*Journal of Mathematical Research with Applications*
May, 2021, Vol. 41, No. 3, pp. 303–322
DOI:10.3770/j.issn:2095-2651.2021.03.008
Http://jmre.dlut.edu.cn

# The Horvitz-Thompson Weighting Method for Quantile Regression Estimation in the Presence of Missing Covariates

**Zhaoji CHU**[1], **Lingnan TAI**[1], **Wei XIONG**[1,2], **Xu GUO**[1,4], **Maozai TIAN**[1,3,5,*]

1. *Center for Applied Statistics, School of Statistics, Renmin University of China, Beijing* 100872*, P. R. China;*
2. *School of Statistics, University of International Business and Economics, Beijing* 100029*, P. R. China;*
3. *Department of Medical Engineering and Technology, Xinjiang Medical University, Xinjiang* 830011*, P. R. China;*
4. *School of Statistics, Beijing Normal University, Beijing* 100875*, P. R. China;*
5. *School of Statistics and Information, Xinjiang University of Finance and Economics, Xinjiang* 830012*, P. R. China*

**Abstract** The lack of covariate data is one of the hotspots of modern statistical analysis. It often appears in surveys or interviews, and becomes more complex in the presence of heavy tailed, skewed, and heteroscedastic data. In this sense, a robust quantile regression method is more concerned. This paper presents an inverse weighted quantile regression method to explore the relationship between response and covariates. This method has several advantages over the naive estimator. On the one hand, it uses all available data and the missing covariates are allowed to be heavily correlated with the response; on the other hand, the estimator is uniform and asymptotically normal at all quantile levels. The effectiveness of this method is verified by simulation. Finally, in order to illustrate the effectiveness of this method, we extend it to the more general case, multivariate case and nonparametric case.

**Keywords** Robust quantile regression; missing covariates; selection probability; Kernel estimator; weighting method

**MR(2020) Subject Classification** 62G05; 62G08; 62G20

## 1. Introduction

Quantile regression was first introduced in [1], which is gradually emerging as a significant and unified statistical methodology for estimating models of conditional quantile functions. This kind of regression offers a systematic strategy for examining how covariates influence the location, scale and shape of the entire response distribution. It has been extensively used in economics, finance, insurance and medical research.

With the applications of quantile regression becoming universal and widespread, various extended models and estimation methods are springing up. For the simple parametric quantile

Received March 17, 2020; Accepted January 28, 2021
Supported by the National Natural Science Foundation of China (Grant No. 11861042) and the China Statistical Research Project (Grant No. 2020LZ25).
* Corresponding author
E-mail address: mztian@ruc.edu.cn (Maozai TIAN)

regression model, minimization of the check function is routinely used. A typical transformation such as Box-Cox transformation could be employed when the considered variable has a non-normal distribution [2]. In consideration of more complex cases, a nonparametric or semiparametric approach could be applied [3,4]. In the case of non-parametric model, kernel weighting method is one of the popular estimation procedure. However, some problems arise in application: first, this estimator is not a distribution function, then cannot be easily and directly calculated; second, quantile curves based on these conditional estimators may suffer from quantile crossing one another, which is absolutely absurd. To solve the first problem, Hall et al. [4] introduced a so-called adjusted Nadaraya-Waston estimator. Yu and Jones [3] usd a "double-kernel" to solve the second problem. Likewise, semiparametric quantile regression has also been discussed in great detail of literatures [5]. Especially in medicine of most concerned nowadays, the need for varied quantile curves rather a simple reference chart emerges when the measurements are strongly dependent on the covariates such as age, income level etc. Therefore, in such case, Cloe and Green [5] employed the weighted quantile regression model, with the response being weight and the covariate being age. Even better, when covariates are missing, censored quantile regression provides a powerful tool in survival analysis. Compared to classical Cox proportional hazard model, it relaxes the proportionality assumptions and can naturally accommodate heterogeneity of data. Moreover, by applying quantile regression survival time can be modeled directly, hence having practical significance.

Missing Data is a long-standing topic in statistical analysis. Among different missing-data patterns, nonresponse is a common problem in survey. But we do not study missing response but focus on missing covariate problems in our paper, which always occur to practical analysis and application. For instance, in a study on the association between acute graft versus host disease of bone marrow transplants, 97 females are follow-up surveyed by Fred Hutchinson Cancer Research Center. The covariate-donor's previous pregnancy status, of great interest, however missed for 31 patients due to the incompleteness of the donors' medical history.

In classical mean regression, various of missing mechanisms as well as estimating approaches have been finely developed to solve the missing covariate problems. There are four kinds of approaches: (1) complete case method (CC), which only just uses fully observed sample to estimate the interest parameters. However, it is well known that the CC analysis can be biased when the data are not missing completely at random (MCAR), this method is not good choice with complex data missing mechanisms although it is easy to implement. (2) imputation-based methods, which involve various methods such as single imputation, multiple imputation and Bayesian imputation [6] etc.. (3) likelihood-based methods, which assume the joint distribution for both covariates and responses and use observed data likelihood by integrating the missing covariate to obtain the model parameters, usually, the EM algorithms or quasi-Newton algorithms are used to resolve this problem [7] etc.. (4) weight-based methods, which use the inverse of some response probability as weight to adjust the observed portion and make unbiased estimators [8,9]. Otherwise, Ibrahim et al. [7] made detailed and full reviews about these methods in generalized linear model (GLM), [10] proposed semiparametric regression imputation estimator, marginal average estimator, and (marginal) propensity score weighted estimator in semiparametric partially linear regression model with missing response data.

Whereas in quantile regression framework, it is a newly topic to deal with missing data problem and there has not been a lot of works. For the missing values of nonresponse, Yoon [11] proposed a two-step estimating method in quantile regression scenario and demonstrates the consistency and asymptotic normality of the proposed estimator. In missing covariates patterns, Wei et al. [12] developed a multiple imputation method of estimating when missing mechanism is missing at random. Sherwood et al. [13] proposed a weighted quantile regression method to analyze health care cost data and developed a modified BIC for variable selection. Wei and Yang [14] constructed unbiased estimating equations, which is an extension of joint modeling method and provided an iterative EM-type algorithm. Recently, Tai et al. [15] built a multiple weighted estimating equations method to resolve the missing problem with quantile regression. Han et al. [16] proposed the multiple robust method in quantile regression to deal with the missing data.

Based on the above literature, we aim to construct a quantile regression method to deal with missing covariate which is established on the weight-based method. The weight is called selection probability, which is similar to Horvitz-Thompson weighting scheme (first proposed in [17]) that has little assumption of the distribution to covariates. Different from other estimate method of selection probability, we use the nonparametric method rather than parametric method (like logistic regression which may suffer from model misspecification risk). Our approach has analogous inspiration to [18] but there are different ideas and skills in proving theoretical properties. Furthermore, we extend the linear quantile regression to nonparametric quantile regression and express the main formal of missing quantile regression model.

The rest of the paper is organized as follows. The proposed method as well as its asymptotic properties are studied in Section 2, where bandwidth selection method is also discussed. A simulation study is presented in Section 3. In Section 4, several extensions are made to illustrate the usefulness of our proposed method. Section 5 concludes this paper with some discussions. Some tables are included in the Appendix.

## 2. Methodology

To solve the missing covariate problem in quantile regression, we propose a new estimator called inverse probability N-W weighted estimator. Further, we prove the asymptotic properties of the estimator.

### 2.1. Inverse probability N-W weighted estimator

We observe $(Y_i, X_i, \delta_i)$, where $Y_i$ is a one-dimensional response, $X_i$ is a one-dimensional covariate, and $\delta_i$ is a missing indicator, $\delta_i = 1$ when $X_i$ is observed and $\delta_i = 0$ when $X_i$ is missing. A case is considered where $X$ is missing at random (MAR), but not missing completely at random (MCAR). The missing probability is allowed to depend on the responses. That is, the

MAR assumption in this paper is that $X$ and $\delta$ are conditionally independent given response $Y$,

$$\pi_i = P(\delta_i = 1|Y_i, X_i) = P(\delta_i = 1|Y_i) = \pi(Y_i), \tag{2.1}$$

where $\pi_i$ is the selection probability mentioned above.

In common two-stage procedures, the selection probabilities are supposed to be known, which is impossible to achieve in most real missing data applications. To solve this problem, we proposed a Nadaraya-Waston (N-W) type estimator to estimate the selection probability $\pi_i$.

For the general linear model,

$$Y = X^T\beta + \varepsilon,$$

where $\beta$ is the unknown parameter, $\varepsilon$ is the error term. Given $X = x$, the $\tau$% conditional quantile function of response $Y$ is

$$Q_Y(\tau|\mathbf{x}) = \mathbf{x}^T\boldsymbol{\beta}, \tag{2.2}$$

where $Q_Y(\tau|\mathbf{x})$ represents the $\tau$th conditional quantile of $Y$ and we assume $Pr(\epsilon \le 0|\mathbf{x}) = \tau$. Thus, given a data set $\{(\mathbf{x}_i, y_i)\}_{i=1}^n$, $\beta$ can be estimated by minimizing $\sum_{i=1}^n \rho_\tau(y_i - \mathbf{x}_i^T\beta)$, where $\mathbf{x}_i = (1, x_i)^T$, $\rho_\tau(x) = \tau x I_{[0,\infty)}(x) + (\tau - 1)xI_{(-\infty,0)}(x)$ is the check function.

To accommodate the missing problem in observed data, we propose a Horvitz-Thompson inverse selection weighted method, in the case of the selection probability $\pi_i$ can be obtained in advance. Then $\beta$ can be estimated by

$$\widehat{\boldsymbol{\beta}} = \arg\min_{\beta} \sum_{i=1}^n \rho_\tau(y_i - \mathbf{x}_i^T\boldsymbol{\beta})\frac{\delta_i}{\pi_i}. \tag{2.3}$$

It can be seen from Eq. (2.3) that if the $i$th observation is missing ($\delta_i = 0$), it would not be included in the objective function, whereas it may play a big part by adjusting the weights $\pi$ in Eq. (2.3). In real applications, $\pi_i$ is usually unknown. Thus due to the strong correlation between $Y_i$ and $\delta_i$, we propose a Nadaraya-Waston type estimator to estimate selection probability $\pi_i$, that is

$$\widehat{\pi}_i = \widehat{\pi}(y_i) = \frac{\sum_{j=1}^n \delta_j K_h(y_i - y_j)}{\sum_{j=1}^n K_h(y_i - y_j)}, \tag{2.4}$$

where $K_h(\cdot) = K(\cdot/h)/h$ and $h$ is a bandwidth. By plugging into the $\hat{\pi}_i$ in Eq. (2.3), estimate of $\beta$ can be obtained by

$$\widehat{\boldsymbol{\beta}}^* = \arg\min_{\boldsymbol{\beta}} \sum_{i=1}^n \rho_\tau(y_i - \mathbf{x}_i^T\boldsymbol{\beta})\frac{\delta_i}{\widehat{\pi}_i}. \tag{2.5}$$

### 2.2. Asymptotic properties

Let $Y_1, \ldots, Y_n$ be independently and identically distributed with $F$. The conditional distribution functions of $Y_i$ is $P(Y_i < y|\mathbf{x}_i) = F(y)$, and we define

$$Q_{Y_i}(\tau|\mathbf{x}_i) = F^{-1}(\tau|\mathbf{x}_i) \equiv \xi(\tau),$$

where $\xi(\tau)$ is the real $\tau$th quantile of the distribution $F$. Then we have the following two theorems for the estimators with known and unknown selection probabilities.

Assumptions and conditions:

(1) The distribution $F$ is absolutely continuous, with continuous densities $f(\xi)$ uniformly bounded away from 0 and $\infty$ at the point $\xi(\tau)$.

(2) Define $D_0(\tau) = \lim_{n\to\infty} n^{-1}\sum \frac{1}{\pi_i}\mathbf{x}_i\mathbf{x}_i^T$, and $D_1(\tau) = \lim_{n\to\infty} n^{-1} f(\xi(\tau))\sum \mathbf{x}_i\mathbf{x}_i^T$. Both $D_0(\tau)$ and $D_1(\tau)$ are the positive definite matrices.

(3) $\max_{i=1,\dots,n} \parallel \mathbf{x}_i \parallel /\sqrt{n} \to 0$.

(4) The selection probability $\pi(Y) \geqslant c > 0$.

(5) The kernel function $K(\cdot)$ is a symmetric probability density with support $[-1, 1]$.

(6) $|\pi(Y)-\widehat{\pi}(Y)| = o_p(1)$ uniformly. $\widehat{\pi}(Y) \geqslant c^* > 0$ and $\widehat{\pi}(Y)$ has bounded partial derivatives up to order 2 almost surely.

(7) The density of $Y$, $f(y)$, has bounded derivatives up to order 2 on support $\mathcal{C}$, and satisfies $0 < \inf_{y\in\mathcal{C}} f(y) \leqslant \sup_{y\in\mathcal{C}} f(y) < \infty$.

**Theorem 2.1** *Under Conditions (1)–(4), we have*

$$\sqrt{n}(\widehat{\boldsymbol{\beta}}(\tau) - \boldsymbol{\beta}(\tau)) \to N(0, \tau(1-\tau)D_0^{-1}(\tau)D_1(\tau)D_0^{-1}(\tau)).$$

Theorem 2.1 is established based on the Nadaraya-Watson estimator of the unknown selection probability $\pi_i$.

**Theorem 2.2** *Under Conditions (1)–(7), we have*

$$\sqrt{n}(\widehat{\boldsymbol{\beta}}^*(\tau) - \boldsymbol{\beta}(\tau)) \to N(0, \tau(1-\tau)D_0^{-1}(\tau)D_1(\tau)D_0^{-1}(\tau)).$$

Theorems 2.1 and 2.2 suggest that, though the estimators are obtained via different assumptions and approaches, asymptotic properties are the same. That is to say, our proposed Nadaraya-Watson type estimator has an identical efficiency as if it is known ahead.

### 2.3. Bandwidth selection

To obtain a more efficient estimator $\hat{\pi}_i$, bandwidth $h$ plays a crucial role in balancing between bias and variance, as in Eq. (2.4). Thus in this section we present an effective method to choose an optimal bandwidth. Härdle [19] demonstrated the consistency of Nadaraya-Watson estimator of any unknown nonparametric function. Similar to that, the asymptotic mean square error (AMSE) of estimator $\widehat{\pi}_i$ is

$$\text{AMSE}\{\widehat{\pi}_h(y)\} = \frac{1}{nh}\frac{\sigma^2(y)}{f_Y(y)}\|K\|_2^2 + \frac{h^4}{4}\{\pi''(y) + 2\frac{\pi'(y)f_Y'(y)}{f_Y(y)}\}^2\mu_2^2(K),$$

where $\sigma^2(y) = \text{Var}(\pi(Y)|Y=y)$. Minimizing the AMSE with respect to $h$, we have the optimal bandwidth for $\widehat{\pi}(y)$, that is

$$h_{opt} = [\frac{\sigma^2(y)\|K\|_2^2}{n\{\pi''(y) + 2\frac{\pi'(y)f_Y'(y)}{f_Y(y)}\}^2\mu_2^2(K)f_Y(y)}]^{\frac{1}{5}}.$$

Note that in the above equation, $h \sim n^{-1/5}$ and both $\pi(y)$ and $f_Y(y)$ are unknown. An "ad hoc" plug-in bandwidth selection is to estimate $\pi(y)$ by a third or higher degree polynomial parametric regression, and to estimate $f_Y(y)$ through nonparametric way such as usual kernel density estimation.

## 3. Simulation studies

In this section, to investigate the efficiency and nice properties of our proposed estimators, we conduct some simulations. In each simulation we generate $n = 1000$ observations and compare the performances of weighted and original estimators for $\beta_0$ and $\beta_1$ in model (3.1), each refers to the intercept (-1.33) and coefficient (1.67), denoted by $orq$ (original quantile regression) and $wrq$ (weighted quantile regression). The original estimator ($orq$) is the same as complete case method (CC), which ignores all the missing observations and just uses the complete sample case to perform estimation without weight. Whereas, the weighted estimator ($wrq$) also only uses complete data but realizes estimation by constructing the inverse probability weighted estimation equation based on the selection probability. In the weighted estimators, $wrq_{\text{-}real}$ denotes the weight of estimator coming from the real selection probability, then the weights come from estimated selection probabilities $\hat{\pi}_i$s via Logistic, Local Polynomial and Nadaraya-Waston estimator are denoted as $wrq_{\text{-}log}$, $wrq_{\text{-}loess}$ and $wrq_{\text{-}nw}$, respectively. Especially, to examine the effects of quantile $\tau$ on the response, $\tau = 0.05, 0.1, 0.25, 0.5, 0.75, 0.9, 0.95$ are considered.

We generate 1000 observations $\{(x_i, y_i)\}$ from

$$Y = 1.67X - 1.33 + 0.6\varepsilon \tag{3.1}$$

where $X \sim U(2, 5)$, a uniform distribution, the error $\varepsilon$ follows a standard normal distribution. To model the missing data, the selection probability $\pi$ is assumed to be a piecewise function of $y_i$,

$$\pi(y_i) = \begin{cases} p_1, & y_i > 4 \\ 0.99, & y_i \leqslant 4 \end{cases} \tag{3.2}$$

where $p_1$ is set to be $0.1, \ldots, 0.9$, respectively. Thus $\delta(y_i)$ has a Bernoulli distribution, i.e.,

$$\delta(y_i) \sim \text{Bernoulli}(\pi(y_i)).$$

If $\delta(y_i) = 0$, we regard the $i$th observation $x_i$ as missing.

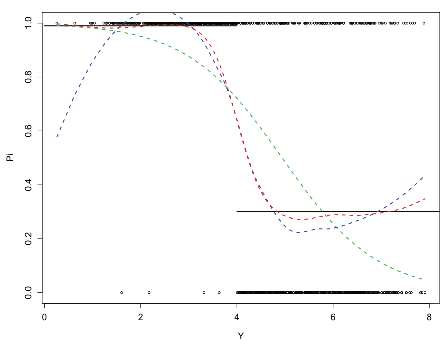


Figure 1 Logistic (green dotted line), Local Polynomial (blue dotted line) and Nadaraya-Waston (red dotted line) estimates of selection probability when $p_1 = 0.3$. The black circles represent $\delta(X_i)$, indicating the missingness of $i$th observation. The solid line is the real probabilities we assumed. If the estimates are beyond the interval [0,1], 0.01 or 0.99 is taken instead

To see the performance of estimators $\hat{\pi}$s, we take $p_1 = 0.3$ as an example. The estimated $\hat{\pi}_i$s is depicted in Figure 1. It can be observed from Figure 1 that the Nadaraya-Waston estimator fits $\pi_i$ best.

To estimate the unknown coefficients $\beta_0$ and $\beta_1$, first we use the real selection probability (3.2) denoted as $\hat{\boldsymbol{\beta}}_{wrq\text{-}real} = (\hat{\beta}_{wrq\text{-}real,0}, \hat{\beta}_{wrq\text{-}real,1})$, that is

$$\hat{\boldsymbol{\beta}}_{wrq\text{-}real}(\tau) = \arg\min_{\beta} \sum_{i=1}^{n} \rho_\tau(y_i - \mathbf{x}_i^T \boldsymbol{\beta}) \frac{\delta(y_i)}{\pi(y_i)}.$$

Then we employ the estimated selection probability $\hat{\pi}_i$s for the missing observations to obtain estimate $\boldsymbol{\beta}$, denoted as $\hat{\boldsymbol{\beta}}_{wrq\text{-}log}$, $\hat{\boldsymbol{\beta}}_{wrq\text{-}loess}$ and $\hat{\boldsymbol{\beta}}_{wrq\text{-}nw}$, respectively. To compare these estimators, we use relative error $re_1$ and $re_0$, where $re_0 = |(\hat{\beta}_0 - \beta_0)/\beta_0|$ and $re_1 = |(\hat{\beta}_1 - \beta_1)/\beta_1|$. Results are reported in Table 1, where the mean relative error of 1000 simulations are also reported.

| $p_1$ | Method | $\tau = 0.05$ | | $\tau = 0.1$ | | $\tau = 0.25$ | | $\tau = 0.5$ | | $\tau = 0.75$ | | $\tau = 0.9$ | | $\tau = 0.95$ | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | re$_0$ | re$_1$ | re$_0$ | re$_1$ | re$_0$ | re$_1$ | re$_0$ | re$_1$ | re$_0$ | re$_1$ | re$_0$ | re$_1$ | re$_0$ | re$_1$ |
| 0.1 | $orq$ | 0.182 | 0.108 | 0.192 | 0.107 | 0.213 | 0.104 | 0.207 | 0.086 | 0.188 | 0.062 | 0.274 | 0.046 | 0.514 | 0.042 |
| | $wrq_{\text{-}real}$ | 0.109 | 0.051 | 0.1 | 0.042 | 0.1 | 0.034 | 0.12 | 0.031 | 0.198 | 0.035 | 0.448 | 0.046 | 0.964 | 0.058 |
| | $wrq_{\text{-}log}$ | 0.293 | 0.142 | 0.285 | 0.129 | 0.267 | 0.107 | 0.266 | 0.084 | 0.349 | 0.072 | 0.889 | 0.087 | 2.626 | 0.14 |
| | $wrq_{\text{-}loess}$ | 0.121 | 0.057 | 0.112 | 0.047 | 0.125 | 0.04 | 0.181 | 0.042 | 0.326 | 0.051 | 0.696 | 0.067 | 1.275 | 0.075 |
| | $wrq_{\text{-}nw}$ | 0.112 | 0.051 | 0.099 | 0.041 | 0.094 | 0.033 | 0.109 | 0.029 | 0.171 | 0.031 | 0.359 | 0.038 | 0.775 | 0.049 |
| 0.2 | $orq$ | 0.115 | 0.064 | 0.106 | 0.058 | 0.093 | 0.046 | 0.086 | 0.033 | 0.128 | 0.026 | 0.292 | 0.026 | 0.614 | 0.032 |
| | $wrq_{\text{-}real}$ | 0.085 | 0.038 | 0.075 | 0.03 | 0.076 | 0.025 | 0.093 | 0.023 | 0.152 | 0.026 | 0.33 | 0.033 | 0.702 | 0.042 |
| | $wrq_{\text{-}log}$ | 0.271 | 0.125 | 0.292 | 0.125 | 0.341 | 0.124 | 0.437 | 0.125 | 0.582 | 0.119 | 0.918 | 0.115 | 1.381 | 0.102 |
| | $wrq_{\text{-}loess}$ | 0.091 | 0.04 | 0.075 | 0.03 | 0.072 | 0.024 | 0.088 | 0.022 | 0.143 | 0.024 | 0.312 | 0.031 | 0.672 | 0.04 |
| | $wrq_{\text{-}nw}$ | 0.087 | 0.038 | 0.075 | 0.03 | 0.074 | 0.024 | 0.086 | 0.022 | 0.137 | 0.024 | 0.285 | 0.029 | 0.596 | 0.037 |
| 0.3 | $orq$ | 0.088 | 0.045 | 0.072 | 0.035 | 0.065 | 0.027 | 0.074 | 0.02 | 0.138 | 0.019 | 0.308 | 0.024 | 0.617 | 0.031 |
| | $wrq_{\text{-}real}$ | 0.076 | 0.032 | 0.066 | 0.026 | 0.066 | 0.021 | 0.08 | 0.02 | 0.129 | 0.021 | 0.285 | 0.028 | 0.585 | 0.035 |
| | $wrq_{\text{-}log}$ | 0.173 | 0.08 | 0.192 | 0.081 | 0.246 | 0.087 | 0.329 | 0.091 | 0.486 | 0.096 | 0.802 | 0.097 | 1.3 | 0.095 |
| | $wrq_{\text{-}loess}$ | 0.077 | 0.033 | 0.067 | 0.026 | 0.063 | 0.021 | 0.073 | 0.018 | 0.12 | 0.02 | 0.259 | 0.025 | 0.534 | 0.031 |
| | $wrq_{\text{-}nw}$ | 0.074 | 0.032 | 0.064 | 0.025 | 0.063 | 0.021 | 0.074 | 0.018 | 0.12 | 0.02 | 0.255 | 0.026 | 0.525 | 0.031 |
| 0.4 | $orq$ | 0.075 | 0.034 | 0.068 | 0.028 | 0.061 | 0.021 | 0.078 | 0.017 | 0.139 | 0.019 | 0.292 | 0.023 | 0.574 | 0.03 |
| | $wrq_{\text{-}real}$ | 0.069 | 0.029 | 0.063 | 0.024 | 0.06 | 0.019 | 0.073 | 0.018 | 0.113 | 0.019 | 0.255 | 0.025 | 0.543 | 0.032 |
| | $wrq_{\text{-}log}$ | 0.118 | 0.052 | 0.133 | 0.054 | 0.173 | 0.06 | 0.242 | 0.066 | 0.364 | 0.07 | 0.625 | 0.073 | 1.066 | 0.075 |
| | $wrq_{\text{-}loess}$ | 0.07 | 0.029 | 0.064 | 0.024 | 0.058 | 0.018 | 0.067 | 0.016 | 0.106 | 0.017 | 0.235 | 0.023 | 0.501 | 0.029 |
| | $wrq_{\text{-}nw}$ | 0.069 | 0.029 | 0.063 | 0.024 | 0.058 | 0.018 | 0.068 | 0.016 | 0.107 | 0.018 | 0.232 | 0.023 | 0.501 | 0.029 |
| 0.5 | $orq$ | 0.07 | 0.03 | 0.06 | 0.023 | 0.061 | 0.018 | 0.077 | 0.017 | 0.128 | 0.018 | 0.277 | 0.023 | 0.553 | 0.029 |
| | $wrq_{\text{-}real}$ | 0.066 | 0.027 | 0.059 | 0.022 | 0.059 | 0.018 | 0.07 | 0.017 | 0.11 | 0.018 | 0.233 | 0.023 | 0.496 | 0.029 |
| | $wrq_{\text{-}log}$ | 0.088 | 0.038 | 0.092 | 0.037 | 0.119 | 0.04 | 0.171 | 0.046 | 0.268 | 0.05 | 0.505 | 0.057 | 0.898 | 0.061 |
| | $wrq_{\text{-}loess}$ | 0.066 | 0.027 | 0.058 | 0.021 | 0.058 | 0.018 | 0.068 | 0.016 | 0.106 | 0.017 | 0.222 | 0.021 | 0.478 | 0.028 |
| | $wrq_{\text{-}nw}$ | 0.065 | 0.027 | 0.058 | 0.021 | 0.058 | 0.018 | 0.067 | 0.016 | 0.105 | 0.017 | 0.227 | 0.022 | 0.481 | 0.028 |
| 0.6 | $orq$ | 0.066 | 0.027 | 0.059 | 0.022 | 0.058 | 0.017 | 0.075 | 0.016 | 0.124 | 0.017 | 0.247 | 0.021 | 0.484 | 0.026 |
| | $wrq_{\text{-}real}$ | 0.064 | 0.026 | 0.057 | 0.021 | 0.055 | 0.017 | 0.066 | 0.015 | 0.105 | 0.017 | 0.219 | 0.021 | 0.444 | 0.026 |
| | $wrq_{\text{-}log}$ | 0.075 | 0.031 | 0.075 | 0.029 | 0.093 | 0.03 | 0.129 | 0.033 | 0.206 | 0.037 | 0.384 | 0.042 | 0.694 | 0.046 |
| | $wrq_{\text{-}loess}$ | 0.064 | 0.026 | 0.056 | 0.02 | 0.054 | 0.016 | 0.064 | 0.015 | 0.1 | 0.016 | 0.209 | 0.02 | 0.438 | 0.026 |
| | $wrq_{\text{-}nw}$ | 0.063 | 0.025 | 0.055 | 0.02 | 0.054 | 0.016 | 0.065 | 0.015 | 0.101 | 0.016 | 0.209 | 0.02 | 0.434 | 0.026 |
| 0.7 | $orq$ | 0.062 | 0.024 | 0.056 | 0.02 | 0.056 | 0.016 | 0.07 | 0.015 | 0.112 | 0.016 | 0.23 | 0.02 | 0.455 | 0.025 |
| | $wrq_{\text{-}real}$ | 0.061 | 0.024 | 0.055 | 0.02 | 0.054 | 0.016 | 0.066 | 0.015 | 0.103 | 0.016 | 0.214 | 0.02 | 0.432 | 0.025 |
| | $wrq_{\text{-}log}$ | 0.066 | 0.026 | 0.064 | 0.024 | 0.071 | 0.022 | 0.096 | 0.024 | 0.154 | 0.027 | 0.298 | 0.031 | 0.574 | 0.036 |
| | $wrq_{\text{-}loess}$ | 0.061 | 0.024 | 0.055 | 0.02 | 0.054 | 0.016 | 0.064 | 0.014 | 0.1 | 0.016 | 0.207 | 0.019 | 0.414 | 0.024 |
| | $wrq_{\text{-}nw}$ | 0.061 | 0.024 | 0.055 | 0.02 | 0.054 | 0.016 | 0.064 | 0.014 | 0.101 | 0.016 | 0.209 | 0.019 | 0.417 | 0.024 |
| 0.8 | $orq$ | 0.061 | 0.024 | 0.055 | 0.02 | 0.052 | 0.015 | 0.067 | 0.015 | 0.105 | 0.016 | 0.208 | 0.019 | 0.413 | 0.023 |
| | $wrq_{\text{-}real}$ | 0.06 | 0.024 | 0.054 | 0.019 | 0.051 | 0.015 | 0.065 | 0.015 | 0.1 | 0.016 | 0.198 | 0.019 | 0.407 | 0.023 |
| | $wrq_{\text{-}log}$ | 0.061 | 0.024 | 0.057 | 0.02 | 0.056 | 0.017 | 0.076 | 0.018 | 0.122 | 0.02 | 0.238 | 0.024 | 0.462 | 0.028 |
| | $wrq_{\text{-}loess}$ | 0.06 | 0.024 | 0.054 | 0.019 | 0.05 | 0.015 | 0.063 | 0.014 | 0.098 | 0.015 | 0.195 | 0.019 | 0.407 | 0.023 |
| | $wrq_{\text{-}nw}$ | 0.06 | 0.024 | 0.054 | 0.019 | 0.05 | 0.015 | 0.064 | 0.014 | 0.098 | 0.015 | 0.195 | 0.019 | 0.407 | 0.023 |
| 0.9 | $orq$ | 0.058 | 0.023 | 0.052 | 0.018 | 0.053 | 0.015 | 0.063 | 0.014 | 0.097 | 0.015 | 0.203 | 0.019 | 0.415 | 0.023 |
| | $wrq_{\text{-}real}$ | 0.058 | 0.023 | 0.052 | 0.018 | 0.052 | 0.015 | 0.061 | 0.013 | 0.095 | 0.015 | 0.203 | 0.019 | 0.413 | 0.023 |
| | $wrq_{\text{-}log}$ | 0.058 | 0.022 | 0.053 | 0.019 | 0.054 | 0.016 | 0.065 | 0.015 | 0.101 | 0.016 | 0.209 | 0.02 | 0.424 | 0.024 |
| | $wrq_{\text{-}loess}$ | 0.058 | 0.023 | 0.052 | 0.018 | 0.053 | 0.015 | 0.062 | 0.014 | 0.094 | 0.014 | 0.2 | 0.018 | 0.411 | 0.023 |
| | $wrq_{\text{-}nw}$ | 0.058 | 0.023 | 0.052 | 0.018 | 0.052 | 0.015 | 0.062 | 0.014 | 0.095 | 0.014 | 0.202 | 0.019 | 0.412 | 0.023 |

Table 1 Mean relative error of several estimators

Parts of Table 1 display the results of $orq$ and $wrq_{\text{-}real}$ methods which are obtained only based on complete data case. Their difference is whether to employ the weights in the objective check function. As a result, in about 77.8% of our simulation results, the performance of $wrq_{\text{-}real}$ is better than $orq$ with given $p_1$s and $\tau$s. It is also evident from Table 1 that $wrq_{\text{-}real}$ method appears better when $\tau$ is small, and $orq$ gradually takes active part as $\tau$ gets bigger, which is due to both the sparseness of data for extreme quantile level and the weighting scheme. The weighting scheme strengthens the role of the several unmissingness data in objective function. Another phenomenon is the influence of $p_1$. We find that the relative errors are getting closer with the increase of $p_1$. That is reasonable, when $p_1$ is big enough, Eq. (3.2) indicates that $\pi(y_i)$ is almost a constant, making the weighting scheme useless. Over 90 percent of the results shows the superiority of the previous method. And only in extreme situations such as $\tau$-level is high and $p_1$ (3.2) is small enough, $orq$ is likely to be better than $wrq_{\text{-}nw}$, which is caused by leverage effect mentioned above.

In addition, it can be seen from Table 1 that $wrq_{\text{-}nw}$ has better performance than $wrq_{\text{-}real}$. That is to say, compared with real selection probability, N-W type estimator is much more preferable, which is obscure and a little bit interesting. However, leverage effect is also one of the reason for explaining this phenomenon, where smooth nonparametric estimate of $\pi(y_i)$ can counterbalance parts of the disadvantages caused by leverage effect and the weight. It is approved by Figure 1. The estimated $\hat{\pi}_i$ is to deviate from $\pi_i$ in the edge of range of $y_i$, that can neutralize some disadvantages caused by leverage effect. Some theoretical explanations can be found in Robins et al. [8].

| Method | $orq$ | $wrq_{\text{-}real}$ | $wrq_{\text{-}log}$ | $wrq_{\text{-}loess}$ | $wrq_{\text{-}nw}$ |
|---|---|---|---|---|---|
| $orq$ | - | 22.22% | 96.03% | 15.08% | 8.73% |
| $wrq_{\text{-}real}$ | 77.78% | - | 98.41% | 30.95% | 11.11% |
| $wrq_{\text{-}log}$ | 3.97% | 1.59% | - | 1.59% | 1.59% |
| $wrq_{\text{-}loess}$ | 84.92% | 69.05% | 98.41% | - | 39.68% |
| $wrq_{\text{-}nw}$ | 91.27% | 88.89% | 98.41% | 60.32% | - |

Table 2 Comparative result of mean relative error

As mentioned above, different estimators of $\pi_i$, such as Logistic [13], Local Polynomial and Nadaraya-Waston estimators, are also compared and presented in Table 1. For the sake of simplicity, Table 2 is also provided. Numbers in Table 2 are the comparative results of the corresponding methods, for instance, 77.78% in Row 2 and Column 1 mean that $wrq_{\text{-}real}$ is better performed than $orq$ in percentage of 77.78. From the result, $wrq_{\text{-}log}$ is even worse than $orq$ in most cases, that's because the selection probability (3.2) is a step function and not suitable for logistic estimation. In practice, non-parametric regression is usually better than parametric method for unknown function. Furthermore, even in most cases of the 1000 simulations $wrq_{\text{-}loess}$ is better than $orq$ or $wrq_{\text{-}real}$, it is inferior to $wrq_{\text{-}nw}$ in almost 60% cases.

From the results and analysis of Tables 1 and 2, we find that in most cases the performance of each estimator is roughly of the relationship as follows: $wrq_{\text{-}nw} > wrq_{\text{-}loess} > wrq_{\text{-}real} > orq > wrq_{\text{-}log}$. Thus we would like to further compare between these estimators by using relative error criteria, since estimators with small relative error are usually wanted and then investigate its dependence on sample size $n$. As the matter of convenience, we choose 3 typical methods, $wrq_{-nw}$, $wrq_{-real}$ and $orq$, to compare. Several sample size cases are considered here, $n = 200, n = 300, n = 500$ and $n = 1000$. Results of the proportion in 1000 replications are listed in Tables 3–5. $P(re_{estimaor1} < re_{estimator2})$ is a measure of estimator accuracy, it shows the proportion of trials in which estimator has smaller relative error.

It is evident from Table 3 that $wrq$ methods (stands for $wrq_{-real}$ and $wrq_{-nw}$) perform better when sample size $n$ gets larger. And $wrq_{-nw}$ is obviously more superior to $wrq_{-real}$. Tables 4 and 5 provide the relative errors of intercept and slope, respectively. It seems that, comparing with $wrq_{-real}$, $wrq_{-nw}$ has better estimation accuracy in terms of coefficient than the intercept. That is another evidence of the crucial influence of leverage effect to the estimators. We also illustrate the performance of $wrq$ methods in Figure 2. It is obvious that compared with $orq$ methods, $wrq$ estimates are closer to the true value except for $wrq_{-log}$.

| Sample Size | $P(\mathrm{re}_{wrq_{-real}} < \mathrm{re}_{orq})$ | $P(\mathrm{re}_{wrq_{-nw}} < \mathrm{re}_{orq})$ | $P(\mathrm{re}_{wrq_{-nw}} < \mathrm{re}_{wrq_{-real}})$ |
|---|---|---|---|
| $n = 200$ | 46.03% | 69.05% | 90.48% |
| $n = 300$ | 49.21% | 70.63% | 84.92% |
| $n = 500$ | 57.94% | 79.37% | 85.71% |
| $n = 1000$ | 77.78% | 91.27% | 88.89% |

Table 3 The effects of sample size on relative errors

| Sample Size | $P(\mathrm{re}_{wrq_{-real},0} < \mathrm{re}_{orq,0})$ | $P(\mathrm{re}_{wrq_{-nw},0} < \mathrm{re}_{orq,0})$ | $P(\mathrm{re}_{wrq_{-nw},0} < \mathrm{re}_{wrq_{-real},0})$ |
|---|---|---|---|
| $n = 200$ | 46.03% | 68.25% | 88.89% |
| n = 300 | 52.38% | 74.60% | 85.71% |
| $n = 500$ | 65.08% | 84.13% | 84.13% |
| $n = 1000$ | 82.54% | 93.65% | 87.30% |

Table 4 The effects of sample size on relative errors of intercept

| Sample Size | $P(\mathrm{re}_{wrq_{-real},1} < \mathrm{re}_{orq,1})$ | $P(\mathrm{re}_{wrq_{-nw},1} < \mathrm{re}_{orq,1})$ | $P(\mathrm{re}_{wrq_{-nw},1} < \mathrm{re}_{wrq_{-real},1})$ |
|---|---|---|---|
| $n = 200$ | 46.03% | 69.84% | 92.06% |
| $n = 300$ | 46.03% | 66.67% | 84.13% |
| $n = 500$ | 50.79% | 74.60% | 87.30% |
| $n = 1000$ | 73.02% | 88.89% | 90.48% |

Table 5 The effects of sample size on relative errors of slope

In addition to analyzing the relative error, we also simulate the Bias (BIAS), sample standard deviation error (SD), estimated standard error (SE) and coverage probability based on nominal target coverage of 95% (COV) of estimators among different estimated methods with varied quantiles and $p_1$ values (see Appendix). The estimated standard errors of estimators are calculated based on the bootstrap method. They are similar to the results in Table 1, that $wrq_{-nw}$ has better estimation effect with both intercept and slope.

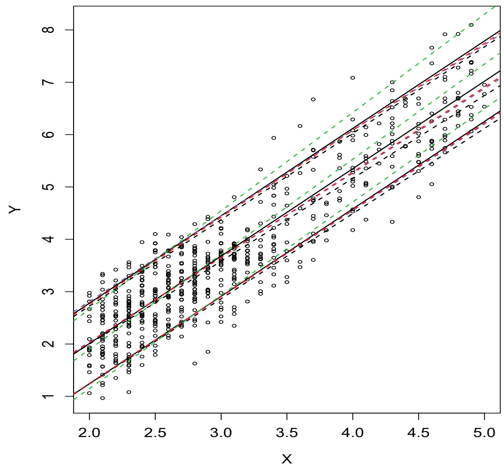


Figure 2 Comparison of the estimated lines ($p_1 = 0.3$, $n = 1000$, $\tau = 0.1, 0.5, 0.9$). The solid lines are the real regression lines as the standard of comparison, the black, purple, green and red dotted lines are based on $orq$, $wrq_{\text{-}real}$, $wrq_{\text{-}log}$ and $wrq_{\text{-}nw}$ methods, respectively.

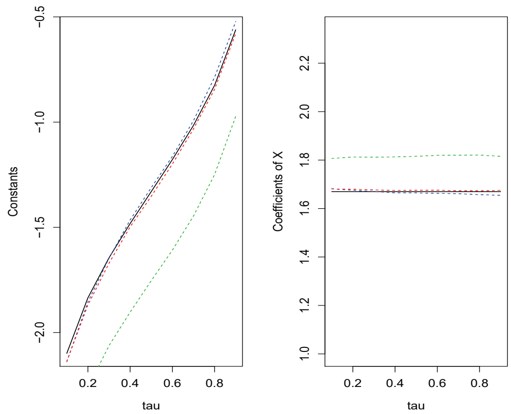


Figure 3 Comparison of the medians of estimated coefficients ($p_1 = 0.3$, $m = 1000$, $n = 1000$, $\tau = 0.1, \ldots, 0.9$). The solid lines are the real coefficients as the standard of comparison, the green, blue and red dotted lines are based on $wrq_{\text{-}log}$, $wrq_{\text{-}loess}$ and $wrq_{\text{-}nw}$ methods, respectively.

Figure 3 depicts the comparison results of the medians of estimated coefficients with various quantile level $\tau$s. It is strongly supported from this figure that $wrq_{\text{-}nw}$ and $wrq_{\text{-}loess}$ estimators are more robust and efficient than $wrq_{\text{-}log}$ estimators. In addition, the performance of $wrq_{\text{-}nw}$ estimator of coefficient is better than $wrq_{\text{-}loess}$, especially when $\tau > 0.5$. Considering with the estimation circles in Figure 1, we can confirm that $wrq_{\text{-}nw}$ is a better estimation than $wrq_{\text{-}loess}$ with respect to our example.

## 4. Generalization

In this section, we would like to generalize our proposed methods to more complicated cases. First multivariate case is considered, second nonparametric techniques are further applied to extend our linear parametric model specification. Local polynomial method is mainly considered here to take an example. Note almost all the nonparametric methods can be applied here.

### 4.1. Multivariate linear model

In the above section, $X$ is assumed to be a one-dimensional variable, which seems to be more restricted. In this section, we generalize our method to the more general settings: that is $X$ can be multivariate, high-dimensional. In the missing data scenario, some of the covariates are assumed to be missing at random while the others not. Thus let $X = (U, T)$, and $U$ is MAR, while $Y$ and $T$ are fully observed. Here, $U$ and $T$ are $d_1$ and $d_2$-dimensional random vectors respectively. Define missing indicator $\delta_i$ for the $i$th individual. $\delta_i = 1$ when $U_i$ is observed, $\delta_i = 0$ when $U_i$ is missing. Then the MAR assumption implies

$$\pi_i = P(\delta_i = 1|Y_i, U_i, T_i) = P(\delta_i = 1|Y_i, T_i) = \pi(\mathbf{S}_i),$$

where $\mathbf{S}_i = (Y_i, T_i)$. Furthermore, $\pi_i$ can be estimated by

$$\widehat{\pi}_i = \widehat{\pi}(\mathbf{s}_i) = \frac{\sum_{j=1}^n \delta_j \mathcal{K}_{\mathbf{H}}(\mathbf{s}_i - \mathbf{s}_j)}{\sum_{j=1}^n \mathcal{K}_{\mathbf{H}}(\mathbf{s}_i - \mathbf{s}_j)},$$

where $\mathcal{K}(\cdot)$ is a multivariate kernel function. $\mathbf{H}$ is a bandwidth matrix. In such case, $\widehat{\pi}_i$ is a weighted average of those $\delta_i$ where $\mathbf{s}_j$ lies in a ball or cube around $\mathbf{s}_i$. More details can be referred to Härdle et al. [18]. After chosen the appropriate kernel and bandwidth, $\widehat{\pi}_i$ can be estimated finely and $\beta$ can be obtained via

$$\widehat{\boldsymbol{\beta}}^* = \arg\min_{\boldsymbol{\beta}} \sum_{i=1}^n \rho_\tau(y_i - \mathbf{x}_i^T \boldsymbol{\beta}) \frac{\delta_i}{\widehat{\pi}(\mathbf{s}_i)}.$$

**Remark 4.1** In the multivariate case where $U = (U_1, \ldots, U_{d_1})$, we mean $\delta = 0$ provided any variable in $U$ is missing, one or two or even more. Thus under this circumstance, just using CC data to conduct estimation is a great loss of information, leading to the inefficiency of the resulting estimator. Therefore on the occasion the imputation method seems to have more advantages. However, the assumption in this paper is that the missing covariates are strongly related with the observed variables, therefore we suppose that most of $U$ is missing for simplicity and easiness to handle. In terms of the selection of the multivariate kernel function, a type of multivariate kernels, which is called spherical or radial-symmetric can be obtained from univariate kernel functions by taking $\mathcal{K}(u) \propto K(\|u\|)$, where $\|u\| = \sqrt{u^T u}$ denotes the Euclidean norm of the vector $u$.

### 4.2. Nonparametric model

When the relationship between $X$ and $Y$ is not linear, nonparametric model is a good choice. In this section, we consider generalization of proposed methods to nonparametric model. To explore the relationship between $X$ and $Y$, we assume that

$$Y = m(X) + \varepsilon,$$

where $m(x)$ is a unknown function, $\varepsilon$ is the error term satisfying $E(\varepsilon) = 0$. We define the $\tau$th conditional quantile of $Y$ as $q_\tau(x)$. To estimate $q_\tau(x)$, local linear method is employed. The idea of local linear quantile regression is to approximate $q_\tau(x)$ by a linear function

$$q_\tau(t) \approx q_\tau(x) + q'_\tau(x)(t - x) \equiv a + b(t - x)$$

for $t$ in a neighborhood of $x$. Thus, estimating $q_\tau(x)$ and $q'_\tau(x)$ is equivalent to estimating $a$ and $b$. Therefore in the missing covariate case, given a sample $\{x_i, y_i\}_{i=1}^n$, $\widehat{a}$ and $\widehat{b}$ can be obtained by minimizing

$$\sum_{i=1}^{n} \rho_\tau(y_i - a - b(x_i - x))K(\frac{x - x_i}{h_1})\frac{\delta_i}{\widehat{\pi}(y_i)},$$

where $h_1$ is a smoothing bandwidth, $\widehat{\pi}$ can be obtained by Eq. (2.4). Note that to have a good estimate in nonparametric model, we should choose two bandwidths $h$ and $h_1$, each representing the smoothness of selection probability and local linear function. Yu and Jones [3] presented the optimal bandwidth in local linear quantile regression.

## 5. Conclusion and discussions

Missing covariates is a focus topic in statistical analysis. In this paper, a weighted quantile regression method is considered to deal with this problem. By adding effective weights into the original objective function, the resulting estimator is robust, efficient and consistent. This method is also demonstrated to have more extensions beyond simple linear regression. The missing probability could be estimated by various parametric and nonparametric methods. And it's verified that the method using nonparametric estimated missing probability has the same efficiency as it is known in advance. Results of the simulation suggest the usefulness and practicality of the proposed estimator. In addition, our method can be generalized to more complicated models, such as nonparametric quantile regression models. For instance, we provide a brief description of local linear quantile regression model, and estimate it by the weighted method. Actually, More models and missing assumptions should be concerned using the provided weighted estimation method, which is dependent upon future research.

## 6. Proof of theorems

Assumptions and conditions:

(1) The distribution $F$ is absolutely continuous, with continuous densities $f(\xi)$ uniformly bounded away from 0 and $\infty$ at the point $\xi(\tau)$.

(2) Define $D_0(\tau) = \lim_{n\to\infty} n^{-1}\sum \frac{1}{\pi_i}\mathbf{x}_i\mathbf{x}_i^T$, and $D_1(\tau) = \lim_{n\to\infty} n^{-1} f(\xi(\tau))\sum \mathbf{x}_i\mathbf{x}_i^T$. Both $D_0(\tau)$ and $D_1(\tau)$ are the positive definite matrices

(3) $\max_{i=1,\ldots,n} \parallel \mathbf{x}_i \parallel /\sqrt{n} \to 0$.

(4) The selection probability $\pi(Y) \geqslant c > 0$.

(5) The kernel function $K(\cdot)$ is a symmetric probability density with support $[-1, 1]$.

(6) $|\pi(Y) - \widehat{\pi}(Y)| = o_p(1)$ uniformly. $\widehat{\pi}(Y) \geqslant c^* > 0$ and $\widehat{\pi}(Y)$ has bounded partial derivatives up to order 2 almost surely.

(7) The density of Y, $f(y)$, has bounded derivatives up to order 2 on support $\mathcal{C}$, and satisfies $0 < \inf_{y \in \mathcal{C}} f(y) \leqslant \sup_{y \in \mathcal{C}} f(y) < \infty$.

**Proof of Theorem 2.1** The behavior of $\sqrt{n}(\widehat{\boldsymbol{\beta}}(\tau) - \boldsymbol{\beta}(\tau))$ follows from consideration of the objective function

$$Z_n(\eta) = \sum_{i=1}^{n} [\rho_\tau(u_i - \mathbf{x}_i^T \eta / \sqrt{n}) - \rho_\tau(u_i)] \frac{\delta_i}{\pi_i},$$

where $u_i = y_i - \mathbf{x}_i^T \boldsymbol{\beta}$. The function $Z_n(\eta)$ is obviously convex and is minimized at $\widehat{\eta}_n = \sqrt{n}(\widehat{\boldsymbol{\beta}}(\tau) - \boldsymbol{\beta}(\tau))$. Using Knight's identity,

$$\rho_\tau(u - v) - \rho_\tau(u) = -v\psi_\tau(u) + \int_0^v (I(u \leq s) - I(u \leq 0)) \mathrm{d}s,$$

with $\psi_\tau(u) = \tau - I\ (u < 0)$. We may write $Z_n(\eta) = Z_{1n}(\eta) + Z_{2n}(\eta)$, where

$$Z_{1n}(\eta) = -\frac{1}{\sqrt{n}} \sum_{i=1}^{n} (\mathbf{x}_i^T \eta \psi_\tau(u_i)) \frac{\delta_i}{\pi_i}$$

$$Z_{2n}(\eta) = \sum_{i=1}^{n} \Big( \int_0^{v_{ni}} (I(u_i \leq s) - I(u_i \leq 0)) \mathrm{d}s \Big) \frac{\delta_i}{\pi_i} \equiv \sum_{i=1}^{n} Z_{2ni}(\eta)$$

and $v_{ni} = \mathbf{x}_i^T \eta / \sqrt{n}$. We have

$$\begin{aligned}
EZ_{1n}(\eta) &= -\frac{1}{\sqrt{n}} \sum_{i=1}^{n} E(\mathbf{x}_i^T \eta \psi_\tau(u_i) \frac{\delta_i}{\pi_i}) \\
&= -\frac{1}{\sqrt{n}} \sum_{i=1}^{n} E(E(\mathbf{x}_i^T \eta \psi_\tau(u_i) \frac{\delta_i}{\pi_i} | y_i, \mathbf{x}_i)) \\
&= -\frac{1}{\sqrt{n}} \sum_{i=1}^{n} E(\mathbf{x}_i^T \eta \psi_\tau(u_i) E(\frac{\delta_i}{\pi_i} | y_i, \mathbf{x}_i)) \\
&= -\frac{1}{\sqrt{n}} \sum_{i=1}^{n} E(\mathbf{x}_i^T \eta \psi_\tau(u_i) \frac{E(\delta_i | y_i, \mathbf{x}_i)}{\pi_i}) \\
&= -\frac{1}{\sqrt{n}} \sum_{i=1}^{n} E(\mathbf{x}_i^T \eta \psi_\tau(u_i)) \\
&= -\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \mathbf{x}_i^T \eta E(\tau - I(u_i < 0)) = 0.
\end{aligned}$$

$$\begin{aligned}
\mathrm{Var} Z_{1n}(\eta) &= \frac{1}{n} \sum_{i=1}^{n} E(\mathbf{x}_i^T \eta \psi_\tau(u_i) \frac{\delta_i}{\pi_i})^2 \\
&= \frac{1}{n} \sum_{i=1}^{n} E(E((\mathbf{x}_i^T \eta)^2 \psi_\tau^2(u_i) \frac{\delta_i}{\pi_i^2} | y_i, \mathbf{x}_i)) \\
&= \frac{1}{n} \sum_{i=1}^{n} E((\mathbf{x}_i^T \eta)^2 \psi_\tau^2(u_i) E(\frac{\delta_i}{\pi_i^2} | y_i, \mathbf{x}_i))
\end{aligned}$$

$$= \frac{1}{n}\sum_{i=1}^{n} E((\mathbf{x}_i^T\eta)^2\psi_\tau^2(u_i)/\pi_i)$$
$$= \frac{1}{n}\sum_{i=1}^{n}(\mathbf{x}_i^T\eta)^2 E(\tau^2/\pi_i + (1-2\tau)I(u_i<0)/\pi_i)$$
$$= \frac{(\tau(1-\tau))}{n}\sum_{i=1}^{n}\frac{(\mathbf{x}_i^T\eta)^2}{\pi_i}$$

It follows from the Lindeberg-Feller central limit theorem, using Condition (2), that $Z_{1n}(\eta) \to -\eta^T W$ where $W \sim N(0, \tau(1-\tau)D_0(\tau))$.

As for $Z_{2n}(\eta)$, we have

$$\begin{aligned}
EZ_{2n}(\eta) &= \sum_{i=1}^{n} EZ_{2ni}(\eta)\\
&= \sum_{i=1}^{n} E\Big(\Big(\int_0^{v_{ni}}(I(u_i\le s)-I(u_i\le 0))\mathrm{d}s\Big)\frac{\delta_i}{\pi_i}\Big)\\
&= \sum_{i=1}^{n} E\Big(E\Big(\Big(\int_0^{v_{ni}}(I(u_i\le s)-I(u_i\le 0))\mathrm{d}s\Big)\frac{\delta_i}{\pi_i}|y_i,\mathbf{x}_i\Big)\Big)\\
&= \sum_{i=1}^{n} E\Big(\int_0^{v_{ni}}(I(u_i\le s)-I(u_i\le 0))\mathrm{d}s\Big)\\
&= \sum_{i=1}^{n}\int_0^{v_{ni}}(F(\xi+s)-F(\xi))\mathrm{d}s\\
&= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\int_0^{\mathbf{x}_i^T\eta}(F(\xi+t/\sqrt{n})-F(\xi\sqrt{n}))\mathrm{d}t\\
&= n^{-1}\sum_{i=1}^{n}\int_0^{\mathbf{x}_i^T\eta}\sqrt{n}(F(\xi+t/\sqrt{n})-F(\xi\sqrt{n}))\mathrm{d}t\\
&= n^{-1}f(\xi)\sum_{i=1}^{n}\int_0^{\mathbf{x}_i^T\eta}t\mathrm{d}t+o(1)\\
&= (2n)^{-1}f(\xi)\sum_{i=1}^{n}\eta^T\mathbf{x}_i\mathbf{x}_i^T\eta+o(1)\\
&\to \frac{1}{2}\eta^T D_1(\tau)\eta.
\end{aligned}$$

The first equation follows from the independence in Condition (1).

$$\begin{aligned}
\mathrm{Var}(Z_{2n}(\eta)) &= \sum_{i=1}^{n}\mathrm{Var}(Z_{2ni}(\eta)) \le \sum_{i=1}^{n}E(Z_{2ni}(\eta))^2\\
&\le \max|Z_{2ni}(\eta)|\sum_{i=1}^{n}E(Z_{2ni}(\eta))\\
&\le \frac{1}{c\sqrt{n}}\max|\mathbf{x}_i^T\eta|\sum_{i=1}^{n}EZ_{2ni}(\eta)\to 0,
\end{aligned}$$

where $c = \min(\pi_i) > 0$ mentioned in Condition (4). The above bound and Condition (3) implies that

$$Z_n(\eta) \to Z_0(\eta) = -\eta^T W + \frac{1}{2}\eta^T D_1(\tau)\eta.$$

The convexity of the limiting objective function, $Z_0(\eta)$, assure the uniqueness of the minimizer, and consequently,

$$\widehat{\eta}_n = \sqrt{n}(\widehat{\boldsymbol{\beta}}(\tau) - \boldsymbol{\beta}(\tau)) = \text{argmin} Z_n(\eta) \to \widehat{\eta}_0 = \text{argmin} Z_0(\eta).$$

Finally, we see that $\widehat{\eta}_0 = D_1^{-1}(\tau)W$ and the result follows. □

**Proof of Theorem 2.2** Similarly, the objective function

$$Z_n^*(\eta) = \sum_{i=1}^n [\rho_\tau(u_i - \mathbf{x}_i^T\eta/\sqrt{n}) - \rho_\tau(u_i)]\frac{\delta_i}{\widehat{\pi}_i},$$

is convex and is minimized at

$$\widehat{\eta}_n^* = \sqrt{n}(\widehat{\boldsymbol{\beta}}^*(\tau) - \boldsymbol{\beta}(\tau)).$$

Using Knight's identity, we get

$$Z_n^*(\eta) = Z_{1n}^*(\eta) + Z_{2n}^*(\eta),$$

where

$$Z_{1n}^*(\eta) = -\frac{1}{\sqrt{n}}\sum_{i=1}^n (\mathbf{x}_i^T\eta\psi_\tau(u_i))\frac{\delta_i}{\widehat{\pi}_i},$$
$$Z_{2n}^*(\eta) = \sum_{i=1}^n \Big(\int_0^{v_{ni}} (I(u_i \le s) - I(u_i \le 0))\mathrm{d}s\Big)\frac{\delta_i}{\widehat{\pi}_i} \equiv \sum_{i=1}^n Z_{2ni}^*(\eta)$$

and $v_{ni} = \mathbf{x}_i^T\eta/\sqrt{n}$.

By a linearization technique in [20],

$$\begin{aligned}
\frac{\delta_i}{\widehat{\pi}_i} &= \frac{\delta_i}{\pi_i} + (\frac{\delta_i}{\widehat{\pi}_i} - \frac{\delta_i}{\pi_i}) = \frac{\delta_i}{\pi_i} + (\frac{\delta_i(\pi_i - \widehat{\pi}_i)}{\pi_i^2} + \frac{\delta_i(\pi_i - \widehat{\pi}_i)^2}{\pi_i^2\widehat{\pi}_i}) \\
&= \frac{\delta_i}{\pi_i} + \frac{\delta_i(\pi_i - \widehat{\pi}_i)}{\pi_i^2}(1 + \frac{\pi_i - \widehat{\pi}_i}{\widehat{\pi}_i}) \\
&= \frac{\delta_i}{\pi_i} + \frac{\delta_i(\pi_i - \widehat{\pi}_i)}{\pi_i^2}(1 + o_p(1)),
\end{aligned}$$

we can derive

$$\begin{aligned}
Z_{1n}^*(\eta) &= -\frac{1}{\sqrt{n}}\sum_{i=1}^n (\mathbf{x}_i^T\eta\psi_\tau(u_i))\frac{\delta_i}{\widehat{\pi}_i} \\
&= -\frac{1}{\sqrt{n}}\sum_{i=1}^n \mathbf{x}_i^T\eta\psi_\tau(u_i)(\frac{\delta_i}{\pi_i} + \frac{\delta_i(\pi_i - \widehat{\pi}_i)}{\pi_i^2}(1 + o_p(1))) \\
&= -\frac{1}{\sqrt{n}}\sum_{i=1}^n \mathbf{x}_i^T\eta\psi_\tau(u_i)\frac{\delta_i}{\pi_i} - \frac{1}{\sqrt{n}}\sum_{i=1}^n \mathbf{x}_i^T\eta\psi_\tau(u_i)\frac{\delta_i(\pi_i - \widehat{\pi}_i)}{\pi_i^2}(1 + o_p(1)) \\
&= A_1 + A_2(1 + o_p(1)).
\end{aligned}$$

From the Proof of Theorem 2.1, we have $E(A_1)=0$, and it can be verified that

$$
\begin{aligned}
A_2 =& -\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{\delta_i(\pi_i-\widehat{\pi}_i)}{\pi_i^2}\mathbf{x}_i^T\eta\psi_\tau(u_i)\\
=& -\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{\delta_i}{\pi_i^2}\frac{\sum_{j=1}^{n}(\pi_i-\delta_i)K_h(y_i-y_j)}{\sum_{j=1}^{n}K_h(y_i-y_j)}\mathbf{x}_i^T\eta\psi_\tau(u_i)\\
=& -\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{\delta_i}{\pi_i^2}\frac{\sum_{j=1}^{n}(\pi_i-\delta_i)K_h(y_i-y_j)}{n\widehat{f}(y_i)}\mathbf{x}_i^T\eta\psi_\tau(u_i)\\
=& -\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{\delta_i}{\pi_i^2}\frac{\sum_{j=1}^{n}(\pi_i-\delta_i)K_h(y_i-y_j)}{nf(y_i)}\mathbf{x}_i^T\eta\psi_\tau(u_i)+o_p(1)\\
=& -\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{\delta_i}{\pi_i^2f(y_i)nh}\mathbf{x}_i^T\eta\psi_\tau(u_i)\sum_{j=1}^{n}(\pi_i-\delta_i)K_h(y_i-y_j)+o_p(1)\\
=& -\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{\delta_i}{\pi^2(y_i)f(y_i)nh}\mathbf{x}_i^T\eta\psi_\tau(u_i)\sum_{j=1}^{n}(\pi(y_i)-\pi(y_j))K_h(y_i-y_j)-\\
&\frac{1}{\sqrt{n}}\sum_{j=1}^{n}(\delta_j-\pi(y_j))\frac{1}{nh}\sum_{i=1}^{n}\frac{\delta_i}{\pi^2(y_i)f(y_i)}\mathbf{x}_i^T\eta\psi_\tau(u_i)K(\frac{y_i-y_j}{h})+o_p(1)\\
=&o_p(1)-\frac{1}{\sqrt{n}}\sum_{j=1}^{n}\frac{\delta_j-\pi(y_j)}{\pi(y_j)}E(\mathbf{x}_i^T\eta\psi_\tau(u_i)|y_j)+o_p(1)\\
=& -\frac{1}{\sqrt{n}}\sum_{j=1}^{n}\frac{\delta_j-\pi(y_j)}{\pi(y_j)}E(\mathbf{x}_i^T\eta\psi_\tau(u_i)|y_j)+o_p(1),
\end{aligned}
$$

then

$$
\begin{aligned}
EA_2 =&E(-\frac{1}{\sqrt{n}}\sum_{j=1}^{n}\frac{\delta_j-\pi(y_j)}{\pi(y_j)}E(\mathbf{x}_i^T\eta\psi_\tau(u_i)|y_j)+o_p(1))\\
=& -\frac{1}{\sqrt{n}}\sum_{j=1}^{n}E(\frac{\delta_j-\pi(y_j)}{\pi(y_j)}E(\mathbf{x}_i^T\eta\psi_\tau(u_i)|y_j))+E(o_p(1))\to 0
\end{aligned}
$$

and we have

$$EZ_{1n}^*(\eta)=E(A_1)+E(A_2)(1+o_p(1))\to 0.$$

For the variance, we have

$$
\begin{aligned}
\mathrm{Var}Z_{1n}^*(\eta) =& \frac{1}{n}\sum_{i=1}^{n}E(\mathbf{x}_i^T\eta\psi_\tau(u_i)\frac{\delta_i}{\widehat{\pi}_i})^2\\
=& \frac{1}{n}\sum_{i=1}^{n}E(\mathbf{x}_i^T\eta\psi_\tau(u_i)(\frac{\delta_i}{\pi_i}+\frac{\delta_i(\pi_i-\widehat{\pi}_i)}{\pi_i^2}(1+o_p(1))))^2\\
=& \frac{1}{n}\sum_{i=1}^{n}E(\mathbf{x}_i^T\eta\psi_\tau(u_i)(\frac{\delta_i}{\pi_i^2}+\frac{2\delta_i(\pi_i-\widehat{\pi}_i)}{\pi_i^3}(1+o_p(1))+\frac{\delta_i(\pi_i-\widehat{\pi}_i)^2}{\pi_i^4}(1+o_p(1))))\\
=& \frac{1}{n}\sum_{i=1}^{n}E(\mathbf{x}_i^T\eta\psi_\tau(u_i)(\frac{\delta_i}{\pi_i^2}+\frac{\delta_i(\pi_i-\widehat{\pi}_i)}{\pi_i^3}(2+o_p(1))))
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{1}{n}\sum_{i=1}^{n} E(\mathbf{x}_i^T \eta\psi_\tau(u_i)\frac{\delta_i}{\pi_i^2}) + \frac{1}{n}\sum_{i=1}^{n} E(\mathbf{x}_i^T \eta\psi_\tau(u_i)\frac{\delta_i(\pi_i - \widehat{\pi}_i)}{\pi_i^3})(2+o_p(1)) \\
&= \mathrm{Var} Z_{1n}(\eta) + \frac{1}{n}E(\sum_{i=1}^{n}\frac{\delta_i}{\pi^3(y_i)}(\pi(y_i)-\widehat{\pi}(y_i))\mathbf{x}_i^T\eta\psi_\tau(u_i))(2+o_p(1)) \\
&= \mathrm{Var} Z_{1n}(\eta) + \frac{1}{n}E(\sum_{i=1}^{n}\frac{\delta_i}{\pi^3(y_i)}\frac{\sum_{j=1}^{n}(\pi(y_i)-\delta_j)K_h(y_i-y_j)}{nf(y_i)}\mathbf{x}_i^T\eta\psi_\tau(u_i) + o_p(1))\cdot \\
&\quad (2+o_p(1)) \\
&= \mathrm{Var} Z_{1n}(\eta) + \frac{1}{n}E(\sum_{j=1}^{n}\frac{\pi(y_j)-\delta_j}{\pi^2(y_j)}E(\mathbf{x}_i^T\eta\psi_\tau(u_i)|y_i) + o_p(1))(2+o_p(1)) \\
&= \mathrm{Var} Z_{1n}(\eta).
\end{aligned}
$$

Then it can be verified that $Z_{1n}^*(\eta) \to -\eta^T W$ where $W \sim N(0, \tau(1-\tau)D_0(\tau))$, using the Lindeberg-Feller central limit theorem.

As for $Z_{2n}^*(\eta)$, we have

$$
\begin{aligned}
EZ_{2n}^*(\eta) =& \sum_{i=1}^{n} EZ_{2ni}^*(\eta) = \sum_{i=1}^{n} E\Big(\Big(\int_0^{v_{ni}}(I(u_i\le s)-I(u_i\le 0))\mathrm{d}s\Big)\frac{\delta_i}{\widehat{\pi}_i}\Big) \\
=& \sum_{i=1}^{n} E\Big(\Big(\int_0^{v_{ni}}(I(u_i\le s)-I(u_i\le 0))\mathrm{d}s\Big)(\frac{\delta_i}{\pi_i} + \frac{\delta_i(\pi_i-\widehat{\pi}_i)}{\pi_i^2}(1+o_p(1)))\Big) \\
=& \sum_{i=1}^{n} E\Big(\int_0^{v_{ni}}(I(u_i\le s)-I(u_i\le 0))\mathrm{d}s\frac{\delta_i}{\pi_i}\Big) + \\
& \Big(\sum_{i=1}^{n} E\Big(\int_0^{v_{ni}}(I(u_i\le s)-I(u_i\le 0))\mathrm{d}s\frac{\delta_i}{\pi_i^2}(\pi_i-\widehat{\pi}_i)\Big)\Big)(1+o_p(1)) \\
=& EZ_{2n}(\eta) + E\Big(\sum_{i=1}^{n}\frac{\delta_i}{\pi^2(y_i)}\frac{\sum_{j=1}^{n}(\pi(y_i)-\delta_j)K_h(y_i-y_j)}{nf(y_i)}\cdot \\
& \Big(\int_0^{v_{ni}}(I(u_i\le s)-I(u_i\le 0))\mathrm{d}s\Big) + o_p(1)\Big)(1+o_p(1)) \\
=& EZ_{2n}(\eta) + E\Big(\sum_{j=1}^{n}\frac{\pi(y_i)-\delta_j}{\pi(y_i)}E\Big(\int_0^{v_{ni}}(I(u_i\le s)-I(u_i\le 0))\mathrm{d}s|y_i\Big) + o_p(1)\Big)(1+o_p(1)) \\
=& EZ_{2n}(\eta) \to \frac{1}{2}\eta^T D_1(\tau)\eta
\end{aligned}
$$

and

$$
\begin{aligned}
\mathrm{Var}(Z_{2n}^*(\eta)) =& \sum_{i=1}^{n}\mathrm{Var}(Z_{2ni}^*(\eta)) \le \sum_{i=1}^{n} E(Z_{2ni}^*(\eta))^2 \\
\le& \max|Z_{2ni}^*(\eta)|\sum_{i=1}^{n}E(Z_{2ni}^*(\eta)) \\
\le& \frac{1}{c^*\sqrt{n}}\max|\mathbf{x}_i^T\eta|\sum_{i=1}^{n}EZ_{2ni}^*(\eta) \to 0.
\end{aligned}
$$

Therefore, as same as the Proof of Theorem 2.1, we can have

$$Z_n^*(\eta) \to Z_0^*(\eta) = -\eta^T W + \frac{1}{2}\eta^T D_1(\tau)\eta.$$

As a result,

$$\widehat{\eta}_n^* = \sqrt{n}(\widehat{\boldsymbol{\beta}}^*(\tau) - \boldsymbol{\beta}(\tau)) = \mathrm{argmin} Z_n^*(\eta) \to \widehat{\eta}_0^* = \mathrm{argmin} Z_0^*(\eta) = D_1^{-1}(\tau)W.$$

We complete the proof of Theorem 2.2. $\square$

## Appendix

The contents of Tables 6 and 7 are shown here.

| $p_1$ | Method | $\tau=0.1$ | | | | $\tau=0.25$ | | | | $\tau=0.5$ | | | | $\tau=0.75$ | | | | $\tau=0.9$ | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | BIAS | SD | SE | COV | BIAS | SD | SE | COV | BIAS | SD | SE | COV | BIAS | SD | SE | COV | BIAS | SD | SE | COV |
| | *orq* | -0.4116 | 0.5387 | 0.4740 | 0.86 | -0.3809 | 0.4448 | 0.4092 | 0.87 | -0.2933 | 0.4088 | 0.4946 | 0.88 | -0.1531 | 0.3870 | 0.6070 | 0.94 | -0.0218 | 0.4465 | 0.5335 | 0.93 |
| | $wrq_{real}$ | 0.0863 | 0.5967 | 0.2915 | 0.89 | 0.0194 | 0.4857 | 0.3934 | 0.94 | 0.0071 | 0.4462 | 0.6110 | 0.95 | -0.0265 | 0.5542 | 0.7760 | 0.96 | -0.1620 | 0.8123 | 1.3433 | 0.88 |
| 0.1 | $wrq_{log}$ | 0.7699 | 0.8119 | 1.0611 | 0.79 | 0.5244 | 0.6680 | 0.7111 | 0.85 | 0.3552 | 0.6676 | 0.6996 | 0.91 | 0.0152 | 1.1913 | 0.6832 | 0.91 | -0.6476 | 1.6390 | 0.6722 | 0.86 |
| | $wrq_{loess}$ | 0.1771 | 0.6265 | 0.3236 | 0.88 | -0.0876 | 0.5569 | 0.3603 | 0.93 | -0.2293 | 0.6021 | 1.1606 | 0.94 | -0.3445 | 0.8116 | 1.4569 | 0.95 | -0.4147 | 0.9669 | 1.6799 | 0.90 |
| | $wrq_{nw}$ | 0.2719 | 0.5760 | 0.3242 | 0.87 | 0.1026 | 0.4824 | 0.3288 | 0.92 | 0.0950 | 0.4232 | 0.4023 | 0.93 | 0.0778 | 0.4508 | 0.6926 | 0.96 | -0.0407 | 0.6215 | 1.0793 | 0.93 |
| | *orq* | -0.1737 | 0.4435 | 0.3359 | 0.94 | -0.1031 | 0.3561 | 0.7171 | 0.94 | -0.0364 | 0.3318 | 0.2986 | 0.95 | 0.0516 | 0.3168 | 0.2116 | 0.94 | 0.1170 | 0.3670 | 0.4268 | 0.92 |
| | $wrq_{real}$ | 0.0494 | 0.4506 | 0.7235 | 0.94 | -0.0032 | 0.3600 | 0.4560 | 0.95 | 0.0012 | 0.3358 | 0.2663 | 0.97 | -0.0061 | 0.3693 | 0.3337 | 0.96 | -0.0563 | 0.5303 | 0.8051 | 0.94 |
| 0.2 | $wrq_{log}$ | 0.6295 | 0.5332 | 0.5942 | 0.81 | 0.6022 | 0.4872 | 0.4591 | 0.76 | 0.5439 | 0.4826 | 0.6102 | 0.78 | 0.4084 | 0.6280 | 0.6220 | 0.85 | 0.1367 | 0.7310 | 0.6019 | 0.92 |
| | $wrq_{loess}$ | 0.0822 | 0.4550 | 0.3801 | 0.91 | -0.0214 | 0.3517 | 0.3224 | 0.94 | -0.0365 | 0.3212 | 0.2020 | 0.96 | -0.0569 | 0.3406 | 0.2179 | 0.97 | -0.1035 | 0.4970 | 0.7372 | 0.96 |
| | $wrq_{nw}$ | 0.1058 | 0.4449 | 0.5115 | 0.92 | 0.0391 | 0.3451 | 0.2934 | 0.94 | 0.0618 | 0.3079 | 0.1868 | 0.95 | 0.0589 | 0.3300 | 0.2210 | 0.94 | 0.0309 | 0.4478 | 0.6560 | 0.95 |
| | *orq* | -0.0321 | 0.4090 | 0.2721 | 0.94 | 0.0089 | 0.3090 | 0.3385 | 0.95 | 0.0355 | 0.2677 | 0.2980 | 0.93 | 0.0844 | 0.2939 | 0.2749 | 0.93 | 0.1256 | 0.3487 | 0.3574 | 0.93 |
| | $wrq_{real}$ | 0.0520 | 0.4024 | 0.2626 | 0.93 | 0.0256 | 0.3210 | 0.3974 | 0.94 | -0.0136 | 0.3001 | 0.3431 | 0.95 | -0.0197 | 0.3545 | 0.4052 | 0.93 | -0.0395 | 0.4785 | 0.2735 | 0.93 |
| 0.3 | $wrq_{log}$ | 0.4498 | 0.3950 | 0.2552 | 0.81 | 0.4398 | 0.3474 | 0.2316 | 0.80 | 0.4419 | 0.3917 | 0.3390 | 0.82 | 0.4050 | 0.4851 | 0.3693 | 0.80 | 0.3038 | 0.5819 | 0.3339 | 0.88 |
| | $wrq_{loess}$ | 0.0794 | 0.4094 | 0.2935 | 0.90 | 0.0391 | 0.3224 | 0.3479 | 0.94 | -0.0267 | 0.2717 | 0.2895 | 0.97 | -0.0378 | 0.3129 | 0.3640 | 0.95 | -0.0619 | 0.4266 | 0.3121 | 0.96 |
| | $wrq_{nw}$ | 0.0819 | 0.3919 | 0.2438 | 0.91 | 0.0600 | 0.3112 | 0.3389 | 0.94 | 0.0178 | 0.2760 | 0.2987 | 0.96 | 0.0162 | 0.3197 | 0.3458 | 0.94 | 0.0134 | 0.4134 | 0.3074 | 0.94 |
| | *orq* | -0.0153 | 0.3661 | 0.5811 | 0.93 | 0.0195 | 0.2782 | 0.3246 | 0.95 | 0.0716 | 0.2589 | 0.2402 | 0.93 | 0.1195 | 0.2671 | 0.4035 | 0.93 | 0.1480 | 0.3362 | 0.1632 | 0.92 |
| | $wrq_{real}$ | 0.0165 | 0.3567 | 0.4728 | 0.93 | 0.0107 | 0.2764 | 0.3743 | 0.96 | 0.0163 | 0.2786 | 0.3127 | 0.95 | 0.0168 | 0.3032 | 0.4040 | 0.95 | -0.0118 | 0.4213 | 0.2500 | 0.93 |
| 0.4 | $wrq_{log}$ | 0.2934 | 0.3565 | 0.3560 | 0.88 | 0.3341 | 0.3020 | 0.4285 | 0.83 | 0.3484 | 0.3038 | 0.3997 | 0.81 | 0.3607 | 0.3695 | 0.2526 | 0.85 | 0.3229 | 0.5017 | 0.1373 | 0.83 |
| | $wrq_{loess}$ | 0.0295 | 0.3511 | 0.4507 | 0.94 | 0.0155 | 0.2611 | 0.3238 | 0.96 | 0.0169 | 0.2445 | 0.2685 | 0.96 | 0.0155 | 0.2882 | 0.3535 | 0.97 | -0.0166 | 0.3861 | 0.1264 | 0.95 |
| | $wrq_{nw}$ | 0.0272 | 0.3470 | 0.4239 | 0.94 | 0.0293 | 0.2605 | 0.3302 | 0.96 | 0.0435 | 0.2546 | 0.2776 | 0.94 | 0.0453 | 0.2765 | 0.3525 | 0.95 | 0.0387 | 0.3789 | 0.1381 | 0.94 |
| | *orq* | 0.0281 | 0.3594 | 0.4041 | 0.94 | 0.0352 | 0.2702 | 0.3296 | 0.95 | 0.0626 | 0.2397 | 0.2685 | 0.93 | 0.0743 | 0.2698 | 0.2613 | 0.92 | 0.1069 | 0.3193 | 0.3161 | 0.93 |
| | $wrq_{real}$ | 0.0305 | 0.3411 | 0.3727 | 0.94 | 0.0133 | 0.2633 | 0.3869 | 0.95 | -0.0016 | 0.2495 | 0.3047 | 0.95 | -0.0232 | 0.2969 | 0.2790 | 0.94 | -0.0231 | 0.3755 | 0.3300 | 0.95 |
| 0.5 | $wrq_{log}$ | 0.2239 | 0.3289 | 0.3314 | 0.91 | 0.2278 | 0.2677 | 0.3044 | 0.87 | 0.2377 | 0.2556 | 0.2670 | 0.88 | 0.2368 | 0.3262 | 0.3626 | 0.88 | 0.2382 | 0.4075 | 0.3242 | 0.90 |
| | $wrq_{loess}$ | 0.0572 | 0.3516 | 0.3868 | 0.94 | 0.0160 | 0.2564 | 0.3496 | 0.95 | -0.0061 | 0.2440 | 0.2956 | 0.95 | -0.0269 | 0.2724 | 0.2626 | 0.94 | -0.0315 | 0.3512 | 0.3390 | 0.95 |
| | $wrq_{nw}$ | 0.0527 | 0.3450 | 0.3805 | 0.94 | 0.0239 | 0.2550 | 0.3655 | 0.94 | 0.0122 | 0.2428 | 0.2878 | 0.95 | 0.0003 | 0.2783 | 0.2828 | 0.94 | 0.0064 | 0.3509 | 0.3326 | 0.94 |
| | *orq* | 0.0239 | 0.3324 | 0.2795 | 0.97 | 0.0515 | 0.2772 | 0.3081 | 0.91 | 0.0693 | 0.2493 | 0.2115 | 0.92 | 0.0925 | 0.2520 | 0.2359 | 0.92 | 0.1150 | 0.3105 | 0.3361 | 0.92 |
| | $wrq_{real}$ | 0.0148 | 0.3224 | 0.2390 | 0.96 | 0.0179 | 0.2706 | 0.2773 | 0.92 | 0.0112 | 0.2575 | 0.2239 | 0.92 | 0.0134 | 0.2769 | 0.2749 | 0.93 | 0.0006 | 0.3451 | 0.3425 | 0.95 |
| 0.6 | $wrq_{log}$ | 0.1453 | 0.3133 | 0.2188 | 0.94 | 0.1640 | 0.2666 | 0.2668 | 0.88 | 0.1789 | 0.2595 | 0.2110 | 0.89 | 0.1903 | 0.2758 | 0.2492 | 0.89 | 0.2104 | 0.3583 | 0.2904 | 0.90 |
| | $wrq_{loess}$ | 0.0239 | 0.3181 | 0.2076 | 0.97 | 0.0171 | 0.2707 | 0.2684 | 0.91 | 0.0078 | 0.2557 | 0.2181 | 0.92 | 0.0007 | 0.2745 | 0.2735 | 0.93 | -0.0004 | 0.3443 | 0.3483 | 0.94 |
| | $wrq_{nw}$ | 0.0207 | 0.3144 | 0.2367 | 0.97 | 0.0246 | 0.2708 | 0.2634 | 0.91 | 0.0194 | 0.2516 | 0.2114 | 0.92 | 0.0231 | 0.2733 | 0.2631 | 0.92 | 0.0226 | 0.3397 | 0.3238 | 0.94 |
| | *orq* | 0.0408 | 0.3183 | 0.7210 | 0.94 | 0.0524 | 0.2637 | 0.4011 | 0.93 | 0.0467 | 0.2475 | 0.1589 | 0.93 | 0.0468 | 0.2572 | 0.2093 | 0.94 | 0.0810 | 0.3117 | 0.4332 | 0.94 |
| | $wrq_{real}$ | 0.0210 | 0.3086 | 0.6848 | 0.95 | 0.0170 | 0.2657 | 0.3935 | 0.93 | -0.0031 | 0.2490 | 0.1540 | 0.94 | -0.0114 | 0.2659 | 0.2118 | 0.95 | -0.0046 | 0.3221 | 0.4259 | 0.95 |
| 0.7 | $wrq_{log}$ | 0.1034 | 0.2989 | 0.6352 | 0.94 | 0.1117 | 0.2559 | 0.3637 | 0.89 | 0.1177 | 0.2429 | 0.1631 | 0.91 | 0.1175 | 0.2689 | 0.1926 | 0.92 | 0.1409 | 0.3230 | 0.5317 | 0.95 |
| | $wrq_{loess}$ | 0.0328 | 0.2992 | 0.6429 | 0.96 | 0.0183 | 0.2584 | 0.3768 | 0.93 | -0.0060 | 0.2404 | 0.1546 | 0.94 | -0.0142 | 0.2593 | 0.2341 | 0.95 | -0.0104 | 0.3071 | 0.4236 | 0.96 |
| | $wrq_{nw}$ | 0.0260 | 0.3000 | 0.6567 | 0.96 | 0.0208 | 0.2578 | 0.3733 | 0.93 | 0.0026 | 0.2408 | 0.1595 | 0.94 | -0.0034 | 0.2605 | 0.2198 | 0.95 | 0.0102 | 0.3072 | 0.4210 | 0.96 |
| | *orq* | -0.0083 | 0.3265 | 0.2385 | 0.94 | 0.0168 | 0.2683 | 0.2279 | 0.93 | 0.0249 | 0.2483 | 0.2080 | 0.91 | 0.0243 | 0.2614 | 0.2565 | 0.93 | 0.0339 | 0.3120 | 0.4291 | 0.93 |
| | $wrq_{real}$ | -0.0136 | 0.3302 | 0.2331 | 0.94 | -0.0068 | 0.2661 | 0.2259 | 0.93 | -0.0081 | 0.2493 | 0.2206 | 0.92 | -0.0153 | 0.2665 | 0.2539 | 0.93 | -0.0192 | 0.3271 | 0.4449 | 0.93 |
| 0.8 | $wrq_{log}$ | 0.0323 | 0.3279 | 0.2371 | 0.94 | 0.0501 | 0.2646 | 0.2291 | 0.93 | 0.0549 | 0.2431 | 0.2124 | 0.93 | 0.0626 | 0.2650 | 0.2595 | 0.92 | 0.0723 | 0.3228 | 0.4699 | 0.92 |
| | $wrq_{loess}$ | -0.0109 | 0.3248 | 0.2290 | 0.94 | -0.0048 | 0.2623 | 0.2325 | 0.93 | -0.0110 | 0.2432 | 0.2152 | 0.92 | -0.0143 | 0.2604 | 0.2568 | 0.93 | -0.0144 | 0.3218 | 0.4228 | 0.92 |
| | $wrq_{nw}$ | -0.0126 | 0.3240 | 0.2331 | 0.94 | -0.0019 | 0.2622 | 0.2206 | 0.93 | -0.0059 | 0.2458 | 0.2144 | 0.92 | -0.0076 | 0.2616 | 0.2619 | 0.93 | -0.0084 | 0.3199 | 0.4214 | 0.93 |
| | *orq* | 0.0162 | 0.3113 | 0.5630 | 0.92 | 0.0192 | 0.2418 | 0.4098 | 0.95 | 0.0159 | 0.2242 | 0.2297 | 0.93 | 0.0230 | 0.2491 | 0.2346 | 0.93 | 0.0386 | 0.3132 | 0.3505 | 0.93 |
| | $wrq_{real}$ | 0.0052 | 0.3157 | 0.5531 | 0.92 | 0.0100 | 0.2424 | 0.4111 | 0.95 | 0.0003 | 0.2251 | 0.2268 | 0.93 | 0.0029 | 0.2556 | 0.2368 | 0.94 | 0.0155 | 0.3171 | 0.3510 | 0.94 |
| 0.9 | $wrq_{log}$ | 0.0279 | 0.3114 | 0.5617 | 0.92 | 0.0330 | 0.2402 | 0.4142 | 0.95 | 0.0301 | 0.2250 | 0.2246 | 0.93 | 0.0371 | 0.2566 | 0.2284 | 0.92 | 0.0518 | 0.3121 | 0.3557 | 0.93 |
| | $wrq_{loess}$ | 0.0115 | 0.3128 | 0.5248 | 0.93 | 0.0122 | 0.2419 | 0.4011 | 0.95 | 0.0012 | 0.2233 | 0.2280 | 0.94 | 0.0034 | 0.2504 | 0.2279 | 0.94 | 0.0208 | 0.3166 | 0.3364 | 0.94 |
| | $wrq_{nw}$ | 0.0076 | 0.3138 | 0.5475 | 0.92 | 0.0131 | 0.2410 | 0.4076 | 0.95 | 0.0019 | 0.2243 | 0.2268 | 0.94 | 0.0049 | 0.2529 | 0.2319 | 0.93 | 0.0216 | 0.3149 | 0.3486 | 0.94 |

Table 6 The SD, SE and COV of intercept

| $p_1$ | Method | $\tau=0.1$ | | | | $\tau=0.25$ | | | | $\tau=0.5$ | | | | $\tau=0.75$ | | | | $\tau=0.9$ | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | BIAS | SD | SE | COV | BIAS | SD | SE | COV | BIAS | SD | SE | COV | BIAS | SD | SE | COV | BIAS | SD | SE | COV |
| | *orq* | 0.1831 | 0.1839 | 0.1824 | 0.83 | 0.1792 | 0.1553 | 0.1370 | 0.81 | 0.1546 | 0.1450 | 0.1675 | 0.84 | 0.1129 | 0.1394 | 0.2379 | 0.90 | 0.0703 | 0.1609 | 0.2103 | 0.89 |
| | $wrq_{real}$ | -0.0291 | 0.1994 | 0.0890 | 0.88 | -0.0064 | 0.1636 | 0.1242 | 0.92 | -0.0023 | 0.1478 | 0.2363 | 0.95 | 0.0101 | 0.1761 | 0.2453 | 0.94 | 0.0535 | 0.2372 | 0.3148 | 0.87 |
| 0.1 | $wrq_{log}$ | -0.2646 | 0.2469 | 0.3028 | 0.75 | -0.1936 | 0.2120 | 0.2567 | 0.81 | -0.1468 | 0.2132 | 0.2253 | 0.89 | -0.0544 | 0.3202 | 0.186 | 0.91 | 0.1488 | 0.4145 | 0.1888 | 0.85 |
| | $wrq_{loess}$ | -0.0704 | 0.2088 | 0.1072 | 0.84 | 0.0165 | 0.1854 | 0.0998 | 0.91 | 0.0605 | 0.1803 | 0.3072 | 0.95 | 0.0967 | 0.2189 | 0.361 | 0.93 | 0.1316 | 0.2611 | 0.4371 | 0.87 |
| | $wrq_{nw}$ | -0.0869 | 0.1890 | 0.1028 | 0.85 | -0.0294 | 0.1638 | 0.1001 | 0.91 | -0.0275 | 0.1467 | 0.1521 | 0.94 | -0.0184 | 0.1557 | 0.2572 | 0.95 | 0.0243 | 0.1968 | 0.2803 | 0.90 |
| | *orq* | 0.0879 | 0.1471 | 0.1139 | 0.91 | 0.0693 | 0.1180 | 0.2463 | 0.91 | 0.0494 | 0.1121 | 0.1004 | 0.93 | 0.0222 | 0.1047 | 0.0752 | 0.94 | 0.0034 | 0.1223 | 0.1631 | 0.93 |
| | $wrq_{real}$ | -0.0177 | 0.1463 | 0.2662 | 0.91 | 0.0011 | 0.1162 | 0.1511 | 0.94 | -0.0012 | 0.1071 | 0.0648 | 0.96 | 0.0001 | 0.114 | 0.1066 | 0.96 | 0.0163 | 0.156 | 0.2524 | 0.94 |
| 0.2 | $wrq_{log}$ | -0.2125 | 0.1614 | 0.1402 | 0.76 | -0.2094 | 0.1528 | 0.1318 | 0.71 | -0.1969 | 0.1577 | 0.2021 | 0.74 | -0.1585 | 0.1957 | 0.2289 | 0.81 | -0.068 | 0.2188 | 0.2141 | 0.91 |
| | $wrq_{loess}$ | -0.0276 | 0.1474 | 0.1382 | 0.89 | 0.0066 | 0.1150 | 0.0790 | 0.95 | 0.0103 | 0.1040 | 0.0502 | 0.95 | 0.0154 | 0.1055 | 0.0583 | 0.96 | 0.0315 | 0.1447 | 0.227 | 0.95 |
| | $wrq_{nw}$ | -0.0331 | 0.1433 | 0.1848 | 0.89 | -0.0106 | 0.1119 | 0.0721 | 0.93 | -0.0177 | 0.1001 | 0.0497 | 0.95 | -0.0176 | 0.1064 | 0.0624 | 0.95 | -0.0081 | 0.1419 | 0.226 | 0.95 |
| | *orq* | 0.0338 | 0.1288 | 0.0929 | 0.94 | 0.0231 | 0.0989 | 0.0989 | 0.94 | 0.0157 | 0.0843 | 0.0914 | 0.93 | 0.0024 | 0.0919 | 0.0801 | 0.95 | -0.0083 | 0.1091 | 0.0987 | 0.95 |
| | $wrq_{real}$ | -0.0183 | 0.1236 | 0.0917 | 0.92 | -0.0090 | 0.1000 | 0.1095 | 0.92 | 0.0039 | 0.0902 | 0.1034 | 0.95 | 0.0053 | 0.1048 | 0.1041 | 0.94 | 0.0133 | 0.1386 | 0.078 | 0.93 |
| 0.3 | $wrq_{log}$ | -0.1517 | 0.1172 | 0.0681 | 0.78 | -0.1495 | 0.1075 | 0.0661 | 0.75 | -0.1541 | 0.1278 | 0.1018 | 0.78 | -0.1464 | 0.1613 | 0.106 | 0.78 | -0.1144 | 0.1929 | 0.1091 | 0.84 |
| | $wrq_{loess}$ | -0.0260 | 0.1252 | 0.1025 | 0.91 | -0.0130 | 0.1004 | 0.0985 | 0.94 | 0.0081 | 0.0815 | 0.0864 | 0.97 | 0.0112 | 0.0929 | 0.0899 | 0.95 | 0.0217 | 0.1235 | 0.0944 | 0.95 |
| | $wrq_{nw}$ | -0.0271 | 0.1195 | 0.0839 | 0.91 | -0.0187 | 0.0967 | 0.0959 | 0.93 | -0.0044 | 0.0843 | 0.0902 | 0.95 | -0.0043 | 0.0969 | 0.0845 | 0.94 | -0.0005 | 0.1244 | 0.0934 | 0.95 |
| | *orq* | 0.0226 | 0.1124 | 0.1845 | 0.95 | 0.0139 | 0.0857 | 0.0814 | 0.96 | -0.0011 | 0.0819 | 0.0821 | 0.94 | -0.0146 | 0.084 | 0.1404 | 0.94 | -0.0222 | 0.1054 | 0.0521 | 0.92 |
| | $wrq_{real}$ | -0.0067 | 0.1084 | 0.1421 | 0.93 | -0.0035 | 0.0847 | 0.0930 | 0.95 | -0.0045 | 0.0861 | 0.0956 | 0.94 | -0.0049 | 0.0911 | 0.1341 | 0.94 | 0.0045 | 0.1239 | 0.0623 | 0.93 |
| 0.4 | $wrq_{log}$ | -0.0985 | 0.1049 | 0.0866 | 0.85 | -0.1124 | 0.0933 | 0.1148 | 0.79 | -0.1177 | 0.0955 | 0.1276 | 0.75 | -0.1244 | 0.1219 | 0.0752 | 0.83 | -0.1139 | 0.1694 | 0.0414 | 0.81 |
| | $wrq_{loess}$ | -0.0097 | 0.1053 | 0.1377 | 0.93 | -0.0047 | 0.0802 | 0.0790 | 0.95 | -0.0048 | 0.0752 | 0.0848 | 0.94 | -0.0045 | 0.0882 | 0.1258 | 0.95 | 0.0058 | 0.1155 | 0.0431 | 0.94 |
| | $wrq_{nw}$ | -0.0093 | 0.1042 | 0.1239 | 0.93 | -0.0088 | 0.0799 | 0.0810 | 0.95 | -0.0124 | 0.0785 | 0.0871 | 0.93 | -0.0126 | 0.0836 | 0.1212 | 0.94 | -0.0105 | 0.1166 | 0.0434 | 0.94 |
| | *orq* | 0.0057 | 0.1075 | 0.1165 | 0.95 | 0.0032 | 0.0813 | 0.0978 | 0.95 | -0.0037 | 0.0721 | 0.0768 | 0.94 | -0.007 | 0.0803 | 0.0794 | 0.94 | -0.017 | 0.0967 | 0.1095 | 0.95 |
| | $wrq_{real}$ | -0.0103 | 0.1023 | 0.1063 | 0.95 | -0.0054 | 0.0778 | 0.1089 | 0.95 | 0.0000 | 0.0736 | 0.0844 | 0.94 | 0.0065 | 0.086 | 0.0858 | 0.95 | 0.0065 | 0.1106 | 0.1041 | 0.95 |
| 0.5 | $wrq_{log}$ | -0.0724 | 0.0941 | 0.0899 | 0.87 | -0.0763 | 0.0791 | 0.0850 | 0.85 | -0.0797 | 0.0762 | 0.0715 | 0.86 | -0.0819 | 0.0994 | 0.1119 | 0.87 | -0.0848 | 0.1293 | 0.1008 | 0.90 |
| | $wrq_{loess}$ | -0.0173 | 0.1043 | 0.1093 | 0.94 | -0.0060 | 0.0751 | 0.0981 | 0.95 | 0.0019 | 0.0710 | 0.0824 | 0.95 | 0.0079 | 0.0773 | 0.0828 | 0.95 | 0.0096 | 0.1018 | 0.1116 | 0.95 |
| | $wrq_{nw}$ | -0.0162 | 0.1017 | 0.1068 | 0.94 | -0.0083 | 0.0752 | 0.1027 | 0.94 | -0.0032 | 0.0708 | 0.0801 | 0.94 | -0.0001 | 0.0802 | 0.0891 | 0.94 | -0.0017 | 0.1032 | 0.1098 | 0.95 |
| | *orq* | 0.0047 | 0.0999 | 0.0943 | 0.96 | -0.0033 | 0.0811 | 0.0858 | 0.92 | -0.0080 | 0.0731 | 0.0685 | 0.93 | -0.015 | 0.0736 | 0.0608 | 0.94 | -0.0208 | 0.0907 | 0.0867 | 0.94 |
| | $wrq_{real}$ | -0.0028 | 0.0959 | 0.0757 | 0.95 | -0.0045 | 0.0785 | 0.0752 | 0.93 | -0.0016 | 0.0746 | 0.0722 | 0.93 | -0.0029 | 0.0793 | 0.0632 | 0.94 | 0.0013 | 0.0973 | 0.0938 | 0.95 |
| 0.6 | $wrq_{log}$ | -0.0453 | 0.0922 | 0.0687 | 0.92 | -0.0517 | 0.0760 | 0.0705 | 0.87 | -0.0565 | 0.0767 | 0.0687 | 0.88 | -0.0614 | 0.0812 | 0.0618 | 0.88 | -0.0689 | 0.1091 | 0.0747 | 0.90 |
| | $wrq_{loess}$ | -0.0051 | 0.0939 | 0.0644 | 0.95 | -0.0037 | 0.0781 | 0.0717 | 0.92 | 0.0000 | 0.0737 | 0.0703 | 0.93 | 0.0016 | 0.078 | 0.0672 | 0.94 | 0.0019 | 0.0976 | 0.0912 | 0.94 |
| | $wrq_{nw}$ | -0.0047 | 0.0930 | 0.0733 | 0.94 | -0.0061 | 0.0784 | 0.0719 | 0.92 | -0.0035 | 0.0727 | 0.0680 | 0.93 | -0.0053 | 0.0783 | 0.0631 | 0.93 | -0.005 | 0.097 | 0.0839 | 0.94 |
| | *orq* | -0.0057 | 0.0925 | 0.1983 | 0.95 | -0.0089 | 0.0750 | 0.1105 | 0.94 | -0.0068 | 0.0706 | 0.0494 | 0.93 | -0.0078 | 0.0728 | 0.0615 | 0.96 | -0.0188 | 0.0914 | 0.1125 | 0.94 |
| | $wrq_{real}$ | -0.0072 | 0.0890 | 0.1897 | 0.95 | -0.0060 | 0.0749 | 0.1050 | 0.94 | 0.0002 | 0.0704 | 0.0466 | 0.94 | 0.002 | 0.0747 | 0.0635 | 0.95 | -0.0015 | 0.0935 | 0.1098 | 0.95 |
| 0.7 | $wrq_{log}$ | -0.0332 | 0.0847 | 0.1711 | 0.93 | -0.0360 | 0.0721 | 0.0902 | 0.90 | -0.0384 | 0.0694 | 0.0462 | 0.89 | -0.0406 | 0.0771 | 0.0565 | 0.92 | -0.049 | 0.0966 | 0.1509 | 0.94 |
| | $wrq_{loess}$ | -0.0100 | 0.0857 | 0.1756 | 0.96 | -0.0059 | 0.0729 | 0.0978 | 0.94 | 0.0016 | 0.0675 | 0.0460 | 0.94 | 0.0031 | 0.0723 | 0.0705 | 0.95 | 0.0006 | 0.0888 | 0.108 | 0.95 |
| | $wrq_{nw}$ | -0.0085 | 0.0860 | 0.1795 | 0.96 | -0.0066 | 0.0728 | 0.0969 | 0.93 | -0.0011 | 0.0677 | 0.0470 | 0.94 | -0.0001 | 0.0729 | 0.0662 | 0.95 | -0.0058 | 0.0893 | 0.1076 | 0.96 |
| | *orq* | 0.0081 | 0.0919 | 0.0622 | 0.94 | 0.0015 | 0.0758 | 0.0670 | 0.93 | -0.0010 | 0.0696 | 0.0608 | 0.92 | -0.0012 | 0.0734 | 0.0746 | 0.93 | -0.003 | 0.0875 | 0.1102 | 0.95 |
| | $wrq_{real}$ | 0.0054 | 0.0924 | 0.0599 | 0.94 | 0.0036 | 0.0748 | 0.0655 | 0.94 | 0.0040 | 0.0699 | 0.0656 | 0.91 | 0.0059 | 0.0744 | 0.0744 | 0.93 | 0.0075 | 0.0911 | 0.1133 | 0.94 |
| 0.8 | $wrq_{log}$ | -0.0089 | 0.0913 | 0.0616 | 0.94 | -0.0141 | 0.0740 | 0.0669 | 0.93 | -0.0162 | 0.0677 | 0.0638 | 0.93 | -0.0192 | 0.0746 | 0.0753 | 0.91 | -0.0222 | 0.0916 | 0.1246 | 0.95 |
| | $wrq_{loess}$ | 0.0047 | 0.0907 | 0.0596 | 0.94 | 0.0032 | 0.0735 | 0.0683 | 0.93 | 0.0047 | 0.0681 | 0.0631 | 0.93 | 0.0056 | 0.0724 | 0.0748 | 0.93 | 0.0056 | 0.0895 | 0.1057 | 0.93 |
| | $wrq_{nw}$ | 0.0051 | 0.0905 | 0.0611 | 0.94 | 0.0023 | 0.0734 | 0.0648 | 0.94 | 0.0034 | 0.0688 | 0.0633 | 0.92 | 0.0037 | 0.073 | 0.0771 | 0.93 | 0.0045 | 0.0891 | 0.1041 | 0.94 |
| | *orq* | -0.0024 | 0.0869 | 0.1605 | 0.93 | -0.0033 | 0.0676 | 0.1119 | 0.93 | -0.0027 | 0.0628 | 0.0646 | 0.93 | -0.0042 | 0.0698 | 0.0667 | 0.93 | -0.0083 | 0.0882 | 0.0955 | 0.94 |
| | $wrq_{real}$ | -0.0012 | 0.0878 | 0.1585 | 0.92 | -0.0025 | 0.0676 | 0.1120 | 0.95 | -0.0002 | 0.0629 | 0.0638 | 0.93 | -0.0005 | 0.0712 | 0.0666 | 0.93 | -0.0038 | 0.0892 | 0.0962 | 0.95 |
| 0.9 | $wrq_{log}$ | -0.0082 | 0.0866 | 0.1599 | 0.93 | -0.0095 | 0.0670 | 0.1131 | 0.94 | -0.0096 | 0.0630 | 0.0626 | 0.94 | -0.0115 | 0.0717 | 0.0653 | 0.92 | -0.0154 | 0.0877 | 0.0982 | 0.94 |
| | $wrq_{loess}$ | -0.0027 | 0.0872 | 0.1482 | 0.93 | -0.0030 | 0.0672 | 0.1099 | 0.94 | -0.0003 | 0.0620 | 0.0647 | 0.94 | -0.0006 | 0.0693 | 0.0645 | 0.93 | -0.0054 | 0.0888 | 0.0923 | 0.94 |
| | $wrq_{nw}$ | -0.0017 | 0.0876 | 0.1551 | 0.93 | -0.0032 | 0.0671 | 0.1120 | 0.94 | -0.0005 | 0.0625 | 0.0641 | 0.94 | -0.0008 | 0.0703 | 0.0655 | 0.93 | -0.0055 | 0.0885 | 0.0961 | 0.94 |

Table 7 The BIAS, SD, SE and COV of slope

# References


[1] R. KOENKER, G. BASSETT. *Regression quantiles*. Econometrica, 1978, **46**: 33–50.

[2] T. J. COLE. *Fitting smoothed centile curves to reference data*. J. R. Stat. Soc. B., 1988, **151**: 385–418.

[3] Keming YU, M.C. JONES. *Local linear quantile regression*. J. Amer. Statist. Assoc., 1998, **98**(441): 228–237.

[4] P. HALL, R. C. L. WOLFF, Qiwei YAO. *Methods for estimating a conditional distribution*. J. Amer. Statist. Assoc., 1999, **94**(445): 154–163.

[5] T. J. COLE, P. J. GREEN. *Smoothing reference centile curves: the LMS nethod and penalized likelihood*. Stat. Med., 1992, **11**: 1305–1319.

[6] A. KONG, J. S. LIU, W. H. WONG. *Sequential imputations and bayesian missing data problems*. J. Amer. Statist. Assoc., 1994, **89**: 278–288.

[7] J. G. IBRAHIM, Minghui CHEN, S. R. LIPSITZ, et al. *Missing-data methods for generalized linear models: a comparative review*. J. Amer. Statist. Assoc., 2005, **100**(469): 332–346.

[8] J. M. ROBINS, A. ROTNITZKY, Lueping ZHAO. *Estimation of regression coefficients when some regressors are not always observed*. J. Amer. Statist. Assoc., 1994, **89**(427): 846–866.

[9] C. Y. WANG, Suojin WANG, Lueping ZHAO, et al. *Weighted semiparametric estimation in regression analysis with missing covariate data*. J. Amer. Statist. Assoc., 1997, **92**(438): 512–525.

[10] Qihua WANG, O. LINTON, W. HÄRDLE. *Semiparametric regression analysis with missing response at random*. J. Amer. Statist. Assoc., 2004, **99**(466): 334–345.

[11] J. YOON. *Quantile Regression Analysis with Missing Response, with Applications to Inequality Measures and Data Combination*. Social Science Electronic Publishing, 2017.

[12] Ying WEI, Yanyuan MA, R. J. CARROLL. *Multiple imputation in quantile regression*. Biometrika, 2012, **99**(2): 423–438.

[13] B. SHERWOOD, Lan WANG, Xiaohua ZHOU. *Weighted quantile regression for analyzing health care cost data with missing covariates*. Stat. Med., 2013, **32**(28): 4967–4979.

[14] Ying WEI, Yunwen YANG. *Quantile regression with covariates missing at random*. Statist. Sinica, 2014, **24**(3): 1277–1299.

[15] Lingnan TAI, Chunyu WANG, Maozai TIAN. *Inverse probability multiple weighted quantile regression estimation and its application with missing data*. Statist. Res., 2018, **35**(9): 115–128.

[16] Peisong HAN, Linglong KONG, Jiwei ZHAO, et al. *A general framework for quantile estimation with incomplete data*. J. R. Stat. Soc. Ser. B. Stat. Methodol., 2019, **81**(2): 305–333.

[17] D. G. HORVITZ, D. J. THOMPSON. *A generalization of sampling without replacement from a finite universe*. J. Amer. Statist. Assoc., 1952, **47**: 663–685.

[18] Xuerong CHEN, A. T. K. WAN, Yong ZHOU. *Efficient quantile regression analysis with missing observations*. J. Amer. Statist. Assoc., 2015, **110**(510): 723–741.

[19] W. HÄRDLE, M. MÜLLER, S. SPERLICH, et al. *Nonparametric and Semiparametric Models*. Springer-Verlag, New York, 2004.

[20] Xu GUO, Wangli XU. *Goodness-of-fit for general linear models with covariates missed at random*. J. Statist. Plann. Inference, 2012, **142**(7): 2047–2058.