9.18	5.17	8.01	17.82
		vii	13.84	7.45	13.47	24.70	9.44	5.25	8.01	18.03
0.5	20	i	12.83	7.33	14.51	23.67	12.50	6.94	11.81	23.45
		ii	12.64	7.48	14.41	23.93	12.35	6.95	11.82	23.61
		iii	12.02	6.54	13.66	22.41	11.42	6.17	10.84	22.24
		iv	12.52	7.31	14.35	23.37	12.27	6.91	11.76	23.51
		v	12.08	6.53	13.67	22.71	11.40	6.09	10.80	22.16
		vi	11.59	7.02	13.76	21.99	11.80	6.74	11.93	22.84
		vii	12.76	7.38	14.52	23.94	12.37	6.99	12.03	23.93
	50	i	11.14	6.30	11.47	20.87	8.90	4.86	7.69	17.41
		ii	11.15	6.28	11.47	21.13	8.90	4.94	7.74	17.43
		iii	9.97	5.22	10.35	19.36	8.24	4.29	7.11	16.67
		iv	11.12	6.27	11.50	20.96	8.95	4.92	7.72	17.45
		v	10.01	5.25	10.33	19.24	8.27	4.25	7.09	16.65
		vi	10.53	6.03	11.10	19.95	8.74	4.85	7.76	17.20
		vii	11.30	6.32	11.54	21.45	8.97	4.89	7.75	17.61
0.95	20	i	5.85	3.92	7.14	10.35	6.83	4.06	7.02	13.67
		ii	5.89	3.92	7.12	10.27	6.77	4.06	7.00	13.71
		iii	6.41	4.13	7.43	11.97	6.32	3.64	6.10	12.73
		iv	5.86	3.91	7.13	10.30	6.76	4.04	6.99	13.66
		v	6.35	4.11	7.43	11.86	6.29	3.61	6.10	12.70
		vi	5.79	3.89	7.06	10.18	6.74	4.03	6.99	13.59
		vii	5.81	3.85	7.02	10.27	6.78	4.03	7.10	13.82
	50	i	5.95	3.61	6.44	11.15	5.48	3.03	4.87	10.80
		ii	5.97	3.61	6.42	11.21	5.45	2.99	4.84	10.80
		iii	5.36	3.15	5.65	10.19	4.90	2.59	4.19	9.99
		iv	5.94	3.61	6.45	11.19	5.49	3.02	4.87	10.81
		v	5.29	3.15	5.59	10.14	4.89	2.60	4.19	9.99
		vi	5.93	3.61	6.44	11.06	5.49	3.01	4.84	10.78
		vii	6.10	3.66	6.50	11.58	5.53	3.01	4.92	10.95

Table 12: Average MAPE for quantile regression with θ=0.9

			T=5				T=10			
ρ	n	weight	β_1	β_2	β_3	β_4	β_1	β_2	β_3	β_4
0	20	i	16.91	10.27	20.07	34.97	14.72	9.49	16.50	29.24
		ii	16.46	10.21	19.89	33.58	14.44	9.39	16.35	29.08
		iii	14.19	7.92	17.15	28.22	12.98	7.76	13.96	25.94
		iv	16.15	10.20	19.83	33.12	14.34	9.37	16.22	28.82
		v	14.30	7.91	17.09	27.96	12.70	7.58	13.66	25.57
		vi	14.52	9.49	18.76	30.75	13.46	9.06	16.56	28.60
		vii	15.66	10.21	19.52	32.55	14.12	9.47	16.53	29.51
	50	i	14.89	8.61	16.67	28.18	11.37	7.16	11.32	22.36
		ii	15.05	8.72	16.66	27.95	11.27	7.17	11.31	22.40
		iii	12.26	6.40	13.28	23.59	9.70	5.68	9.35	20.09
		iv	14.79	8.75	16.54	27.74	11.20	7.07	11.18	22.35
		v	12.08	6.33	13.16	23.42	9.62	5.61	9.22	19.82
		vi	13.53	8.16	15.74	25.58	11.02	7.04	11.52	21.94
		vii	14.77	8.71	16.57	28.15	11.19	7.14	11.34	22.64
0.5	20	i	13.60	8.80	17.43	28.05	14.48	9.15	16.51	29.01
		ii	13.33	8.73	17.31	26.80	14.20	9.16	16.42	28.79
		iii	12.16	7.19	15.43	24.30	12.31	7.35	13.54	25.14
		iv	13.26	8.76	17.25	26.58	13.99	9.02	16.32	28.91
		v	12.18	7.15	15.29	24.40	12.09	7.20	13.35	24.97
		vi	12.16	8.35	16.38	24.98	13.07	8.78	16.32	27.71
		vii	13.19	8.76	17.29	27.43	13.84	9.07	16.53	29.27
	50	i	12.48	7.63	14.51	23.58	11.18	6.92	11.32	22.23
		ii	12.12	7.75	14.58	23.61	11.03	6.88	11.26	22.21
		iii	10.33	5.75	11.91	20.57	9.16	5.31	9.06	19.47
		iv	12.14	7.71	14.55	23.55	11.06	6.89	11.27	22.27
		v	10.09	5.66	11.73	20.60	9.04	5.17	8.93	19.23
		vi	11.50	7.57	14.04	22.52	10.64	6.74	11.22	21.68
		vii	12.36	7.72	14.66	23.73	11.02	6.87	11.32	22.45
0.95	20	i	5.90	4.35	7.55	9.36	7.66	5.19	9.19	15.08
		ii	5.87	4.35	7.53	9.36	7.57	5.20	9.20	15.09
		iii	6.90	4.68	8.76	12.93	7.02	4.36	7.74	14.40
		iv	5.87	4.34	7.51	9.29	7.59	5.18	9.18	15.04
		v	6.85	4.66	8.72	13.01	6.99	4.33	7.71	14.42
		vi	5.80	4.30	7.46	9.21	7.53	5.17	9.14	14.91
		vii	5.81	4.25	7.33	9.24	7.51	5.11	9.19	14.96
	50	i	6.30	4.42	7.64	11.06	6.82	4.23	6.93	13.15
		ii	6.33	4.44	7.64	11.02	6.87	4.24	6.96	13.22
		iii	5.89	3.77	6.84	11.30	5.70	3.27	5.43	11.72
		iv	6.25	4.42	7.60	11.04	6.82	4.25	6.95	13.20
		v	5.91	3.73	6.83	11.19	5.64	3.23	5.38	11.68
		vi	6.25	4.41	7.59	10.94	6.78	4.23	6.93	13.11
		vii	6.32	4.39	7.56	11.13	6.84	4.22	7.01	13.32

Table 13: Average absolute value of estimated bias for quantile regression with θ=0.5

			T=5				T=10			
ρ	n	weight	β_1	β_2	β_3	β_4	β_1	β_2	β_3	β_4
0	20	i	7.18	3.71	3.44	7.14	4.78	3.18	1.30	2.08
		ii	6.75	3.69	3.45	7.07	4.71	3.12	1.26	2.15
		iii	7.41	3.62	3.47	7.54	5.11	3.19	1.43	2.39
		iv	7.04	3.67	3.42	6.98	4.70	3.13	1.28	2.24
		v	7.21	3.53	3.51	7.53	5.20	3.13	1.29	2.44
		vi	5.50	3.26	2.66	6.38	4.19	2.83	1.12	2.80
		vii	6.83	3.67	3.27	6.94	4.56	3.12	1.48	2.56
	50	i	5.51	3.03	2.01	5.17	2.96	2.05	0.57	0.72
		ii	5.28	3.06	2.12	4.74	2.91	1.99	0.53	0.83
		iii	5.93	2.91	2.71	4.86	3.19	2.06	0.57	0.77
		iv	5.40	3.08	2.07	4.71	2.95	2.02	0.55	0.77
		v	5.80	2.82	2.69	5.20	3.14	2.05	0.61	0.75
		vi	4.69	2.71	1.78	4.30	2.82	2.03	0.57	0.76
		vii	5.49	3.06	2.06	4.63	2.92	2.03	0.62	0.87
0.5	20	i	4.64	3.10	2.20	5.77	4.31	2.78	1.13	2.14
		ii	4.21	2.84	2.14	5.35	4.33	2.81	1.12	2.10
		iii	4.74	2.70	2.35	5.88	4.70	2.58	1.24	2.67
		iv	4.40	2.93	2.21	5.29	4.19	2.74	1.12	2.19
		v	4.85	2.60	2.30	5.74	4.70	2.59	1.23	2.64
		vi	3.65	2.59	1.82	5.24	3.94	2.68	1.10	2.22
		vii	4.39	2.87	2.27	5.41	4.13	2.71	1.23	2.57
	50	i	3.77	2.41	1.26	3.40	2.61	1.79	0.47	0.74
		ii	3.76	2.48	1.28	3.49	2.60	1.78	0.46	0.74
		iii	3.93	2.00	1.57	3.80	2.98	1.58	0.55	1.04
		iv	3.69	2.45	1.32	3.46	2.58	1.77	0.45	0.72
		v	3.95	2.03	1.69	3.97	2.91	1.54	0.58	1.11
		vi	3.41	2.39	1.25	3.30	2.50	1.72	0.40	0.69
		vii	3.66	2.42	1.35	3.48	2.60	1.76	0.50	0.94
0.95	20	i	0.71	0.80	0.70	1.85	1.18	0.86	0.43	1.00
		ii	0.75	0.79	0.65	1.77	1.21	0.86	0.43	1.00
		iii	0.96	0.78	0.77	1.85	1.24	0.64	0.35	1.11
		iv	0.67	0.76	0.66	1.81	1.22	0.87	0.44	0.97
		v	0.95	0.80	0.77	1.77	1.25	0.65	0.36	1.14
		vi	0.70	0.78	0.69	1.80	1.10	0.83	0.46	1.05
		vii	0.72	0.74	0.63	1.68	1.17	0.82	0.44	1.29
	50	i	0.83	0.76	0.46	1.31	0.80	0.54	0.18	0.35
		ii	0.75	0.71	0.50	1.28	0.79	0.53	0.18	0.39
		iii	0.86	0.60	0.52	1.42	0.88	0.41	0.18	0.47
		iv	0.76	0.71	0.50	1.33	0.78	0.54	0.20	0.37
		v	0.83	0.60	0.55	1.46	0.91	0.44	0.17	0.50
		vi	0.76	0.71	0.49	1.29	0.76	0.54	0.21	0.35
		vii	0.80	0.75	0.48	1.26	0.76	0.52	0.22	0.52

Table 14: Average absolute value of estimated bias for quantile regression with θ=0.75

			T=5				T=10			
ρ	n	weight	β_1	β_2	β_3	β_4	β_1	β_2	β_3	β_4
0	20	i	5.39	5.67	3.52	8.52	3.53	5.13	2.67	2.38
		ii	4.81	5.61	3.62	7.60	3.58	5.00	2.86	2.43
		iii	5.48	3.86	3.04	7.66	3.90	4.09	1.88	2.25
		iv	4.80	5.50	3.38	8.39	3.52	4.98	2.75	2.47
		v	5.49	3.95	2.92	7.37	3.85	4.08	1.94	2.33
		vi	4.03	5.11	2.85	7.00	2.98	4.75	2.61	2.85
		vii	4.88	5.55	3.41	8.00	3.43	4.99	3.01	2.84
	50	i	4.04	4.92	3.11	4.45	2.87	3.76	2.46	0.98
		ii	4.20	5.00	3.17	4.34	2.86	3.76	2.63	0.99
		iii	4.42	3.33	2.39	4.44	2.75	3.11	1.69	0.90
		iv	4.13	4.95	3.18	4.58	2.89	3.79	2.59	0.99
		v	4.24	3.32	2.41	4.38	2.71	3.08	1.74	0.87
		vi	3.35	4.63	2.86	4.15	2.59	3.71	2.51	0.92
		vii	3.90	5.19	3.14	4.58	2.89	3.79	2.65	1.04
0.5	20	i	2.83	4.66	2.53	5.49	2.87	4.81	2.38	2.60
		ii	2.66	4.85	2.63	5.71	2.76	4.77	2.71	2.78
		iii	3.56	3.05	2.44	5.33	3.17	3.56	1.95	2.67
		iv	2.56	4.69	2.65	5.33	2.70	4.75	2.70	2.78
		v	3.60	3.04	2.37	5.68	3.16	3.49	1.93	2.75
		vi	2.09	4.47	2.54	4.87	2.45	4.60	2.62	2.63
		vii	2.80	4.78	2.81	5.60	2.68	4.77	2.78	3.20
	50	i	2.05	4.30	2.20	3.44	2.35	3.51	2.42	1.07
		ii	2.04	4.30	2.35	3.54	2.42	3.58	2.64	1.10
		iii	2.56	2.59	1.85	3.66	1.99	2.55	1.62	1.23
		iv	2.02	4.30	2.31	3.46	2.44	3.56	2.55	1.16
		v	2.59	2.63	1.91	3.64	2.05	2.51	1.56	1.32
		vi	1.83	4.09	2.22	3.28	2.24	3.51	2.48	0.93
		vii	2.18	4.32	2.32	3.75	2.34	3.51	2.57	1.31
0.95	20	i	1.47	2.25	1.96	0.86	1.29	2.49	2.25	0.86
		ii	1.40	2.27	2.00	0.71	1.35	2.49	2.36	0.88
		iii	0.70	1.44	1.14	1.43	0.61	1.37	1.04	1.01
		iv	1.45	2.25	1.97	0.80	1.39	2.47	2.32	0.87
		v	0.72	1.43	1.23	1.46	0.66	1.36	1.06	1.12
		vi	1.48	2.22	1.94	0.81	1.35	2.47	2.31	0.83
		vii	1.47	2.21	1.93	0.76	1.43	2.43	2.29	1.10
	50	i	1.50	2.36	2.12	0.50	1.96	2.12	2.15	0.45
		ii	1.49	2.36	2.11	0.55	1.99	2.08	2.13	0.45
		iii	0.85	1.37	0.93	0.87	0.94	1.19	0.88	0.46
		iv	1.50	2.36	2.11	0.53	1.94	2.11	2.13	0.48
		v	0.83	1.39	0.99	0.83	0.97	1.20	0.94	0.54
		vi	1.48	2.36	2.12	0.45	1.93	2.10	2.12	0.44
		vii	1.44	2.38	2.11	0.63	2.03	2.07	2.15	0.63

Table 15: Average absolute value of estimated bias for quantile regression with θ=0.9

			T=5				T=10			
ρ	n	weight	β_1	β_2	β_3	β_4	β_1	β_2	β_3	β_4
0	20	i	4.09	7.52	3.76	14.08	2.50	7.44	4.75	5.19
		ii	3.67	7.52	4.45	12.32	2.42	7.33	5.34	4.99
		iii	4.35	4.36	2.60	8.78	2.85	5.39	3.04	3.92
		iv	3.35	7.53	4.56	12.08	2.33	7.31	5.41	4.86
		v	4.54	4.38	2.65	8.48	2.71	5.22	2.93	3.72
		vi	2.17	6.87	4.36	10.91	1.51	6.96	5.28	5.61
		vii	2.95	7.57	4.84	11.60	2.06	7.36	5.96	5.65
	50	i	2.67	6.60	3.51	7.11	1.96	5.92	4.52	2.18
		ii	2.84	6.76	3.74	6.64	1.96	5.92	4.78	2.15
		iii	3.23	3.85	2.60	5.01	1.92	4.16	3.05	1.76
		iv	2.57	6.77	3.71	6.49	1.91	5.82	4.71	2.26
		v	3.10	3.80	2.67	4.99	1.89	4.10	2.97	1.68
		vi	1.88	6.27	3.45	5.42	1.70	5.80	4.53	1.77
		vii	2.60	6.75	3.84	6.85	1.83	5.87	4.99	2.34
0.5	20	i	1.81	6.33	3.99	9.47	2.18	7.09	4.54	5.35
		ii	1.58	6.32	4.37	8.07	1.99	7.10	5.29	5.20
		iii	3.07	3.69	2.57	6.62	2.72	4.81	2.93	3.74
		iv	1.63	6.34	4.42	7.87	1.74	6.94	5.20	5.39
		v	3.09	3.63	2.45	6.87	2.50	4.65	2.84	3.96
		vi	1.47	5.98	4.35	7.25	1.17	6.72	5.23	5.11
		vii	1.46	6.32	4.18	8.73	1.55	6.92	5.41	6.15
	50	i	1.83	5.86	3.55	4.75	1.61	5.70	4.33	2.17
		ii	1.43	6.01	3.80	4.55	1.55	5.65	4.55	2.24
		iii	2.00	3.26	2.35	3.83	1.51	3.67	2.90	1.87
		iv	1.48	5.96	3.81	4.58	1.64	5.66	4.54	2.24
		v	1.86	3.15	2.36	4.26	1.58	3.54	2.78	1.97
		vi	1.68	5.87	3.79	4.14	1.49	5.52	4.43	1.88
		vii	1.48	5.98	3.66	4.49	1.61	5.60	4.63	2.66
0.95	20	i	2.49	3.10	2.24	0.70	2.57	3.91	3.92	0.90
		ii	2.54	3.09	2.29	0.78	2.66	3.93	4.10	0.91
		iii	1.06	2.26	1.57	1.36	1.08	2.34	1.81	1.20
		iv	2.54	3.08	2.30	0.72	2.59	3.92	4.02	0.82
		v	1.09	2.24	1.61	1.52	1.13	2.33	1.82	1.25
		vi	2.59	3.05	2.27	0.81	2.60	3.90	4.00	0.78
		vii	2.51	3.05	2.19	0.56	2.73	3.82	3.94	1.00
	50	i	2.82	3.52	3.22	0.23	3.33	3.52	3.77	0.66
		ii	2.82	3.53	3.25	0.25	3.29	3.53	3.77	0.66
		iii	1.59	2.24	1.40	0.93	1.81	2.09	1.55	0.71
		iv	2.89	3.51	3.28	0.22	3.34	3.53	3.80	0.75
		v	1.59	2.22	1.36	0.83	1.93	2.06	1.61	0.84
		vi	2.85	3.51	3.26	0.19	3.35	3.52	3.81	0.69
		vii	2.83	3.51	3.24	0.26	3.42	3.50	3.83	0.87

6.2 Motivation for using Formula (13)

The preliminary simulation study comparing the use $\widehat{\boldsymbol{\beta}}_{\min}(0.5)$ or $\widehat{\boldsymbol{\beta}}_{\min}(\theta)$ in (13) consisted of Monte Carlo simulations with $R = 500$ replications for these two alternatives, using the quantiles $\theta = 0.75$ and $\theta = 0.9$, with the u_{it} in (25) following the standard normal and standard lognormal distributions, respectively. Tables 16 and 17 give the mean differences over the weights i.-vii. between the mean absolute percent errors (MAPE) of $\widehat{\boldsymbol{\beta}}_{\min}(0.5)$ and $\widehat{\boldsymbol{\beta}}_{\min}(\theta)$. A negative value thus imply that $\widehat{\boldsymbol{\beta}}_{\min}(0.5)$ on average performs better than $\widehat{\boldsymbol{\beta}}_{\min}(\theta)$. As can be seen are most values negative, so that $\widehat{\boldsymbol{\beta}}_{\min}(0.5)$ overall is performing somewhat better than $\widehat{\boldsymbol{\beta}}_{\min}(\theta)$, which give the conclusion that this is to be preferred.

Table 16: Average differences over weights i-vii between MAPE values for using θ=0.5 and θ=θ in Formula (13), when θ=0.75

		T=5				T=10			
ρ	n	β_1	β_2	β_3	β_4	β_1	β_2	β_3	β_4
		Normal							
0	20	-1.19	-0.22	-0.88	-1.00	0.05	-0.03	-0.05	-0.08
	50	-0.82	-0.14	-0.50	-0.68	-0.11	-0.04	-0.10	-0.15
0.5	20	-0.44	-0.17	-0.44	-0.45	-0.46	-0.24	-0.48	-0.36
	50	-0.48	-0.07	-0.40	-0.54	-0.14	-0.06	-0.11	-0.28
0.95	20	-0.01	-0.05	-0.13	0.15	-0.07	-0.07	-0.06	-0.14
	50	-0.08	-0.04	-0.11	0.03	-0.05	-0.05	-0.07	-0.13
		Lognormal							
0	20	-0.66	-0.25	-0.56	-0.66	-0.20	-0.12	-0.18	-0.27
	50	-0.26	-0.11	-0.24	-0.57	-0.03	-0.06	-0.09	-0.13
0.5	20	-0.40	-0.19	-0.41	-0.56	-0.50	-0.22	-0.42	-0.64
	50	-0.21	-0.05	-0.19	-0.32	-0.10	-0.05	-0.13	-0.18
0.95	20	-0.09	-0.03	-0.11	-0.04	-0.10	-0.05	-0.19	-0.08
	50	0.03	-0.04	-0.07	-0.12	-0.04	-0.02	-0.08	-0.03

Table 17: Average differences over weights i-vii between MAPE values for using θ=0.5 and θ=θ in Formula (13), when θ=0.9

		T=5				T=10			
ρ	n	β_1	β_2	β_3	β_4	β_1	β_2	β_3	β_4
		Normal							
0	20	-0.75	-0.40	-0.81	-1.17	-0.63	-0.11	-0.29	-0.88
	50	-0.73	-0.27	-0.53	-0.84	-0.28	-0.11	-0.26	-0.34
0.5	20	0.06	-0.20	-0.22	-0.67	-0.25	-0.26	-0.39	-0.66
	50	-0.41	-0.21	-0.40	-0.69	-0.25	-0.16	-0.29	-0.42
0.95	20	0.10	-0.08	-0.05	0.06	-0.18	-0.16	-0.34	-0.23
	50	-0.10	-0.14	-0.18	-0.15	-0.11	-0.11	-0.21	-0.26
		Lognormal							
0	20	-0.20	-0.65	-0.97	-1.71	-0.52	-0.33	-0.62	-0.97
	50	-1.52	-0.62	-1.42	-1.29	-0.25	-0.17	-0.35	-0.60
0.5	20	-0.50	-0.46	-1.19	-0.95	-0.68	-0.47	-0.81	-1.47
	50	-0.70	-0.26	-0.71	-1.03	-0.33	-0.21	-0.42	-0.89
0.95	20	-0.10	-0.04	-0.08	-0.13	-0.23	-0.17	-0.36	-0.40
	50	-0.15	-0.13	-0.28	-0.25	-0.15	-0.13	-0.28	-0.41

References

Davidian, M. and Giltinan, D.M. (1993). Some general estimation methods for nonlinear mixed-effects models. *Journal of Biopharmaceutical Statistics* **3**, 23-55.

Davidian, M. and Giltinan, D.M. (1995). *Nonlinear Models for Repeated Measurement Data.* Chapman & Hall, London.

Davidian, M. and Giltinan, D.M. (2003). Nonlinear models for repeated measurement data: An overview and update. *Journal of Agricultural, Biological & Environmental Statistics* **8**, 387-419.

Davis, C.S. (2002). *Statistical Methods for the analysis of Repeated Measurements.* Springer, Berlin.

Diggle, P.J., Heagerty, P., Liang, K.-Y. and Zeger, S.L. (2002). *Analysis of Longitudinal Data.* Oxford University Press, Oxford.

He, X., Fu, B. and Fung, W.K. (2003). Median regression for longitudinal data. *Statistics in Medicine* **22** (23), 3655-3669.

He, X. and Kim, M.-O. (2002). On marginal estimation in a semiparametric model for longitudinal data with time-independent covariates. *Metrika* **55**, 67-74.

He, X., Zhu, Z.-Y. and Fung, W.-K. (2002). Estimation in a semiparametric model for longitudinal data with unspecified dependence structure. *Biometrika* **89**, 579-590.

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

Jureckova, J. and Prochazka, B. (1994). Regression quantiles and trimmed least squares estimator in nonlinear regression model. *Journal of Nonparametric Statistics* **3**, 201-222.

Koenker, R. (2004). Quantile regression for longitudinal data. *Journal of Multivariate Analysis* **91**, 74-89.

Koenker, R. and Bassett, G. (1978). Regression quantiles. *Econometrica* **46**, 33-50.

Lipsitz, S.R., Fitzmaurice, G.M., Molenberghs, G. and 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. *Applied Statistics* **46**, 463-476.

Pinheiro, J.C. and Bates, D.M. (2000). *Mixed-Effects Models in S and S-PLUS.* Springer, Berlin.

Yu, K. (2004). *Nonparametric quantile regression for longitudinal data analysis.* Unpublished manuscript.

Zeger, S.L. and Liang, K.-Y. (1992). An overview of methods for the analysis of longitudinal data. *Statistics in Medicine* **11**, 1825-1839.

Zhao, Q. (2001). Asymptotically efficient median regression in the presence of heteroscedasticity of unknown form. *Econometric Theory* **17**, 765-784.

Andreas Karlsson
Department of Information Science/Division of Statistics
Uppsala University
P. O. Box 513
S - 751 20 Uppsala
Sweden
E-mail: andreas.karlsson@dis.uu.se