

Article

Efficient Estimation for the Derivative of Nonparametric Function by Optimally Combining Quantile Information

Xiaoshuang Zhou *, Xiulian Gao, Yukun Zhang, Xiuling Yin and Yanfeng Shen

School of Mathematics and Big Data, Dezhou University, Dezhou 253023, China; gaoxiulian@dzu.edu.cn (X.G.); zhangyukun@dzu.edu.cn (Y.Z.); yinxiuling@dzu.edu.cn (X.Y.); shenyanfeng@dzu.edu.cn (Y.S.)

* Correspondence: zhouxiaoshuang@dzu.edu.cn; Tel.: +86-1526-690-0327

Abstract: In this article, we focus on the efficient estimators of the derivative of the nonparametric function in the nonparametric quantile regression model. We develop two ways of combining quantile regression information to derive the estimators. One is the weighted composite quantile regression estimator based on the quantile weighted loss function; the other is the weighted quantile average estimator based on the weighted average of quantile regression estimators at a single quantile. Furthermore, by minimizing the asymptotic variance, the optimal weight vector is computed, and consequently, the optimal estimator is obtained. Furthermore, we conduct some simulations to evaluate the performance of our proposed estimators under different symmetric error distributions. Simulation studies further illustrate that both estimators work better than the local linear least square estimator for all the symmetric errors considered except the normal error, and the weighted quantile average estimator performs better than the weighted composite quantile regression estimator in most situations.



Citation: Zhou, X.; Gao, X.; Zhang, Y.; Yin, X.; Shen, Y. Efficient Estimation for the Derivative of Nonparametric Function by Optimally Combining Quantile Information. *Symmetry* **2021**, *13*, 2387. <https://doi.org/10.3390/sym13122387>

Academic Editors: Célia Nunes, Miguel Fonseca and Elsa Moreira

Received: 11 November 2021
Accepted: 4 December 2021
Published: 10 December 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: nonparametric regression model; the derivative of nonparametric function; composite quantile regression; weighted quantile average estimator

1. Introduction

Consider the following general nonparametric regression model:

$$Y = m(T) + \sigma(T)\varepsilon, \quad (1)$$

where $Y \in R$ is the response variable, and T is a scalar covariate independent of ε , where ε is the random error. $m(t) = E(Y|T = t)$ is an unknown smooth function, and $\sigma(T)$ is a non-negative function which represents the standard deviation of the model error ε . $E(\varepsilon) = 0$ and $\text{var}(\varepsilon) = 1$. As we all know, lots of estimation methods, including the kernel regression method, the spline smoothing method, the orthogonal series approximation method and the local polynomial method, have been investigated for estimating the nonparametric regression function $m(t)$. The relevant works can be referred but not limited to [1–5]. However, we need to estimate $m'(t)$, that is, the derivative of $m(t)$, in many situations. Therefore, in this article, we intend to derive the efficient estimators of $m'(t)$.

Since the seminal work of Koenker and Bassett [6], quantile regression has been deeply studied in the literature and applied in econometrics, biomedicine and other fields. It referred to Koenker [7] for comprehensive treatments. Based on the quantile regression, Zou and Yuan [8] proposed composite quantile regression (abbreviated as CQR) for linear models. It assumes that the regression coefficients are the same across different quantile regression models by combining the strength across multiple quantile regression models. The advantage of the CQR method is that it can improve the relative efficiency of the relevant estimators significantly and is usually used in regression models with non-normal error distribution. Kai et al. [9] introduced the local CQR for the general nonparametric regression model. As a kind of nonlinear smoother, the local CQR method does not require

finite error variance and hence can work well even when the error distribution has infinite variance. Meanwhile, the local CQR method can significantly improve the estimation efficiency of local linear least squares for some cases. Kai et al. [10] applied the local CQR method to a semiparametric varying coefficient partially linear model, and the results showed that the local CQR method outperformed both the least squares and the single quantile regression. Jiang et al. [11] applied the two step CQR method to the single index model and established its efficiency. Ning and Tang [12] considered the estimation and test issues for CQR method with missing covariates. Recently, many researchers applied the COR method to other various models under different data cases. It can be referred but not limited to [13–19].

To make full use of quantile information, one may consider combining information over different quantiles by the criterion function of the estimation procedure or combining information based on estimators at different quantiles. In this article, we argue that, although a composite quantile regression estimator based on the aggregated criterion function may outperform the ordinary least square estimator for some non-normal distributions, it is noted that simple averaging usually does not make full use of all the information of quantiles. Information at different quantiles are correlated, and improperly using multiple quantile information may even reduce the efficiency. Roughly speaking, simple average delivers good estimators when the error distribution is close to normal. In fact, for nonparametric regression, a simple averaging-based composite quantile regression estimator is asymptotically equivalent to the local least squares estimator. However, the main purpose of combining quantile information is to improve efficiency when the error distribution is not normal and the ordinary least square method does not work well. It is therefore important to combine quantile information appropriately to achieve efficiency. In this paper, we mainly study optimal combination of quantile regression information for estimating the derivative of nonparametric function. As stated above, we intend to propose and develop two ways of combining quantile information. One is the weighted local CQR estimator based on weighted quantile loss functions, and the other is the weighted quantile average estimator based on a weighted average of quantile regression estimators at single quantiles. Our proposed estimators inherit many advantages. Both the theoretical results and simulation studies further illustrate that both the weighted local CQR estimator and the weighted quantile average estimator work better than the common local linear least square estimator for all the symmetric errors considered except the normal error, and the weighted quantile average estimator performs better than the weighted composite quantile regression estimator in most situations.

The rest of the paper is organized as follows. The weighted local CQR estimator and the weighted quantile average estimator are proposed in Section 2. Meanwhile, the main theoretical results including asymptotic normality and the optimal weights are presented. In Section 3, the asymptotic relative efficiency of the weighted local CQR estimator and the weighted quantile average estimator are compared. The feasibility of our proposed method is verified by random simulations in Section 4. The technical proofs of the theoretical results are presented in Section 5. Conclusions are added in Section 2.

2. Methodology

Firstly, we give some conditions and notations, which are required in our subsequent discussions. Let $f(\cdot)$ and $F(\cdot)$ be the density function and cumulative distribution function of the error, respectively. Denote $f_T(\cdot)$ as the marginal density function of the covariate T . Let \mathcal{T} be the σ -field generated by $\{T_1, T_2, \dots, T_n\}$. Choose kernel $K(\cdot)$ as a symmetric density function and allow the following:

$$\mu_j = \int u^j K(u) du \quad \text{and} \quad \nu_j = \int u^j K^2(u) du, j = 0, 1, 2, \dots$$

$$c_k = F^{-1}(\tau_k) \quad \text{and} \quad \tau_{kk'} = \min(\tau_k, \tau_{k'}) - \tau_k \tau_{k'}.$$

The following Conditions (C1)–(C4) are needed for Theorems 1–3:

- (C1) $m(\cdot)$ has a continuous 3-th derivative in the neighborhood of t_0 ;
 (C2) $f_T(\cdot)$, the marginal density function of T , is differentiable and positive in the neighborhood of t_0 ;
 (C3) The conditional variance $\sigma^2(\cdot)$ is continuous in the neighborhood of t_0 ;
 (C4) $f(\cdot)$, the density function of the error ε , is always positive on its support.

2.1. Weighted Composite Quantile Regression Estimation

Let $Q(\tau)$ be the τ -quantile of ε and $Q_{Y|T}(\tau|t)$ be the conditional τ -th quantile of Y given $T = t$. Then, the nonparametric regression model (1) has the following quantile regression representation:

$$Q_{Y|T}(\tau|t) = m(t) + \sigma(t)Q(\tau). \quad (2)$$

Suppose that $(t_i, Y_i), i = 1, \dots, n$ are independent and identically distributed random samples from model (2). We consider estimating $m'(\cdot)$ at t_0 over $\{\tau_k, k = 1, 2, \dots, q\}$ jointly based on the following weighted local quadratic CQR loss function:

$$(\hat{a}_1, \dots, \hat{a}_q, \hat{b}_1, \hat{b}_2) = \underset{a_1, \dots, a_q, b_1, b_2}{\operatorname{argmin}} \sum_{k=1}^q \omega_k \sum_{i=1}^n \rho_{\tau_k} \{Y_i - a_k - b_1(t_i - t) - \frac{b_2}{2}(t_i - t)^2\} K\left(\frac{t_i - t}{h}\right), \quad (3)$$

where $\rho_{\tau_k}(z) = \tau_k z - z I_{(z \leq 0)}, k = 1, \dots, q$ are the quantile loss functions at q quantile positions $\tau_k = k/(q+1)$, respectively, and $\omega_k \geq 0, k = 1, \dots, q$ with $\sum_{k=1}^q \omega_k = 1$ are weights. Hereinafter, we use $\omega = (\omega_1, \dots, \omega_q)^T$ to denote the weight vector in different scenarios whenever no confusion arises. Then, the weighted local quadratic composite quantile regression (WCQR) estimator for $m'(t_0)$ is given by:

$$\hat{m}'_{WCQR}(t_0) = \hat{b}_1. \quad (4)$$

In the subsequent Theorem 1, we present the asymptotic bias, variance and asymptotic normality of $\hat{m}'_{WCQR}(t_0)$, and the proofs can be found in Section 5.

Theorem 1. Suppose that t_0 is an interior point of the support of $f_T(\cdot)$. Under the regularity conditions (C1) – (C4), if $h \rightarrow 0$ and $nh^3 \rightarrow \infty$, then the asymptotic conditional bias and variance of $\hat{m}'_{WCQR}(t_0)$ are given, respectively, by:

$$\operatorname{bias}\{\hat{m}'_{WCQR}(t_0)|\mathcal{T}\} = \frac{1}{6} m'''(t_0) \frac{\mu_4}{\mu_2^2} h^2 + o_p(h^2), \quad (5)$$

$$\operatorname{var}\{\hat{m}'_{WCQR}(t_0)|\mathcal{T}\} = \frac{1}{nh^3} \frac{v_2 \sigma^2(t_0)}{\mu_2^2 f_T(t_0)} R_\omega(q) + o_p\left(\frac{1}{nh^3}\right), \quad (6)$$

where

$$R_\omega(q) = \frac{\sum_{k=1}^q \sum_{k'=1}^q \omega_k \omega_{k'} \tau_{kk'}}{\left\{\sum_{k=1}^q \omega_k f(c_k)\right\}^2}. \quad (7)$$

Furthermore, we have the following asymptotic normal result:

$$\sqrt{nh^3} \left\{ \hat{m}'_{WCQR}(t_0) - m'(t_0) - \frac{1}{6} m'''(t_0) \frac{\mu_4}{\mu_2^2} h^2 \right\} \xrightarrow{D} N\left\{0, \frac{v_2 \sigma^2(t_0)}{\mu_2^2 f_T(t_0)} R_\omega(q)\right\}, \quad (8)$$

where “ \xrightarrow{D} ” stands for convergence in distribution.

Remark 1. If we use equal weights $\omega_k = \frac{1}{q}, k = 1, \dots, q$ over all quantiles, then the asymptotic variance of the unweighted local quadratic CQR estimator is given by $\frac{v_2 \sigma^2(t_0)}{\mu_2^2 f_T(t_0)} R(q)$, where:

$$R(q) = \frac{\sum_{k=1}^q \sum_{k'=1}^q \tau_{kk'}}{\{\sum_{k=1}^q f(c_k)\}^2}.$$

However, the importance of information at different quantiles are often different, and the information at different quantiles are correlated—depending on the error distribution. Thus, it is essential to optimally combine these quantiles' information together.

From Theorem 1, the asymptotic variance of $\hat{m}'_{WCQR}(t_0)$ depends on the weight vector ω through $R_\omega(q)$, thus, a natural way to select the optimal weight vector is to minimize $R_\omega(q)$ in (7). It is easy to know that the optimal weight vector ω^* is the solution of the following optimization problem:

$$\omega^* = \operatorname{argmin} R_\omega(q) = \operatorname{argmin} \frac{\sum_{k=1}^q \sum_{k'=1}^q \omega_k \omega_{k'} \tau_{kk'}}{\{\sum_{k=1}^q \omega_k f(c_k)\}^2}.$$

such that $\sum_{k=1}^q \omega_k = 1$ and $\sum_{k=1}^q \omega_k c_k f(c_k) = 0$.

2.2. Weighted Quantile Average Estimation

As described in Section 2.1, the WCQR estimator combines the information at different quantiles by means of the criterion function. In the following, we consider an alternative approach which combines information based on the estimators at different quantiles. As stated in Section 1, we focus on the estimation of $m'(t)$. For a fixed $\tau_k, 0 < \tau_k < 1$, consider the local quadratic nonparametric quantile regression:

$$(\hat{a}, \hat{b}_1, \hat{b}_2) = \operatorname{argmin}_{a, b_1, b_2} \sum_{i=1}^n \rho_{\tau_k} \{Y_i - a - b_1(t_i - t) - \frac{b_2}{2}(t_i - t)^2\} K\left(\frac{t_i - t}{h}\right). \quad (9)$$

The solution of b_1 in the above optimization, denoted as $\hat{m}'(\tau_k, t_0)$, provides an estimator for $m'(t_0)$. In the following, we construct the weighted quantile averaging nonparametric estimator for $m'(t_0)$ based on the weighted average of $\hat{m}'(\tau_k, t_0)$ over $\tau_k, \tau_k = k/(q+1), k = 1, 2, \dots, q$:

$$\hat{m}'_{WQAE}(t_0) = \sum_{k=1}^q \omega_k \hat{m}'(\tau_k, t_0), \quad (10)$$

where $\omega = (\omega_1, \omega_2, \dots, \omega_q)^T$ satisfies the following conditions:

$$\sum_{k=1}^q \omega_k = 1, \quad (11)$$

$$\sum_{k=1}^q \omega_k F^{-1}(\tau_k) = 0. \quad (12)$$

The introduction of ω can eliminate the bias term caused by the asymmetric random error. Meanwhile, the weight vector ω satisfying conditions (11) and (12) can guarantee the estimation consistency and asymptotic unbiasedness of $\hat{m}'_{WQAE}(t_0)$ asymptotically, which can also be seen from the proof of the asymptotic properties of $\hat{m}'_{WQAE}(t_0)$ in Section 5. In these weight vectors, which satisfy conditions (11) and (12), we can select the optimal one by optimization. The details are discussed in the subsequent Section.

In the subsequent Theorem 2, we present the asymptotic bias, variance and asymptotic normality of $\hat{m}'_{WQAE}(t_0)$, and the proofs can be found in Section 5.

Theorem 2. Suppose that t_0 is an interior point of the support of $f_T(\cdot)$, and the weight vector ω satisfies conditions (10) and (11). Under the regularity Conditions (C1)–(C4) in Section 5, if

$h \rightarrow 0$ and $nh^3 \rightarrow \infty$, then the asymptotic conditional bias and variance of the weighted quantile average estimation $\hat{m}'_{WQAE}(t_0)$ are given, respectively, by:

$$\text{bias}\{\hat{m}'_{WQAE}(t_0)|\mathcal{T}\} = \frac{1}{6}m'''(t_0)\frac{\mu_4}{\mu_2^2}h^2 + o_p(h^2), \quad (13)$$

$$\text{var}\{\hat{m}'_{WQAE}(t_0)|\mathcal{T}\} = \frac{1}{nh^3}\frac{v_2\sigma^2(t_0)}{\mu_2^2 f_T(t_0)}J_\omega(q) + o_p\left(\frac{1}{nh^3}\right), \quad (14)$$

where $J_\omega(q) = \omega^T H \omega$, H is the $q \times q$ matrix with the (k, k') -th element $\frac{\tau_{kk'}}{f(c_k)f(c_{k'})}$, that is

$$J_\omega(q) = \sum_{k=1}^q \sum_{k'=1}^q \frac{\omega_k \omega_{k'} \tau_{kk'}}{f(c_k)f(c_{k'})}. \quad (15)$$

Furthermore, conditioning on \mathcal{T} , we have the following asymptotic normal distribution:

$$\sqrt{nh^3}\{\hat{m}'_{WQAE}(t_0) - m'(t_0) - \frac{1}{6}m'''(t_0)\frac{\mu_4}{\mu_2^2}h^2\} \xrightarrow{D} N\left\{0, \frac{v_2\sigma^2(t_0)}{\mu_2^2 f_T(t_0)}J_\omega(q)\right\}. \quad (16)$$

Remark 2. If we simply use equal weights in (14), then the resulting unweighted quantile average estimator for $m'(t_0)$ has the asymptotic normality in Theorem 2 with $J_\omega(q)$ replaced by:

$$J(q) = \frac{1}{q^2} \sum_{k=1}^q \sum_{k'=1}^q \frac{\tau_{kk'}}{f(c_k)f(c_{k'})}.$$

Remark 3. From Theorem 2, the covariance matrix of $\hat{m}'_{WQAE}(t_0)$ depends on the weight vector through $J_\omega(q)$, thus, a natural way to select the optimal weight vector is to minimize $J_\omega(q)$ in (14).

The following Theorem gives the optimal weight vector and the optimal weighted quantile average estimation of $m'(t_0)$.

Theorem 3. Supposing that the Conditions (C1) – (C4) hold, then the optimal weight vector minimizing $J_\omega(q)$ is:

$$\omega^* = \frac{(c^T H^{-1} c)H^{-1}\mathbf{1} - (\mathbf{1}^T H^{-1} c)H^{-1}c}{(c^T H^{-1} c)(\mathbf{1}^T H^{-1}\mathbf{1}) - (\mathbf{1}^T H^{-1} c)^2}, \quad (17)$$

where c is a q -dimensional column vector with k -th element $c_k = F^{-1}(\tau_k)$ and $\mathbf{1}$ is a q -dimensional column vector with all elements 1. Furthermore, the corresponding conditionally variance of the optimal weighted quantile average estimation of $\hat{m}'_{WQAE}(t_0)$, denoted as $\hat{m}'_{WQAE}^*(t_0)$, is given by:

$$\text{var}\{\hat{m}'_{WQAE}^*(t_0)|\mathcal{T}\} = \frac{1}{nh^3}\frac{v_2\sigma^2(t_0)}{\mu_2^2 f_T(t_0)}J_{\omega^*}(q) + o_p\left(\frac{1}{nh^3}\right),$$

where:

$$J_{\omega^*}(q) = \sum_{k=1}^q \sum_{k'=1}^q \frac{\omega_k^* \omega_{k'}^* \tau_{kk'}}{f(c_k)f(c_{k'})}.$$

Comparing the weighted local quadratic CQR estimation, the weighted quantile average estimation and the local quadratic least squares estimators for $m'(t_0)$, we see that they have the same leading bias term $\frac{1}{6}m'''(t_0)\frac{\mu_4}{\mu_2^2}h^2$, whereas their asymptotic variances are different.

3. Comparison of Asymptotic Efficiency

The WQAE differs from the WCQR estimator in several aspects. While the WCQR estimator is based on the aggregation of several quantile loss functions, the WQAE is

based on a weighted average of separate estimators from different quantiles. As a result, computing the WQAE only involves q separate parameter minimization problems, whereas the WCQR requires solving a larger parameter minimization problem. In addition, to ensure a proper loss function, the weights $\omega_k, k = 1, \dots, n$ in $R_\omega(q)$ are restricted to be non-negative; by contrast, the weights $\omega_k, k = 1, \dots, n$ in $J_\omega(q)$ can be negative. Obviously, it is computationally appealing to impose less constraint on the weights.

From Theorem 1, the mean square of error (MSE) of the WCQR estimation $\hat{m}'_{WCQR}(t_0)$ is given by:

$$MSE\{\hat{m}'_{WCQR}(t_0)\} = \left\{\frac{1}{6}m'''(t_0)\frac{\mu_4}{\mu_2}\right\}^2 h^4 + \frac{1}{nh^3} \frac{v_2\sigma^2(t_0)}{\mu_2^2 f_T(t_0)} R_\omega(q) + o_p\left(h^4 + \frac{1}{nh^3}\right).$$

Thus, the optimal variable bandwidth minimizing $MSE\{\hat{m}'_{WCQR}(t_0)\}$ is:

$$h_{WCQR}^{opt} = R_\omega(q)^{1/7} \left[\frac{27v_2\sigma^2(t_0)}{f_T(t_0)\{m'''(t_0)\mu_4\}^2} \right]^{1/7} n^{-1/7}.$$

From Theorem 2, the MSE of the WQAE $\hat{m}'_{WQAE}(t_0)$ is given by:

$$MSE\{\hat{m}'_{WQAE}(t_0)\} = \left\{\frac{1}{6}m'''(t_0)\frac{\mu_4}{\mu_2}\right\}^2 h^4 + \frac{1}{nh^3} \frac{v_2\sigma^2(t_0)}{\mu_2^2 f_T(t_0)} J_\omega(q) + o_p\left(h^4 + \frac{1}{nh^3}\right).$$

Thus, the optimal variable bandwidth minimizing $MSE\{\hat{m}'_{WQAE}(t_0)\}$ is:

$$h_{WQAE}^{opt} = J_\omega(q)^{1/7} \left[\frac{27v_2\sigma^2(t_0)}{f_T(t_0)\{m'''(t_0)\mu_4\}^2} \right]^{1/7} n^{-1/7}.$$

4. Simulation Studies

In this Section, we implement simulation studies to compare the finite sample performance of our WQAE with those of the WCQR and the local linear least square estimates. In all examples, the kernel function is chosen to be the Gaussian kernel function, and the number of replications is set to be 400. Similar to Kai et al. [9], we use the short-cut strategy to select the bandwidth in our simulation studies.

An important quantity to valuate the performance of different estimators is the average squared errors (ASE), which can be represented by:

$$ASE(\hat{g}) = \frac{1}{n_{grid}} \sum_{i=1}^{n_{grid}} \{\hat{g}(u_k) - g(u_k)\}^2$$

with g being $m'(\cdot)$ in the simulation, where $\{u_k, k = 1, 2, \dots, n_{grid}\}$ are the grid points at which the estimator $\hat{g}(\cdot)$ of g is evaluated. Set $n_{grid} = 200$ and set the grid points to evenly distribute over the interval on which $m'(\cdot)$ is estimated. Furthermore, we can also evaluate the different estimators \hat{g}_1 and \hat{g}_2 of $m'(\cdot)$ via the ratio of the average squared errors (RASE) defined by:

$$RASE(\hat{g}_1, \hat{g}_2) = \frac{ASE(\hat{g}_2)}{ASE(\hat{g}_1)}.$$

In the simulation below, we compare the finite behaviors of these three estimators under different conditions, including different symmetric or asymmetric errors, homoscedastic or heteroscedastic models.

4.1. Homoscedastic Model

We consider the following homoscedastic model:

$$\text{Model 1 : } Y = \sin(2T) + 2\exp(-16T^2) + 0.5\epsilon,$$

where the covariate T follows $N(0, 1)$, and the nonparametric regression function is $m(t) = \sin(2t) + 2\exp(-16t^2)$. In our simulation, we generate 400 random samples, each covering $n = 200$ observations. Our aim is to estimate the derivative of $m(t)$ over $[-1.5, 1.5]$. It is

noted that the Cauchy distribution is the truncated Cauchy distribution on $[-10, 10]$. We investigate the finite sample behaviors of the WCQR, WQAE and LSE (the local linear least squares estimator) separately according to the RASE and ASE under the models under some symmetric and asymmetric errors in the next two examples.

Example 1. In this example, we consider Model 1 under different symmetric error distributions. The mean and standard deviations of RASEs with 400 simulations are established in Table 1. The following can be seen for Table 1:

- (1) Under the standard normal error $N(0, 1)$, the LSE estimator is the best one among these three estimators, and WQAE performs nearly as well as WCQR.
- (2) Under the non-normal and symmetric errors, WCQR is consistently superior with WQAE and LSE methods, LSE especially performs consistently the worst. Obviously, our RASEs are significantly larger than 1.

Table 1. The biases and standard deviations for RASEs in Example 1.

| ε | $RASE(\hat{m}'_{WQAE}, \hat{m}'_{WCQR})$ | | | $RASE(\hat{m}'_{WQAE}, \hat{m}'_{LSE})$ | | |
|----------------|--|--------------------|--------------------|---|--------------------|--------------------|
| | $q = 5$ | $q = 9$ | $q = 19$ | $q = 5$ | $q = 9$ | $q = 19$ |
| $N(0, 1)$ | 0.9823 (0.0784) | 0.9745 (0.1028) | 0.9379 (0.1235) | 0.9217 (0.1332) | 0.9391 (0.1213) | 0.9264 (0.1229) |
| <i>Laplace</i> | 1.0696 (0.1765) | 1.1052 (0.2354) | 1.1143 (0.2689) | 1.3076 (0.3757) | 1.2919 (0.3723) | 1.2298 (0.3326) |
| $t(3)$ | 1.0235 (0.1329) | 1.0536 (0.2092) | 1.1137 (0.2901) | 1.6441 (0.8123) | 1.5678 (0.8002) | 1.4705 (0.6790) |
| $U[-1, 1]$ | 1.2243 (0.2029) | 1.4221 (0.3048) | 1.4765 (0.3832) | 1.1268 (0.2035) | 1.4373 (0.3414) | 1.6057 (0.4565) |
| <i>Cauchy</i> | 1.3066 (0.3960) | 1.5016 (0.5601) | 1.5831 (0.6746) | 2.2993 (0.9127) | 2.2939 (0.9850) | 2.1652 (0.9245) |

Example 2. In this example, we consider Model 1 under different asymmetric error distributions. The mean and standard deviations of ASEs with 400 simulations are presented in Table 2. It is noted that “ $F(4, 6)$ ” denotes the centralized $F(4, 6)$. The following can be seen for Table 2:

- (1) WCQR, whose bias is non-vanishing and does not approach to zero under asymmetric errors, breaks down. So, WCQR is the worst in the case of asymmetric errors.
- (2) For all the asymmetric errors considered, WQAE performs consistently better than local linear least squares.

Table 2. The means and standard deviations for ASEs in Example 2.

| Distribution of ε | $ASE(\hat{m}'_{WQAE})$ | | | $ASE(\hat{m}'_{WCQR})$ | | | $ASE(\hat{m}'_{LSE})$ |
|-------------------------------|------------------------|--------------------|--------------------|------------------------|--------------------|--------------------|-----------------------|
| | $q = 5$ | $q = 9$ | $q = 19$ | $q = 5$ | $q = 9$ | $q = 19$ | |
| $N(0, 1)$ | 0.0297 (0.0185) | 0.0302 (0.0180) | 0.0339 (0.0220) | 0.1302 (0.0275) | 0.0883 (0.0227) | 0.0565 (0.0181) | 0.0426 (0.0264) |
| <i>Laplace</i> | 0.0223 (0.0121) | 0.0226 (0.0132) | 0.0289 (0.0333) | 0.0871 (0.0200) | 0.0614 (0.0268) | 0.0422 (0.0299) | 0.0377 (0.0719) |
| $t(3)$ | 0.0492 (0.0243) | 0.0480 (0.0243) | 0.0483 (0.0245) | 0.1191 (0.0348) | 0.0888 (0.0286) | 0.0665 (0.0248) | 0.0515 (0.0231) |
| $U[-1, 1]$ | 0.0528 (0.0257) | 0.0525 (0.0251) | 0.0543 (0.0252) | 0.0960 (0.0343) | 0.0786 (0.0303) | 0.0647 (0.0269) | 0.0570 (0.0267) |
| <i>Cauchy</i> | 0.0194 (0.0172) | 0.0205 (0.0209) | 0.0213 (0.0231) | 0.0450 (0.0153) | 0.0366 (0.0146) | 0.0300 (0.0155) | 0.0223 (0.0192) |

4.2. Heteroscedastic Model

In this subsection, we mainly consider the following heteroscedastic model:

$$\text{Model 2 : } Y = T \sin(2\pi T) + \sigma(T)\varepsilon,$$

where the covariate T follows $N(0, 1)$, $\sigma(t) = (2 + \cos(2\pi t))/10$, and the nonparametric regression function is $m(t) = t \sin(2\pi t)$. Our aim is to estimate the derivative of $m(t)$ on $[-1.5, 1.5]$. In our simulation, we generate 400 random samples, each having $n = 200$ observations. Similar to the arguments in the homoscedastic model, we evaluate their behaviors under symmetric and asymmetric errors in the following Examples 3 and 4.

Example 3. We consider Model 2 under different symmetric error distributions in this example. The simulation results of biases and standard deviations of RASEs over 400 simulations are summarized in Table 3. We have the following findings:

- (1) Under the standard normal error $N(0, 1)$, LSE outperforms both WQAE and WCQR. Even for normal errors, our RASEs are a little smaller than 1.
- (2) For all the symmetric errors considered except the normal error, WQAE performs significantly better than the LSE, and our RASEs are obviously greater than 1 compared with the LSE. Furthermore, WQAE performs better than WCQR in most cases.

Table 3. The biases and standard deviations for RASEs in Example 3.

| ε | RASE($\hat{m}'_{WQAE}, \hat{m}'_{WCQR}$) | | | RASE($\hat{m}'_{WQAE}, \hat{m}'_{LSE}$) | | |
|---------------|--|--------------------|--------------------|---|--------------------|--------------------|
| | $q = 5$ | $q = 9$ | $q = 19$ | $q = 5$ | $q = 9$ | $q = 19$ |
| $N(0, 1)$ | 0.9841 (0.0518) | 0.9749 (0.0642) | 0.9532 (0.0767) | 0.9328 (0.1211) | 0.9476 (0.1115) | 0.9357 (0.1109) |
| Laplace | 0.9963 | 1.0152 | 1.0130 | 1.0658 | 1.0636 | 1.0417 |
| $t(3)$ | 1.0368 (0.1886) | 1.0657 (0.2442) | 1.0688 (0.2638) | 1.3067 (0.7175) | 1.2584 (0.5569) | 1.1891 (0.4523) |
| $U[-1, 1]$ | 1.0859 (0.1335) | 1.0830 (0.2008) | 0.9387 (0.2842) | 1.1836 (0.2804) | 1.2602 (0.3641) | 1.1277 (0.4413) |
| Cauchy | 1.1273 (0.2488) | 1.1784 (0.3176) | 1.1683 (0.3371) | 1.5532 (0.4678) | 1.5329 (0.4672) | 1.4596 (0.4234) |

Example 4. In this example, we consider Model 2 under different asymmetric error distributions. The simulation results of biases and standard deviations of RASEs over 400 simulations are summarized in Table 4. Similar to Example 2, the simulation results show the following:

- (1) WCQR breaks down for its non-vanishing bias in the case of asymmetric errors. Therefore, WCQR performs the worst among the three estimators.
- (2) For all the asymmetric errors considered, WQAE performs consistently better than local linear least squares.

Table 4. The means and standard deviations for ASEs in Example 4.

| ε | $ASE(\hat{m}'_{WQAE})$ | | | $ASE(\hat{m}'_{WCQR})$ | | | $ASE(\hat{m}'_{LSE})$ |
|----------------|------------------------|--------------------|--------------------|------------------------|--------------------|--------------------|-----------------------|
| | $q = 5$ | $q = 9$ | $q = 19$ | $q = 5$ | $q = 9$ | $q = 19$ | |
| $N(0, 1)$ | 0.0249 (0.0383) | 0.0287 (0.0514) | 0.0348 (0.0562) | 0.0711 (0.0779) | 0.0644 (0.1101) | 0.0553 (0.1245) | 0.0752 (0.2549) |
| <i>Laplace</i> | 0.0312 (0.1021) | 0.0323 (0.0902) | 0.0451 (0.1179) | 0.0831 (0.0824) | 0.0678 (0.0723) | 0.0556 (0.0638) | 0.0724 (0.1357) |
| $t(3)$ | 0.0369 (0.0198) | 0.0365 (0.0183) | 0.0376 (0.0192) | 0.0783 (0.0190) | 0.0602 (0.0169) | 0.0476 (0.0158) | 0.0434 (0.0221) |
| $U[-1, 1]$ | 0.0321 (0.0169) | 0.0302 (0.0173) | 0.0329 (0.0171) | 0.0567 (0.0194) | 0.0465 (0.0177) | 0.0392 (0.0166) | 0.0387 (0.0214) |
| <i>Cauchy</i> | 0.0348 (0.0219) | 0.0375 (0.0316) | 0.0414 (0.0455) | 0.1061 (0.0277) | 0.0833 (0.0274) | 0.0657 (0.0301) | 0.0569 (0.0526) |

5. Proofs

Before embarking on the proofs of Theorems 1–3, we first give and prove Lemma 1. Suppose that the kernel function K has the finite support $[-M, M]$. The following notations are needed to present our theoretical results:

Denote

$$u_k = \sqrt{nh}\{a_k - m(t_0) - \sigma(t_0)c_k\}, k = 1, \dots, q,$$

and

$$v_j = h^j \sqrt{nh}\{j!b_j - m^{(j)}(t_0)\}/j!, j = 1, 2.$$

Set $x_i = (t_i - t_0)/h$, $K_i = K(x_i)$ and $\Delta_{i,k} = \frac{u_k}{\sqrt{nh}} + \frac{v_1 x_i}{\sqrt{nh}} + \frac{v_2 x_i^2}{\sqrt{nh}}$. Write $d_{i,k} = c_k\{\sigma(t_i) - \sigma(t_0)\} + r_{i,2}$, $r_{i,2} = m(t_i) - m(t_0) - m'(t_0)(t_i - t_0) - \frac{m''(t_0)}{2}(t_i - t_0)^2$. Define $\eta_{i,k}^* = I\{\varepsilon_i \leq c_k - d_{i,k}/\sigma(t_i) - \tau_k\}$, and let $W^* = (w_{11}^*, \dots, w_{1q}^*, w_{21}^*, w_{22}^*)^T$, with

$$w_{1k}^* = \frac{\omega_k}{\sqrt{nh}} \sum_{i=1}^n K_i \eta_{i,k}^*, \quad k = 1, \dots, q, \quad w_{2j}^* = \frac{1}{\sqrt{nh}} \sum_{k=1}^q \omega_k \sum_{i=1}^n K_i x_i^j \eta_{i,k}^*, \quad j = 1, 2.$$

Furthermore, let S_{11} be a $q \times q$ diagonal matrix with diagonal elements $\{\omega_k f(c_k), k = 1, \dots, q\}$; S_{12} be a $q \times 2$ matrix with (k, j) -element $\omega_k f(c_k) \mu_j, k = 1, \dots, q$ and $j = 1, 2$. $S_{21} = S_{12}^T$; and S_{22} be a 2×2 diagonal with diagonal elements $\mu_2 \sum_{k=1}^q \omega_k f(c_k)$ and $\mu_4 \sum_{k=1}^q \omega_k f(c_k)$. In addition, let $S_{n,11}$ be a $q \times q$ diagonal matrix with diagonal elements $\{\omega_k f(c_k) \sum_{i=1}^n K_i / (nh \sigma(t_i)), k = 1, \dots, q\}$, let S_{12} be a $q \times 2$ matrix with (k, j) -element $\omega_k f(c_k) \sum_{i=1}^n K_i x_i^j / (nh \sigma(t_i)), k = 1, \dots, q$ and $j = 1, 2$. $S_{n,21} = S_{n,12}^T$, and let $S_{n,22}$ be a 2×2 diagonal with diagonal elements $\sum_{k=1}^q \omega_k f(c_k) \sum_{i=1}^n K_i x_i^2 / (nh \sigma(t_i))$. Similarly, let Σ_{11} be a $q \times q$ matrix with (k, k') -element $v_0 \omega_k \omega_{k'} \tau_{kk'}, k, k' = 1, \dots, q$, let Σ_{12} be a $q \times 2$ matrix with (k, j) -element $v_j \omega_k \sum_{k'=1}^q \omega_{k'} \tau_{kk'}, k = 1, \dots, q$ and $j = 1, 2$. $\Sigma_{21} = \Sigma_{12}^T$, and let Σ_{22} be a 2×2 matrix with (j, j') -element $v_{j+j'} \sum_{k,k'=1}^q \omega_k \omega_{k'} \tau_{kk'}$ for $j, j' = 1, 2$.

Define:

$$S = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix}, S_n = \begin{bmatrix} S_{n,11} & S_{n,12} \\ S_{n,21} & S_{n,22} \end{bmatrix},$$

$$\Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}.$$

Partition $(S^{-1}\Sigma S^{-1})$ into four submatrices as follows:

$$S^{-1}\Sigma S^{-1} = \begin{bmatrix} (S^{-1}\Sigma S^{-1})_{11} & (S^{-1}\Sigma S^{-1})_{12} \\ (S^{-1}\Sigma S^{-1})_{21} & (S^{-1}\Sigma S^{-1})_{22} \end{bmatrix},$$

where $(\cdot)_{11}$ stands for the top left-hand $q \times q$ submatrix, and $(\cdot)_{22}$ stands for the bottom right-hand element.

Lemma 1. Denote $\hat{\theta}_n = (\hat{u}_1, \dots, \hat{u}_q, \hat{v}_1, \hat{v}_2)^T$ as the minimizer of the weighted local quadratic CQR loss. Under regularity Conditions (C1)–(C4), we have:

$$\hat{\theta}_n + \frac{\sigma(t_0)}{f_T(t_0)} S^{-1} E[W_n^* | \mathcal{T}] \xrightarrow{D} \text{MVN}_{q+2} \left\{ 0, \frac{\sigma^2(t_0)}{f_T(t_0)} S^{-1} \Sigma S^{-1} \right\}.$$

Proof of Lemma 1. The proof is similar to that of Theorem 5 of Kai et al. [9]. We divide the whole procedure into three steps:

Step 1 Minimizing the weighted local quadratic CQR loss is equivalent to minimizing $L_n(\theta)$, defined as:

$$L_n(\theta) = \sum_{k=1}^q \omega_k u_k \left\{ \sum_{i=1}^n \frac{K_i \eta_{i,k}^*}{\sqrt{nh}} \right\} + \sum_{j=1}^2 v_j \sum_{k=1}^q \omega_k \left\{ \sum_{i=1}^n \frac{K_i x_i^j \eta_{i,k}^*}{\sqrt{nh}} \right\} + \sum_{k=1}^q \omega_k B_{n,k}(\theta).$$

with respect to $\theta = (u_1, \dots, u_q, v_1, v_2)^T$, where

$$B_{n,k}(\theta) = \sum_{i=1}^n \left(K_i \int_0^{\Delta_{i,k}} \left[I \left\{ \varepsilon_i \leq c_k - \frac{d_{i,k}}{\sigma(t_i)} + \frac{z}{\sigma(t_i)} \right\} - I \left\{ \varepsilon_i \leq c_k - \frac{d_{i,k}}{\sigma(t_i)} \right\} \right] dz \right).$$

For further details, see the proofs of Lemma 2 in Kai et al. [9].

Step 2 Under regularity conditions (C1)–(C4), we have:

$$L_n(\theta) = \frac{1}{2} \theta^T S_n \theta + (W_n^*)^T \theta + o_p(1).$$

For further details, see the proof of Lemma 2 of Kai et al. [9].

Step 3 It is easy to obtain:

$$\frac{1}{nh} \sum_{i=1}^n K_i x_i^j \xrightarrow{P} f_T(t_0) \mu_j, \quad j = 0, 1, 2,$$

where \xrightarrow{P} stands for convergence in probability. Thus we have:

$$S_n \xrightarrow{P} \frac{f_T(t_0)}{\sigma(t_0)} S = \frac{f_T(t_0)}{\sigma(t_0)} \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix}.$$

Together with the results of Step 2, we have:

$$L_n(\theta) = \frac{1}{2} \frac{f_T(t_0)}{\sigma(t_0)} \theta^T S \theta + (W_n^*)^T \theta + o_p(1).$$

Note that the convex function $L_n(\theta) - (W_n^*)^T \theta$ converges in probability to the convex function $\theta^T S \theta f_T(t_0) / (2\sigma(t_0))$. Then, it can be deduced from the convexity lemma of Pollard [20] that the quadratic approximation to $L_n(\theta)$ holds uniformly for θ in any compact set, which leads to:

$$\hat{\theta}_n = - \frac{\sigma(t_0)}{f_T(t_0)} S^{-1} W_n^* + o_p(1).$$

Finally, similar to the procedures of Theorem 5 in Kai et al. [9], we have:

$$W_n^*|\mathcal{T} - E[W_n^*|\mathcal{T}] \xrightarrow{D} \text{MVN}_{q+2}\{0, f_T(t_0)\Sigma\}.$$

This completes the proof.

□

Proof of Remark 1. If we use equal weights $\omega_k = \frac{1}{q}, k = 1, \dots, q$ over all quantiles, then we have:

$$\begin{aligned} R_w(q) &= \frac{\sum_{k=1}^q \sum_{k'=1}^q \frac{1}{q} \frac{1}{q} \tau_{kk'}}{\{\sum_{k=1}^q \frac{1}{q} f(c_k)\}^2} \\ &= \frac{\frac{1}{q^2} \sum_{k=1}^q \sum_{k'=1}^q \tau_{kk'}}{\frac{1}{q^2} \{\sum_{k=1}^q f(c_k)\}^2} \\ &= \frac{\sum_{k=1}^q \sum_{k'=1}^q \tau_{kk'}}{\{\sum_{k=1}^q f(c_k)\}^2} \\ &\doteq R(q). \end{aligned}$$

Therefore, the asymptotic variance of the unweighted local quadratic CQR estimator can be given by $\frac{\nu_2 \sigma^2(t_0)}{\mu_2^2 f_T(t_0)} R(q)$. □

Proof of Remark 2. If we use equal weights $\omega_k = \frac{1}{q}, k = 1, \dots, q$ over all quantiles, then we have:

$$\begin{aligned} J_w(q) &= \sum_{k=1}^q \sum_{k'=1}^q \frac{\frac{1}{q} \frac{1}{q} \tau_{kk'}}{f(c_k) f(c'_k)} \\ &= \frac{1}{q^2} \sum_{k=1}^q \sum_{k'=1}^q \frac{\tau_{kk'}}{f(c_k) f(c'_k)} \\ &\doteq J(q), \end{aligned}$$

then the resulting unweighted quantile average estimator for $m'(t_0)$ has the asymptotic normality in Theorem 2 with $J_w(q)$ replaced by:

$$J(q) = \frac{1}{q^2} \sum_{k=1}^q \sum_{k'=1}^q \frac{\tau_{kk'}}{f(c_k) f(c'_k)}.$$

□

Proof of Theorem 1. We apply Lemma 1 to obtain the asymptotic normality of $\hat{m}'_{WCQR}(t_0)$. It is easy to obtain:

$$S_{12} = (0_{q \times 1}, \mu_2 \omega_k f(c_k)_{q \times 1}),$$

and

$$S_{22} = \text{diag}\{\mu_2 \sum_{k=1}^q \omega_k f(c_k), \mu_4 \sum_{k=1}^q \omega_k f(c_k)\}.$$

$$\text{Since } S_{11} = \text{diag}\{\omega_1 f(c_1), \dots, \omega_q f(c_q)\},$$

$$(S^{-1})_{22} = (S_{22} - S_{21} S_{11}^{-1} S_{12})^{-1} = \text{diag}\left[\left\{\mu_2 \sum_{k=1}^q \omega_k f(c_k)\right\}^{-1}, \left\{(\mu_4 - \mu_2^2) \sum_{k=1}^q \omega_k f(c_k)\right\}^{-1}\right].$$

Note that $(S^{-1})_{21} = -(S^{-1})_{22} S_{21} S_{11}^{-1}$. Thus, we have:

$$(S^{-1})_{21} = (\mathbf{0}_{q \times 1}, \left[\frac{\mu_2}{(\mu_4 - \mu_2^2) \sum_{k=1}^q \omega_k f(c_k)} \right] \mathbf{1}_{q \times 1})^T.$$

where $\mathbf{0}_{q \times 1}$ and $\mathbf{1}_{q \times 1}$ denote the q -dimensional column vector with all entries of 0 and 1, respectively. Denote $e = (1, 0)^T$. By Lemma 1, we can obtain:

$$\begin{aligned} \text{bias}\{\hat{m}'_{\text{WCQR}}(t_0)|\mathcal{T}\} &= -\frac{\sigma(t_0)}{hf_T(t_0)} \frac{1}{\sqrt{nh}} e^T \{(S^{-1})_{21} E[W_{1n}^*|\mathcal{T}] + (S^{-1})_{22} E[W_{2n}^*|\mathcal{T}]\} \\ &= -\frac{\sigma(t_0)}{hf_T(t_0)} \frac{1}{\mu_2 \sum_{k=1}^q \omega_k f(c_k)} \frac{1}{\sqrt{nh}} E[w_{21}^*|\mathcal{T}]. \end{aligned}$$

Note that:

$$E[w_{2j}^*|\mathcal{T}] = \frac{1}{\sqrt{nh}} \sum_{i=1}^n K_i x_i^j \sum_{k=1}^q \omega_k [F\{c_k - \frac{d_{i,k}}{\sigma(t_i)}\} - F(c_k)].$$

Similarly, we have:

$$\sum_{k=1}^q \omega_k [F\{c_k - \frac{d_{i,k}}{\sigma(t_i)}\} - F(c_k)] = -\sum_{k=1}^q \omega_k f(c_k) \frac{r_{i,2}}{\sigma(t_i)} \{1 + o_p(1)\}.$$

Therefore:

$$\text{bias}\{\hat{m}'_{\text{WCQR}}(t_0)|\mathcal{T}\} = \frac{1}{nh^2} \frac{\sigma(t_0)}{f_T(t_0)} \sum_{i=1}^n K_i x_i \frac{r_{i,2}}{\sigma(t_i)} \{1 + o_p(1)\}.$$

It is easy to obtain:

$$\frac{1}{nh} \sum_{i=1}^n K_i x_i \frac{r_{i,2}}{\sigma(t_i)} = \frac{f_T(t_0) m'''(t_0)}{6\sigma(t_0)} \frac{\mu_4}{\mu_2} h^3 \{1 + o_p(1)\}.$$

Thus, we obtain:

$$\text{bias}\{\hat{m}'_{\text{WCQR}}(t_0)|\mathcal{T}\} = \frac{1}{6} m'''(t_0) \frac{\mu_4}{\mu_2} h^2 + o_p(h^2).$$

Furthermore, the conditional variance of $\hat{m}'_{\text{WCQR}}(t_0)$ is:

$$\begin{aligned} \text{var}\{\hat{m}'_{\text{WCQR}}(t_0)|\mathcal{T}\} &= \frac{1}{nh^3} \frac{\sigma^2(t_0)}{f_T(t_0)} e^T e + o_p\left(\frac{1}{nh^3}\right) \\ &= \frac{1}{nh^3} \frac{v_2}{\mu_2^2} \frac{\sigma^2(t_0)}{f_T(t_0)} R_\omega(q) + o_p\left(\frac{1}{nh^3}\right). \end{aligned}$$

which completes the proof. \square

Proof of Theorem 2. It is easy to see that:

$$\begin{aligned} E\{\hat{m}'(\tau_k, t_0)|\mathcal{T}\} &= -\frac{1}{h\sqrt{nh}} \frac{\sigma(t_0)}{f_T(t_0)f(c_k)} \frac{1}{\mu_2} E[w_{21}^*|\mathcal{T}] + m'(t_0) \\ &= -\frac{1}{nh^2} \frac{\sigma(t_0)}{f_T(t_0)f(c_k)} \frac{1}{\mu_2} \sum_{i=1}^n K_i x_i \left[F\left(c_k - \frac{d_{i,k}}{\sigma(t_i)}\right) - F(c_k)\right] + m'(t_0) \\ &= \frac{1}{nh^2} \frac{\sigma(t_0)}{f_T(t_0)} \frac{1}{\mu_2} \sum_{i=1}^n K_i x_i \frac{d_{i,k}}{\sigma(t_i)} (1 + o_p(1)) + m'(t_0) \\ &= \frac{1}{nh^2} \frac{\sigma(t_0)}{f_T(t_0)} \frac{1}{\mu_2} \sum_{i=1}^n K_i x_i \frac{c_k \{\sigma(t_i) - \sigma(t_0)\} + r_{i,2}}{\sigma(t_i)} (1 + o_p(1)) + m'(t_0). \end{aligned}$$

Then, we have:

$$\begin{aligned} \text{bias}\{\hat{m}'_{WQAE}(t_0)|\mathcal{T}\} &= E\{\hat{m}'_{WQAE}(t_0)|\mathcal{T}\} - m'(t_0) \\ &= E\left\{\sum_{k=1}^q \omega_k \hat{m}'(\tau_k, t_0)|\mathcal{T}\right\} - m'(t_0) = \sum_{k=1}^q \omega_k [E\{\hat{m}'(\tau_k, t_0)|\mathcal{T}\} - m'(t_0)] \\ &= \sum_{k=1}^q \omega_k \left[\frac{1}{nh^2} \frac{\sigma(t_0)}{f_T(t_0)} \frac{1}{\mu_2} \sum_{i=1}^n K_i x_i \frac{c_k \{\sigma(t_i) - \sigma(t_0)\} + r_{i,2}}{\sigma(t_i)} (1 + o_p(1)) \right] \\ &= \frac{1}{nh^2} \frac{\sigma(t_0)}{f_T(t_0)} \frac{1}{\mu_2} \sum_{i=1}^n K_i x_i \sum_{k=1}^q \omega_k \frac{c_k \{\sigma(t_i) - \sigma(t_0)\} + r_{i,2}}{\sigma(t_i)} (1 + o_p(1)). \end{aligned}$$

Using the condition $\sum_{k=1}^q \omega_k c_k = 0$, we obtain:

$$\begin{aligned} \text{bias}\{\hat{m}'_{WQAE}(t_0)|\mathcal{T}\} &= \frac{1}{nh^2} \frac{\sigma(t_0)}{f_T(t_0)} \frac{1}{\mu_2} \sum_{i=1}^n K_i x_i \sum_{k=1}^q \omega_k \frac{r_{i,2}}{\sigma(t_i)} (1 + o_p(1)) \\ &= \frac{1}{nh^2} \frac{\sigma(t_0)}{f_T(t_0)} \frac{1}{\mu_2} \sum_{i=1}^n K_i x_i \frac{r_{i,2}}{\sigma(t_i)} (1 + o_p(1)). \end{aligned}$$

By using the following fact:

$$\frac{1}{nh} \sum_{i=1}^n K_i x_i \frac{r_{i,2}}{\sigma(t_i)} = \frac{f_T(t_0) m'''(t_0)}{6\sigma(t_0)} \frac{\mu_4}{\mu_2} h^3 (1 + o_p(1)),$$

we obtain:

$$\text{bias}\{\hat{m}'_{WQAE}(t_0)|\mathcal{T}\} = \frac{1}{6} m'''(t_0) \frac{\mu_4}{\mu_2^2} h^2 + o_p(h^2). \quad (18)$$

$$\begin{aligned} \text{var}\{\hat{m}'_{WQAE}(t_0)|\mathcal{T}\} &= \text{var}\left\{\sum_{k=1}^q \omega_k \hat{m}'(\tau_k, t_0)|\mathcal{T}\right\} \\ &= \text{var}\left\{-\frac{1}{h\sqrt{nh}} \frac{\sigma(t_0)}{f_T(t_0)} \frac{1}{\mu_2} \sum_{k=1}^q \frac{\omega_k}{f(c_k)} \omega_{21}^*|\mathcal{T}\right\} \\ &= \frac{1}{nh^3} \frac{\sigma^2(t_0)}{f_T^2(t_0)} \frac{1}{\mu_2^2} \text{var}\left\{\sum_{k=1}^q \frac{\omega_k}{f(c_k)} \omega_{21}^*|\mathcal{T}\right\}. \end{aligned}$$

Using the similar arguments in Kai et al. [9], we have:

$$\text{var}\left\{\sum_{k=1}^q \frac{\omega_k}{f(c_k)} (\omega_{21}^* - \omega_{21})|\mathcal{T}\right\} = o_p(1).$$

Thus, we have:

$$\begin{aligned} \text{var}\{\hat{m}'_{WQAE}(t_0)|\mathcal{T}\} &= \frac{1}{nh^3} \frac{\sigma^2(t_0)}{f_T^2(t_0)} \frac{1}{\mu_2^2} \text{var}\left\{\sum_{k=1}^q \frac{\omega_k}{f(c_k)} \omega_{21}|\mathcal{T}\right\} \\ &= \frac{1}{n^2 h^4} \frac{\sigma^2(t_0)}{f_T^2(t_0)} \frac{1}{\mu_2^2} \text{var}\left\{\sum_{i=1}^n K_i x_i \sum_{k=1}^q \frac{\omega_k}{f(c_k)} \eta_{i,k}|\mathcal{T}\right\}. \quad (19) \end{aligned}$$

Denote $\xi = \sum_{i=1}^n K_i x_i Q_i$, where $Q_i = \sum_{k=1}^q \frac{\omega_k}{f(c_k)} \eta_{i,k}$. Note that $\text{cov}(\eta_{i,k}, \eta_{i,k'}) = \tau_{kk'}$ and $\text{cov}(\eta_{i,k}, \eta_{j,k'}) = 0$, if $i \neq j$. It is easy to obtain $EQ_i^2 = J_\omega(q)$, where

$$J_\omega(q) = \sum_{k=1}^q \sum_{k'=1}^q \frac{\omega_k \omega_{k'} \tau_{kk'}}{f(c_k) f(c_{k'})}.$$

It is easy to obtain that $E(K_i^2 x_i^2) = hf_T(t_0)v_2(1 + o_p(1))$. Therefore:

$$\text{var}(\xi|\mathcal{T}) = nE\xi^2 E(K_1 x_1)^2 = nhf_T(t_0)v_2 J_\omega(q)\{1 + o_p(1)\}.$$

Combined with the result (13), we have:

$$\text{var}\{\hat{m}'_{WQAE}(t_0)\} = \frac{1}{nh^3} \frac{v_2}{\mu_2^2} \frac{\sigma^2(t_0)}{f_T(t_0)} J_\omega(q) + o_p\left(\frac{1}{nh^3}\right).$$

which completes the proof. \square

Proof of Theorem 3. The optimal ω^* can be obtained by solving the following optimization problem:

$$\min J_\omega(q) = \min \sum_{k=1}^q \sum_{k'=1}^q \frac{\omega_k \omega_{k'} \tau_{kk'}}{f(c_k) f(c_{k'})},$$

where ω satisfies $\sum_{k=1}^q \omega_k = 1$ and $\sum_{k=1}^q \omega_k c_k = 0$. With the help of the Lagrange multiplier method, we can obtain the optimal weight vector. This leads to the corresponding minimum conditional variance of \hat{m}'_{WQAE} . \square

6. Conclusions

In this article, we mainly investigated the efficient estimators of the derivative of the nonparametric function in the nonparametric quantile regression model (1). We developed two ways of combining quantile regression information to derive the estimators of $m'(t)$. One is the weighted composite quantile regression estimator based on the quantile weighted loss functions, and the other is the weighted quantile average estimator based on the weighted average of quantile regression estimators at a single quantile. Furthermore, by minimizing the asymptotic variance, the optimal weight vector is computed, and consequently, the optimal estimator can be obtained. Moreover, we conduct some simulations to compare the performance of our proposed estimators to the local linear least square estimator under different symmetric error distributions. Simulation studies illustrate that both the estimators works better than the local linear least square estimator for all the symmetric errors considered except the normal error, and the weighted quantile average estimator performs better than the weighted composite quantile regression estimator in most situations.

Author Contributions: Conceptualization, methodology and formal analysis, X.Z. and X.G.; validation and data analysis, X.Y.; visualization and supervision, Y.Z.; investigation and resources, Y.S.; writing—original draft preparation, X.Z. and X.Y.; writing—review and editing, X.G.; project administration, X.Z., X.G. and Y.Z. All authors have read and agreed to the published version of the manuscript.

Funding: Zhou's work was supported by the Ministry of Education Humanities and Social Sciences Research Youth Foundation (Grant No. 19YJC910011), the Natural Science Foundation of Shandong Province (Grant No. ZR2020MA021) and the Project of Shandong Province Higher Educational Science and Technology Program (Grant No. J18KB099). Yin's research is supported by the Research and Development Project of Dezhou City in China. Shen's research is supported by the Natural Science Foundation of Shandong Province (Grant No. ZR2018LA003) and the Open Research Fund Program of Data Recovery Key Laboratory of Sichuan Province (Grant No. DRN19020).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Some data, models, or code that support the findings of this study are available from the corresponding author upon reasonable request.

Conflicts of Interest: The authors declare no conflict of interest. The funders X.Z., X.Y. and Y.S. have roles in the design of the study, in the analyses and interpretation of data, in the writing of the manuscript, and in the decision to publish the results.

References

1. Dou, X.; Shirahata, S. Comparisons of B-spline procedures with kernel procedures in estimating regression functions and their derivatives. *J. Jpn. Soc. Comput. Stat.* **2009**, *22*, 57–77. [\[CrossRef\]](#)
2. Ruppert, D. Nonparametric Regression and Spline Smoothing. *J. Am. Stat. Assoc.* **2001**, *96*, 1522–1523. [\[CrossRef\]](#)
3. Fan, J.; Gasser, T.; Gijbels, I.; Brockmann, M.; Engel, J. Local Polynomial Regression: Optimal Kernels and Asymptotic Minimax Efficiency. *Ann. Inst. Stat. Math.* **1997**, *49*, 79–99. [\[CrossRef\]](#)
4. Zhang, X.; King, M.L.; Shang, H.L. Bayesian Bandwidth Selection for a Nonparametric Regression Model with Mixed Types of Regressors. *Econometrics* **2016**, *4*, 24. [\[CrossRef\]](#)
5. Souza-Rodrigues, E.A. Nonparametric Regression with Common Shocks. *Econometrics* **2016**, *4*, 36. [\[CrossRef\]](#)
6. Koenker, R.; Bassett, G. Regression quantiles. *Econometrica* **1978**, *46*, 33–50. [\[CrossRef\]](#)
7. Koenker, R. *Quantile Regression*; Cambridge University Press: Cambridge, UK, 2005.
8. Zou, H.; Yuan, M. Composite quantile regression and the oracle Model Selection Theory. *Ann. Stats* **2008**, *36*, 1108–1126. [\[CrossRef\]](#)
9. Kai, B.; Li, R.; Zou, H. Local CQR smoothing: An efficient and safe alternative to local polynomial regression. *J. R. Stat. Soc. Ser. B* **2010**, *72*, 49–69. [\[CrossRef\]](#) [\[PubMed\]](#)
10. Kai, B.; Li, R.; Zou, H. New efficient estimation and variable selection methods for semiparametric varying-coefficient partially linear models. *Ann. Stat.* **2011**, *39*, 305–332. [\[CrossRef\]](#) [\[PubMed\]](#)
11. Jiang, R.; Zhou, Z.; Qian, W. Single index composite quantile regression. *J. Korean Stat. Soc.* **2011**, *3*, 323–332.
12. Ning, Z.; Tang, L. Estimation and test procedures for composite quantile regression with covariates missing at random. *Stat. Probab. Lett.* **2014**, *95*, 15–25. [\[CrossRef\]](#)
13. Jiang, R. Composite quantile regression for linear errors-in-variables models. *Hacet. J. Math. Stat.* **2015**, *44*, 707–713. [\[CrossRef\]](#)
14. Jiang, R.; Qian, W.; Zhou, Z. Single-index composite quantile regression with heteroscedasticity and general error distributions. *Stat. Pap.* **2016**, *57*, 185–203. [\[CrossRef\]](#)
15. Zhang, R.; Lv, Y.; Zhao, W.; Liu, J. Composite quantile regression and variable selection in single-index coefficient model. *J. Stat. Plan. Inference* **2016**, *176*, 1–21. [\[CrossRef\]](#)
16. Zhao, W.; Lian, H.; Song, X. Composite quantile regression for correlated data. *Comput. Stat. Data Anal.* **2016**, *109*, 15–33. [\[CrossRef\]](#)
17. Luo, S.; Zhang, C.; Wang, M. Composite Quantile Regression for Varying Coefficient Models with Response Data Missing at Random. *Symmetry* **2019**, *11*, 1065. [\[CrossRef\]](#)
18. Sun, J.; Gai, Y.; Lin, L. Weighted local linear composite quantile estimation for the case of general error distributions. *J. Stat. Plan. Inference* **2013**, *143*, 1049–1063. [\[CrossRef\]](#)
19. Zhao, Z.; Xiao, Z. Efficient Regressions via Optimally Combining Quantile Information. *Econom. Theory* **2014**, *30*, 1272–1314. [\[CrossRef\]](#) [\[PubMed\]](#)
20. Pollard, D. Asymptotics for least absolute deviation regression estimators. *Econom. Theory* **1991**, *7*, 186–199. [\[CrossRef\]](#)