Supplementary Material for "Local CQR Smoothing: An Efficient and Safe Alternative to Local Polynomial Regression"

Bo Kai [*] Runze Li [†] Hui Zou [‡]

In this supplement note, we derive the asymptotic bias, variance and normality of local CQR estimator when the variance of error is infinite in Section 1. We further study the asymptotic behavior of the local CQR estimate at the boundary. Some simulation results are presented in Section 3.

1 Proof of infinite variance case

Suppose that

$$Y = m(T) + \epsilon,$$

where ϵ has a density f with mean 0 and variance infinity.

Suppose that t_0 is an interior point of the support of $f_T(\cdot)$. Note that the local p-polynomial CQR estimator is constructed by minimizing

$$\sum_{k=1}^{q}\left[\sum_{i=1}^{n}\rho_{\tau_k}\Big\{y_i - a_k - \sum_{j=1}^{p} b_j(t_i - t_0)^j\Big\}K(\frac{t_i - t_0}{h})\right], \tag{1.1}$$

[*] Bo Kai, Department of Statistics, The Pennsylvania State University. E-mail: bokai@psu.edu. Kai's research is supported by NIDA, NIH grants R21 DA024260 and P50 DA10075 as a research assistant. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIDA or the NIH.

[†] Runze Li is Professor of Statistics, Department of Statistics, The Pennsylvania State University. E-mail: rli@stat.psu.edu. Li's research is supported by National Science Foundation grants DMS 0348869 and DMS 0722351.

[‡] Address for correspondence: Hui Zou is Assistant professor of Statistics, School of Statistics, University of Minnesota. E-mail: hzou@stat.umn.edu. Zou's research is supported by National Science Foundation grant DMS 0706733.

and the local p-polynomial CQR estimators of $m(t_0)$ and $m^{(r)}(t_0)$ are given by

$$\hat{m}(t_0) = \frac{1}{q}\sum_{k=1}^{q}\hat{a}_k, \quad \text{and} \quad \hat{m}^{(r)}(t_0) = r!\hat{b}_r, \quad r = 1, \cdots, p. \tag{1.2}$$

The following notation is needed to present the asymptotic properties of the local p-polynomial CQR estimator. Let S_{11} be a $q \times q$ diagonal matrix with diagonal elements $f(c_k)$, $k = 1, \cdots, q$, S_{12} be a $q \times p$ matrix with (k, j)-element being $f(c_k)\mu_j$, $k = 1, \cdots, q$ and $j = 1, \cdots, p$, $S_{21} = S_{12}^T$, and S_{22} be a $p \times p$ matrix with (j, j')-element being $\sum_{k=1}^{q} f(c_k)\mu_{j+j'}$, for $j, j' = 1, \cdots, p$. Similarly, let Σ_{11} be a $q \times q$ matrix with (k, k')-element $\nu_0\tau_{kk'}$, $k, k' = 1, \cdots, q$, Σ_{12} be a $q \times p$ matrix with (k, j)-element being $\nu_j\sum_{k'=1}^{q}\tau_{kk'}$, $k = 1, \cdots, q$ and $j = 1, \cdots, p$, $\Sigma_{21} = \Sigma_{12}^T$, and Σ_{22} be a $p \times p$ matrix with (j, j')-element being $(\sum_{k,k'=1}^{q}\tau_{kk'})\nu_{j+j'}$, for $j, j' = 1, \cdots, p$. Define

$$S = \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix}, \quad \text{and} \quad \Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}.$$

Partition S^{-1} into four submatrices as follows

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

where and hereafter, we use $(\cdot)_{11}$ to denote the left-top $q \times q$ submatrix and use $(\cdot)_{22}$ to denote the right-bottom $p \times p$ submatrix.

Furthermore, let $u_k = \sqrt{nh}\{a_k - m(t_0) - c_k\}$, $v_j = h^j\sqrt{nh}\{j!b_j - m^{(j)}(t_0)\}/j!$. Let $x_i = (t_i - t_0)/h$, $K_i = K(x_i)$ and $\Delta_{i,k} = \dfrac{u_k}{\sqrt{nh}} + \sum_{j=1}^{p}\dfrac{v_j x_i^j}{\sqrt{nh}}$. Write $r_{i,p} = m(t_i) - \sum_{j=0}^{p} m^{(j)}(t_0)(t_i - t_0)^j/j!$. Define $\eta_{i,k}^*$ to be $I(\epsilon_i \le c_k - r_{i,p}) - \tau_k$. let $W_n^* = (w_{11}^*, \cdots, w_{1q}^*, w_{21}^*, \cdots, w_{2p}^*)^T$ with $w_{1k}^* = \frac{1}{\sqrt{nh}}\sum_{i=1}^{n} K_i\eta_{i,k}^*$ and $w_{2j}^* = \frac{1}{\sqrt{nh}}\sum_{k=1}^{q}\sum_{i=1}^{n} K_i x_i^j \eta_{i,k}^*$. The asymptotic properties of the local p-polynomial CQR estimator are based on the following theorem.

Theorem 1.1. *Denote $\hat{\theta}_n = (\hat{u}_1, \cdots, \hat{u}_q, \hat{v}_1, \cdots, \hat{v}_p)$ be the minimizer of (1.1). Assume that $f_T(t_0) > 0$, $f_T(\cdot)$ and $m^{(p+2)}(\cdot)$ are continuous in a neighborhood of t_0, and $f(\cdot)$ is positive in the neighborhoods of $\{\tau_k\}$. If $h \to 0$ and $nh \to \infty$, then we have*

$$\hat{\theta}_n + \frac{1}{f_T(t_0)}S^{-1}E(W_n^*|\mathbf{T}) \xrightarrow{\mathcal{L}} MVN(\mathbf{0}, \frac{1}{f_T(t_0)}S^{-1}\Sigma S^{-1}).$$

Proof. To apply the identity

$$\rho_\tau(x - y) - \rho_\tau(x) = y(I(x \le 0) - \tau) + \int_0^y [I(x \le z) - I(x \le 0)]dz, \tag{1.3}$$

we write

$$
\begin{aligned}
y_i - a_k - \sum_{j=1}^{p} b_j(t_i - t_0)^j &= \epsilon_i + m(t_i) - a_k - \sum_{j=1}^{p} b_j(t_i - t_0)^j \\
&= (\epsilon_i - c_k) + r_{i,p} - \frac{u_k}{\sqrt{nh}} - \sum_{j=1}^{p} \frac{v_j x_i^j}{\sqrt{nh}} \\
&= (\epsilon_i - c_k) + r_{i,p} - \Delta_{i,k} \ ,
\end{aligned}
$$

Minimizing (1.1) is equivalent to minimizing

$$
L_n(\theta) = \sum_{i=1}^{n} \left\{ K_i \sum_{k=1}^{q} \Big[\rho_{\tau_k}\big((\epsilon_i - c_k) + r_{i,p} - \Delta_{i,k}\big) - \rho_{\tau_k}\big((\epsilon_i - c_k) + r_{i,p}\big) \Big] \right\}.
$$

Using the identity (1.3) and with some straightforward calculations, it follows that

$$
\begin{aligned}
L_n(\theta) &= \sum_{i=1}^{n} \left\{ K_i \sum_{k=1}^{q} \Delta_{i,k} \left[I(\epsilon_i \le c_k - r_{i,p}) - \tau_k \right] \right\} \\
&\quad + \sum_{i=1}^{n} \left\{ K_i \sum_{k=1}^{q} \int_0^{\Delta_{i,k}} \left[I(\epsilon_i \le c_k - r_{i,p} + z) - I(\epsilon_i \le c_k - r_{i,p}) \right] dz \right\} \\
&= \sum_{k=1}^{q} u_k \left(\sum_{i=1}^{n} \frac{K_i \eta_{i,k}^*}{\sqrt{nh}} \right) + \sum_{j=1}^{p} v_j \left(\sum_{k=1}^{q} \sum_{i=1}^{n} \frac{K_i x_i^j \eta_{i,k}^*}{\sqrt{nh}} \right) + \sum_{k=1}^{q} B_{n,k}(\theta),
\end{aligned}
$$

where

$$
B_{n,k}(\theta) = \sum_{i=1}^{n} \left\{ K_i \int_0^{\Delta_{i,k}} \Big[I(\epsilon_i \le c_k - r_{i,p} + z) - I(\epsilon_i \le c_k - r_{i,p}) \Big] dz \right\}.
$$

Let $S_{n,11}$ be a $q \times q$ diagonal matrix with diagonal elements $f(c_k) \sum_{i=1}^{n} K_i / nh$, $k = 1, \cdots, q$; $S_{n,12}$ be a $q \times p$ matrix with (k, j)-element $f(c_k) \sum_{i=1}^{n} K_i x_i^j / nh$, $j = 1, \cdots, p$; $S_{n,22}$ be a $p \times p$ matrix with (j, j') element $\sum_{k=1}^{q} f(c_k) \sum_{i=1}^{n} K_i x_i^{j+j'} / nh$. Denote

$$
S_n = \begin{pmatrix} S_{n,11} & S_{n,12} \\ S_{n,12}^T & S_{n,22} \end{pmatrix}.
$$

We write $L_n(\theta)$ as

$$
\begin{aligned}
L_n(\theta) &= \sum_{k=1}^{q} u_k \left(\sum_{i=1}^{n} \frac{K_i \eta_{i,k}^*}{\sqrt{nh}} \right) + \sum_{j=1}^{p} v_j \left(\sum_{k=1}^{q} \sum_{i=1}^{n} \frac{K_i x_i^j \eta_{i,k}^*}{\sqrt{nh}} \right) \\
&\quad + \sum_{k=1}^{q} E_\epsilon [B_{n,k}(\theta) | \mathbf{T}] + \sum_{k=1}^{q} R_{n,k}(\theta),
\end{aligned}
$$

where $R_{n,k}(\theta) = B_{n,k}(\theta) - E_\epsilon[B_{n,k}(\theta)|\mathbf{T}]$.

By similar arguments, we can show that $\sum_{k=1}^q E_\epsilon[B_{n,k}(\theta)|\mathbf{T}] = \frac{1}{2}\theta^T S_n \theta + o_p(1)$ and $R_{n,k}(\theta) = o_p(1)$. Together with $\sum_{i=1}^n K_i x_i^j/nh \xrightarrow{P} f_T(t_0)\mu_j$, and

$$S_n \xrightarrow{P} f_T(t_0)S = f_T(t_0)\begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix},$$

we have

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

Since the convex function $L_n(\theta) - (W_n^*)^T\theta$ converges in probability to the convex function $\dfrac{f_T(t_0)}{2}\theta^T S\theta$, it follows from the convexity lemma that for any compact set Θ, the quadratic approximation to $L_n(\theta)$ holds uniformly for θ in any compact set, which leads to

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

Denote $\eta_{i,k} = I(\epsilon_i \le c_k) - \tau_k$ and $W_n = (w_{11}, \cdots, w_{1q}, w_{21}, \cdots, w_{2p})^T$ with $w_{1k} = \dfrac{1}{\sqrt{nh}}\displaystyle\sum_{i=1}^n K_i\eta_{i,k}$ and $w_{2j} = \dfrac{1}{\sqrt{nh}}\displaystyle\sum_{k=1}^q\sum_{i=1}^n K_i x_i^j \eta_{i,k}$. By the Cramer-Wald theorem, it is easy to see that the CLT for $W_n|\mathbf{T}$ holds

$$\frac{W_n|\mathbf{T} - E[W_n|\mathbf{T}]}{\sqrt{Var[W_n|\mathbf{T}]}} \xrightarrow{\mathcal{L}} MVN(\mathbf{0}, I_{(p+q)\times(p+q)}). \tag{1.4}$$

Note that

$$Cov(\eta_{i,k}, \eta_{i,k'}) = \tau_{kk'}, \qquad Cov(\eta_{i,k}, \eta_{j,k'}) = 0, \quad if \quad i \neq j.$$

and $\sum_{i=1}^n K_i^2 x_i^j/nh \xrightarrow{P} f_T(t_0)\nu_j$, Therefore, $Var[W_n|\mathbf{T}] \xrightarrow{P} f_T(t_0)\Sigma$. Combined with (1.4), we have

$$W_n|\mathbf{T} \xrightarrow{\mathcal{L}} MVN(\mathbf{0}, f_T(t_0)\Sigma).$$

Moreover, we have $Var(w_{1k}^* - w_{1k}|\mathbf{T}) = \frac{1}{nh}\sum_{i=1}^n K_i^2 Var(\eta_{i,k}^* - \eta_{i,k}) \le \frac{1}{nh}\sum_{i=1}^n K_i^2\{F(c_k + |r_{i,p}|) - F(c_k)\} = o_p(1)$ and also $Var(w_{2j}^* - w_{2j}|\mathbf{T}) = \frac{1}{nh}\sum_{i=1}^n K_i^2 x_i^j Var(\sum_{k=1}^q \eta_{i,k}^* - \eta_{i,k}) \le \frac{q^2}{nh}\sum_{i=1}^n K_i^2 x_i^j \max_k\{F(c_k + |r_{i,p}|) - F(c_k)\} = o_p(1)$, thus

$$Var(W_n^* - W_n|\mathbf{T}) = o_p(1).$$

So by Slutsky's theorem, conditioning on $\mathbf{T}$, we have

$$W_n^*|\mathbf{T} - E(W_n^*|\mathbf{T}) \xrightarrow{\mathcal{L}} MVN(\mathbf{0}, f_T(t_0)\Sigma).$$

Therefore,

$$\hat{\theta}_n + \frac{1}{f_T(t_0)} S^{-1} E(W_n^*|\mathbf{T}) \xrightarrow{\mathcal{L}} MVN(\mathbf{0}, \frac{1}{f_T(t_0)} S^{-1} \Sigma S^{-1}). \tag{1.5}$$

This completes the proof. □

The asymptotic properties of local CQR estimators $\hat{m}(t)$ and $\hat{m}'(t)$ are two special cases of the general result.

Theorem 1.2. *Under the regularity conditions in Theorem 1.1, the asymptotic conditional bias and variance of the local linear CQR estimator $\hat{m}(t_0)$ are given by*

$$Bias(\hat{m}(t_0)|\mathbf{T}) = \frac{1}{2} m''(t_0) \mu_2 h^2 + o_p(h^2), \tag{1.6}$$

$$Var(\hat{m}(t_0)|\mathbf{T}) = \frac{1}{nh} \frac{\nu_0}{f_T(t_0)} R_1(q) + o_p(\frac{1}{nh}). \tag{1.7}$$

Furthermore, conditioning on $\mathbf{T}$, we have

$$\sqrt{nh}\{\hat{m}(t_0) - m(t_0) - \frac{1}{2} m''(t_0)\mu_2 h^2\} \xrightarrow{\mathcal{L}} N\left(0, \frac{\nu_0}{f_T(t_0)} R_1(q)\right). \tag{1.8}$$

Proof of Theorem 1.2. The asymptotic normality follows Theorem 1.1 with $p = 1$. Let us calculate the conditional bias and variance, respectively. Denote by $e_{q\times 1}$ the vector that contains q 1's. When $p = 1$, S is a diagonal matrix with diagonal elements $f(c_1), \cdots, f(c_q), \mu_2 \sum_{k=1}^q f(c_k)$. So the asymptotic conditional bias of $\hat{m}(t_0) = \frac{1}{q}\sum_{k=1}^q \hat{a}_k$ is

$$\begin{aligned}
Bias(\hat{m}(t_0)|\mathbf{T}) &= \frac{1}{q}\sum_{k=1}^q c_k - \frac{1}{q\cdot\sqrt{nh}} \frac{1}{f_T(t_0)} e_{q\times 1}^T (S^{-1})_{11} E(W_{1n}^*|\mathbf{T}) \\
&= \frac{1}{q}\sum_{k=1}^q c_k - \frac{1}{q\cdot nh} \frac{1}{f_T(t_0)} \sum_{i=1}^n K_i \sum_{k=1}^q \frac{1}{f(c_k)} \{F(c_k - r_{i,p}) - F(c_k)\} \\
&= \frac{1}{nh} \frac{1}{f_T(t_0)} \sum_{i=1}^n K_i r_{i,p} \{1 + o_p(1)\}.
\end{aligned}$$

By using the fact that

$$\frac{1}{nh}\sum_{i=1}^n K_i r_{i,p} = \frac{f_T(t_0) m''(t_0)}{2} \mu_2 h^2 \{1 + o_p(1)\},$$

we obtain

$$Bias(\hat{m}(t_0)|\mathbf{T}) = \frac{1}{2} m''(t_0)\mu_2 h^2 + o_p(h^2). \tag{1.9}$$

Furthermore, the conditional variance of $\hat{m}(t_0)$ is

$$\begin{aligned} Var(\hat{m}(t_0)|\mathbf{T}) &= \frac{1}{nh}\frac{1}{f_T(t_0)}\frac{1}{q^2}e_{q\times 1}^T(S^{-1}\Sigma S^{-1})_{11}e_{q\times 1} + o_p(\frac{1}{nh}) \\ &= \frac{1}{nh}\frac{\nu_0}{f_T(t_0)}R_1(q) + o_p(\frac{1}{nh}), \end{aligned} \tag{1.10}$$

which completes the proof. □

Theorem 1.3. *Under the regularity conditions in Theorem 1.1, the asymptotic conditional bias and variance of $\hat{m}'(t_0)$ from local quadratic CQR is given by*

$$Bias(\hat{m}'(t_0)|\mathbf{T}) = \frac{1}{6}m'''(t_0)\frac{\mu_4}{\mu_2}h^2 + o_p(h^2), \tag{1.11}$$

$$Var(\hat{m}'(t_0)|\mathbf{T}) = \frac{1}{nh^3}\frac{\nu_2}{\mu_2^2 f_T(t_0)}R_2(q) + o_p(\frac{1}{nh^3}). \tag{1.12}$$

Furthermore, conditioning on $\mathbf{T}$*, we have the following asymptotic normal distribution*

$$\sqrt{nh^3}\left(\hat{m}'(t_0) - m'(t_0) - \frac{1}{6}m'''(t_0)\frac{\mu_4}{\mu_2}h^2\right) \xrightarrow{\mathcal{L}} N\left(0, \frac{\nu_2}{\mu_2^2 f_T(t_0)}R_2(q)\right). \tag{1.13}$$

Proof of Theorem 1.3. We apply Theorem 1.1 to get the asymptotic normality. Denote by e_r the p-vector $(0, 0, \cdots, 1, 0, \cdots, 0)^T$ with 1 on the r^{th} position. When $p = 2$, S_{12} and S_{22} have the following forms

$$S_{12} = \begin{pmatrix} \mathbf{0}_{q\times 1} & \mu_2\Big(f(c_k)\Big)_{q\times 1} \end{pmatrix}, S_{22} = \begin{pmatrix} \mu_2\sum_{k=1}^q f(c_k) & 0 \\ 0 & \mu_4\sum_{k=1}^q f(c_k) \end{pmatrix}.$$

Thus,

$$(S^{-1})_{22} = (S_{22} - S_{21}S_{11}^{-1}S_{12})^{-1} = \begin{pmatrix} \frac{1}{\mu_2\sum_{k=1}^q f(c_k)} & 0 \\ 0 & \frac{1}{(\mu_4-\mu_2^2)\sum_{k=1}^q f(c_k)} \end{pmatrix},$$

$$(S^{-1})_{21} = -(S^{-1})_{22}S_{21}S_{11}^{-1} = \begin{pmatrix} \mathbf{0}_{1\times q} \\ \left(\frac{\mu_2}{(\mu_4-\mu_2^2)\sum_{k=1}^q f(c_k)}\right)_{1\times q} \end{pmatrix},$$

since $S_{11} = \operatorname{diag}(f(c_1), \cdots, f(c_q))$. By Theorem 1.1

$$\begin{aligned} Bias(\hat{m}'(t_0)|\mathbf{T}) &= -\frac{1}{hf_T(t_0)}\frac{1}{\sqrt{nh}}e_1^T\left\{(S^{-1})_{21}E(W_{1n}^*|\mathbf{T}) + (S^{-1})_{22}E(W_{2n}^*|\mathbf{T})\right\} \\ &= -\frac{1}{hf_T(t_0)}\frac{1}{\mu_2\sum_{k=1}^q f(c_k)}\frac{1}{\sqrt{nh}}E(w_{21}^*|\mathbf{T}). \end{aligned}$$

Note that

$$E(w_{2j}^*|\mathbf{T}) = \frac{1}{\sqrt{nh}} \sum_{i=1}^{n} K_i x_i^j \sum_{k=1}^{q} \{F(c_k - r_{i,p}) - F(c_k)\}$$

Therefore, $Bias(\hat{m}'(t_0)|\mathbf{T})$ is equal to $\frac{1}{\cdot nh^2} \frac{1}{f_T(t_0)} \sum_{i=1}^{n} K_i x_i r_{i,p}\{1 + o_p(1)\}$. Still using the fact that with $p = 2$

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

we obtain

$$Bias(\hat{m}'(t_0)|\mathbf{T}) = \frac{1}{6} m'''(t_0) \frac{\mu_4}{\mu_2} h^2 + o_p(h^2). \tag{1.14}$$

Furthermore, the conditional variance of $\hat{m}(t_0)$ is

$$\begin{aligned} Var(\hat{m}'(t_0)|\mathbf{T}) &= \frac{1}{nh^3} \frac{1}{f_T(t_0)} e_1^T (S^{-1} \Sigma S^{-1})_{22} e_1 + o_p(\frac{1}{nh^3}), \\ &= \frac{1}{nh^3} \frac{\nu_2}{\mu_2^2 f_T(t_0)} R_2(q) + o_p(\frac{1}{nh^3}). \end{aligned} \tag{1.15}$$

which completes the proof. □

2 Asymptotic boundary behavior of local CQR estimators

Back to the general nonparametric regression model

$$Y = m(T) + \sigma(T)\epsilon, \tag{2.1}$$

Now we study the behavior of the estimator at the boundary of the support of T. Without loss of generality, assume $f_T(\cdot)$ has support on $[0, 1]$. We consider the left boundary point $t = ch$, where c is a positive constant. let

$$\mu_j(c) = \int_{-c}^{\infty} u^j K(u) du \quad \nu_j(c) = \int_{-c}^{\infty} u^j K^2(u) du, \quad j = 0, 1, 2, \ldots$$

We first establish asymptotic theory of the local p-polynomial CQR estimators at $t = ch$, and then discuss the special case of $p = 1$ and 2.

The following notation is needed to present the asymptotic properties of the local p-polynomial CQR estimator. Let $S_{11}(c)$ be a $q \times q$ diagonal matrix with diagonal elements $f(c_k)$, $k = 1, \cdots, q$, $S_{12}(c)$ be a $q \times p$ matrix with (k, j)-element being $f(c_k)\mu_j(c)$, $k = 1, \cdots, q$ and $j = 1, \cdots, p$, $S_{21}(c) = S_{12}^T(c)$, and $S_{22}(c)$ be a $p \times p$ matrix with (j, j')-element being $\sum_{k=1}^{q} f(c_k)\mu_{j+j'}(c)$, for $j, j' = 1, \cdots, p$. Similarly, Let $\Sigma_{11}(c)$ be a $q \times q$ matrix with (k, k')-element $\nu_0(c)\tau_{kk'}$, $k, k' = 1, \cdots, q$, $\Sigma_{12}(c)$ be a $q \times p$ matrix with (k, j)-element being $\nu_j(c)\sum_{k'=1}^{q} \tau_{kk'}$, $k = 1, \cdots, q$ and $j = 1, \cdots, p$, $\Sigma_{21}(c) = \Sigma_{12}^T(c)$, and $\Sigma_{22}(c)$ be a $p \times p$ matrix with (j, j')-element being $(\sum_{k,k'=1}^{q} \tau_{kk'})\nu_{j+j'}(c)$, for $j, j' = 1, \cdots, p$. Define

$$S(c) = \begin{pmatrix} S_{11}(c) & S_{12}(c) \\ S_{21}(c) & S_{22}(c) \end{pmatrix}, \quad \text{and} \quad \Sigma(c) = \begin{pmatrix} \Sigma_{11}(c) & \Sigma_{12}(c) \\ \Sigma_{21}(c) & \Sigma_{22}(c) \end{pmatrix}.$$

Partition $S^{-1}(c)$ into four submatrices as follows

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

where and hereafter, we use $(\cdot)_{11}$ to denote the left-top $q \times q$ submatrix and use $(\cdot)_{22}$ to denote the right-bottom $p \times p$ submatrix.

Furthermore, let $u_k = \sqrt{nh}\{a_k - m(t) - \sigma(t)c_k\}$, $v_j = h^j\sqrt{nh}\{j!b_j - m^{(j)}(t)\}/j!$. Let $x_i = (t_i - t)/h$, $K_i = K(x_i)$ and $\Delta_{i,k} = \dfrac{u_k}{\sqrt{nh}} + \sum_{j=1}^{p} \dfrac{v_j x_i^j}{\sqrt{nh}}$. Write $d_{i,k} = c_k[\sigma(t_i) - \sigma(t)] + r_{i,p}$ with $r_{i,p} = m(t_i) - \sum_{j=0}^{p} m^{(j)}(t)(t_i - t)^j/j!$. Define $\eta_{i,k}^*$ to be $I(\epsilon_i \le c_k - \frac{d_{i,k}}{\sigma(t_i)}) - \tau_k$. let $W_n^* = (w_{11}^*, \cdots, w_{1q}^*, w_{21}^*, \cdots, w_{2p}^*)^T$ with $w_{1k}^* = \frac{1}{\sqrt{nh}}\sum_{i=1}^{n} K_i \eta_{i,k}^*$ and $w_{2j}^* = \frac{1}{\sqrt{nh}}\sum_{k=1}^{q}\sum_{i=1}^{n} K_i x_i^j \eta_{i,k}^*$.

Theorem 2.1. *Denote $\hat{\theta}_n = (\hat{u}_1, \cdots, \hat{u}_q, \hat{v}_1, \cdots, \hat{v}_p)$ be the minimizer of (1.1). Assume that $f_T(0+) > 0$, $f_T(\cdot), m^{(p+1)}(\cdot)$ and $\sigma^2(\cdot)$ are right continuous at the point 0, and $f(\cdot)$ is positive in the neighborhoods of $\{\tau_k\}$. If $h \to 0$ and $nh \to \infty$, then we have*

$$\hat{\theta}_n + \frac{\sigma(0+)}{f_T(0+)} S^{-1}(c) E(W_n^* | \mathbf{T}) \xrightarrow{\mathcal{L}} MVN(\mathbf{0}, \frac{\sigma^2(0+)}{f_T(0+)} S^{-1}(c)\Sigma(c)S^{-1}(c)).$$

The proof is quite similar to the one for interior points, so we omit it here. Now let's look at the asymptotic behavior of local CQR estimators $\hat{m}(t)$ and $\hat{m}'(t)$ at the boundary.

Theorem 2.2. *Under the regularity conditions in Theorem 2.1, if the error distribution is symmetric, then the asymptotic conditional bias and variance of the local linear CQR estimator $\hat{m}(t)$ are given by*

$$Bias(\hat{m}(t)|\mathbf{T}) = \frac{1}{2} a(c) m''(0+) h^2 + o_p(h^2), \tag{2.2}$$

$$Var(\hat{m}(t)|\mathbf{T}) = \frac{1}{nh} \frac{b(c)\sigma^2(0+)}{f_T(0+)} R_1(q) + o_p(\frac{1}{nh}). \tag{2.3}$$

where $a(c) = \dfrac{\mu_2^2(c) - \mu_1(c)\mu_3(c)}{\mu_0(c)\mu_2(c) - \mu_1^2(c)}$ *and* $b(c) = \dfrac{\mu_2^2(c)\nu_0(c) - 2\mu_1(c)\mu_2(c)\nu_1(c) + \mu_1^2(c)\nu_2(c)}{\{\mu_0(c)\mu_2(c) - \mu_1^2(c)\}^2}$.

Furthermore, conditioning on $\mathbf{T}$*, we have*

$$\sqrt{nh}\{\hat{m}(t) - m(t) - \frac{1}{2}a(c)m''(0+)h^2\} \xrightarrow{\mathcal{L}} N\left(0, \frac{b(c)\sigma^2(0+)}{f_T(0+)}R_1(q)\right). \tag{2.4}$$

Proof of Theorem 2.2. The asymptotic normality follows Theorem 2.1 with $p = 1$. Let us calculate the conditional bias and variance, respectively. Denote by $e_{q\times 1}$ the vector that contains q 1's. The asymptotic conditional bias of $\hat{m}(t) = \frac{1}{q}\sum_{k=1}^{q} \hat{a}_k$ is

$$Bias(\hat{m}(t)|\mathbf{T}) \quad = \quad \frac{1}{q}\sigma(t)\sum_{k=1}^{q} c_k - \frac{1}{q\cdot\sqrt{nh}}\frac{\sigma(0+)}{f_T(0+)}(e_{q\times 1}^T\ 0)S^{-1}E(W_n^*|\mathbf{T})$$

Note that the error is symmetric, thus $\sum_{k=1}^{q} c_k = 0$, and similarly we can show that

$$E(w_{1k}^*|\mathbf{T}) = f(c_k)\frac{f_T(0+)m''(0+)}{2\sigma(0+)}\mu_2(c)h^2\{1 + o_p(1)\} \quad k = 1, \cdots, q,$$

and

$$E(w_{21}^*|\mathbf{T}) = \{\sum_{k=1}^{q} f(c_k)\}\frac{f_T(0+)m''(0+)}{2\sigma(0+)}\mu_3(c)h^2\{1 + o_p(1)\}.$$

Therefore,

$$\begin{aligned} Bias(\hat{m}(t)|\mathbf{T}) \quad &= \quad -\frac{1}{q\cdot\sqrt{nh}}\frac{\sigma(t_0)}{f_T(t_0)}(e_{q\times 1}^T\ 0)S^{-1}E(W_n^*|\mathbf{T}) \\ &= \quad \frac{1}{2}\frac{\mu_2^2(c) - \mu_1(c)\mu_3(c)}{\mu_0(c)\mu_2(c) - \mu_1^2(c)}m''(0+)h^2 + o_p(h^2) \\ &= \quad \frac{1}{2}a(c)m''(0+)h^2 + o_p(h^2). \end{aligned}$$

Furthermore, the conditional variance of $\hat{m}(t_0)$ is

$$\begin{aligned} Var(\hat{m}(t)|\mathbf{T}) \quad &= \quad \frac{1}{nh}\frac{\sigma^2(0+)}{f_T(0+)}\frac{1}{q^2}e_{q\times 1}^T(S^{-1}\Sigma S^{-1})_{11}e_{q\times 1} + o_p(\frac{1}{nh}) \\ &= \quad \frac{1}{nh}\frac{\sigma^2(0+)}{f_T(0+)}\frac{\mu_2^2(c)\nu_0(c) - 2\mu_1(c)\mu_2(c)\nu_1(c) + \mu_1^2(c)\nu_2(c)}{\{\mu_0(c)\mu_2(c) - \mu_1^2(c)\}^2}R_1(q) + o_p(\frac{1}{nh}) \\ &= \quad \frac{1}{nh}\frac{b(c)\sigma^2(0+)}{f_T(0+)}R_1(q) + o_p(\frac{1}{nh}), \end{aligned} \tag{2.5}$$

which completes the proof. □

From Theorem 2.2, it can be seen that the leading team of the asymptotic bias of the local linear CQR estimator is the same as that of the local linear LS estimator. This relationship is the same as that when x is an interior point. Furthermore, the relationship between the asymptotic variances of the local CQR and that of LS estimators at boundary is also the same as that for interior points, i.e., they are different by the factor R_2. Thus, Theorem 2.2 clearly indicates that the local CQR estimator shares the property of the automatical boundary correction, a nice property of local linear least squares estimator.

Theorem 2.3. *Under the regularity conditions in Theorem 2.1, if the error distribution is symmetric, then the asymptotic conditional bias and variance of the local quadratic CQR estimator $\hat{m}'(t)$ are given by*

$$Bias(\hat{m}'(t)|\mathbf{T}) = \frac{1}{2}a^*(c)m''(0+)h^2 + o_p(h^2), \tag{2.6}$$

$$Var(\hat{m}'(t)|\mathbf{T}) = \frac{1}{nh^3}\frac{b^*(c)\sigma^2(0+)}{f_T(0+)}R_2(q) + o_p(\frac{1}{nh^3}), \tag{2.7}$$

where $a^*(c)$ and $b^*(c)$ are constants that depend only on c and the kernel K.

Furthermore, conditioning on $\mathbf{T}$, we have

$$\sqrt{nh^3}\{\hat{m}'(t) - m(t) - \frac{1}{6}a^*(c)m'''(0+)h^2\} \xrightarrow{\mathcal{L}} N\left(0, \frac{b^*(c)\sigma^2(0+)}{f_T(0+)}R_2(q)\right). \tag{2.8}$$

Proof of Theorem 2.3. We apply Theorem 2.1 to get the asymptotic normality. Denote by e_r the p-vector $(0, 0, \cdots, 1, 0, \cdots, 0)^T$ with 1 on the r^{th} position. When $p = 2$, we have

$$E(w_{1k}^*|\mathbf{T}) = f(c_k)\frac{f_T(0+)m'''(0+)}{6\sigma(0+)}\mu_2(c)h^3\{1 + o_p(1)\} \quad k = 1, \cdots, q,$$

and

$$E(w_{2j}^*|\mathbf{T}) = \{\sum_{k=1}^{q} f(c_k)\}\frac{f_T(0+)m'''(0+)}{6\sigma(0+)}\mu_{2+j}(c)h^3\{1 + o_p(1)\} \quad j = 1, 2.$$

Therefore,

$$\begin{aligned} Bias(\hat{m}'(t)|\mathbf{T}) &= -\frac{\sigma(0+)}{hf_T(0+)}\frac{1}{\sqrt{nh}}e_1^T\left\{(S^{-1})_{21}E(W_{1n}^*|\mathbf{T}) + (S^{-1})_{22}E(W_{2n}^*|\mathbf{T})\right\} \\ &= \frac{1}{6}a^*(c)m'''(0+)h^2 + o_p(h^2). \end{aligned}$$

Furthermore, the conditional variance of $\hat{m}'(t)$ is

$$\begin{aligned} Var(\hat{m}'(t)|\mathbf{T}) &= \frac{1}{nh^3}\frac{\sigma^2(0+)}{f_T(0+)}e_1^T(S^{-1}\Sigma S^{-1})_{22}e_1 + o_p(\frac{1}{nh^3}), \\ &= \frac{1}{nh^3}\frac{b^*(c)\sigma^2(0+)}{f_T(0+)}R_2(q) + o_p(\frac{1}{nh^3}). \end{aligned} \tag{2.9}$$

This completes the proof. □

From Theorem 2.3, it can be seen that the asymptotic bias of local CQR estimator at boundary is of order h^2, and its asymptotic variance is of order $1/nh^3$. Thus, the orders of the asymptotic bias and variance are the same as those of local quadratic regression. Thus, the local quadratic CQR estimator possesses the property of automatic boundary correction.

3 Simulation studies

In this section, we provide two simulation examples. The first example is to demonstrate the performance of local CQR estimate when the error follows a Cauchy distribution. The second one is to compare the boundary behavior of the local CQR estimator and the local least squares estimator.

Example 1 (Infinite error variance). We generated 400 data set, each consisting of $n = 200$ observations, from

$$Y = \sin(2T) + 2\exp(-16T^2) + 0.5\epsilon, \tag{3.1}$$

where T follows $N(0, 1)$. In our simulation, the error ϵ follows the Cauchy distribution. Thus, the error variance is infinite. For the local polynomial CQR estimator, we consider $q = 5$, 9 and 19, and estimate $m(\cdot)$ and $m'(\cdot)$ over $[-1.5, 1.5]$. The mean and standard deviation of RASE over 400 simulations are summarized in Table 1. To see how the proposed estimate behaves at a typical point, Table 1 also depicts the biases and standard deviations of $\hat{m}(t)$ and $\hat{m}'(t)$ at $t = 0.75$. In Table 1, CQR_5, CQR_9 and CQR_{19} correspond to the local CQR estimate with $q = 5$, 9 and 19, respectively. From the Table 1, we can see that the RASE of local CQR estimate is much less than that of local LS estimate. This is because the local LS estimator is not a consistent estimator for the regression function, while the local CQR estimator is. This is also evidenced from the standard deviation of the local estimator at $t = 0.75$.

Table 1: Simulation Results for Example 1

	$\hat{m}$			$\hat{m}'$		
	RASE	$t = 0.75$		**RASE**	$t = 0.75$	
	Mean(SD)	**Bias**	**Std**	**Mean(SD)**	**Bias**	**Std**
Cauchy						
LS	—	-0.0881	7.8740	—	5.1324	87.7494
CQR_5	$10228.96_{(125981.63)}$	-0.0241	0.2965	$14386.19_{(160902.10)}$	0.0716	1.5997
CQR_9	$4798.64_{(51545.41)}$	-0.0713	0.9690	$14243.84_{(158913.53)}$	0.0686	1.6133
CQR_{19}	$1120.90_{(12889.99)}$	-0.0929	1.2995	$14224.16_{(159441.96)}$	0.0727	1.6064

Example 2. We generated 400 data set, each consisting of $n = 200$ observations, from

$$Y = \sin(2T) + 2\exp(-16T^2) + 0.5\epsilon, \tag{3.2}$$

where T follows $N(0,1)$. In our simulation, the error ϵ follows $0.95N(0,1) + 0.05N(0,10^2)$. Figure 1 depicts the 400 estimated coefficient functions of CQR_9 for all 400 simulations. Results for CQR_5 and CQR_{19} is similar, so we opt not to present here. Figure 2 depicts the plots of estimate of the regression function and its derivative based on a typical data set. From Figures 1 and 2, it can be clearly seen that the local CQR estimator improve over the local least squares estimator for both interior and boundary points.

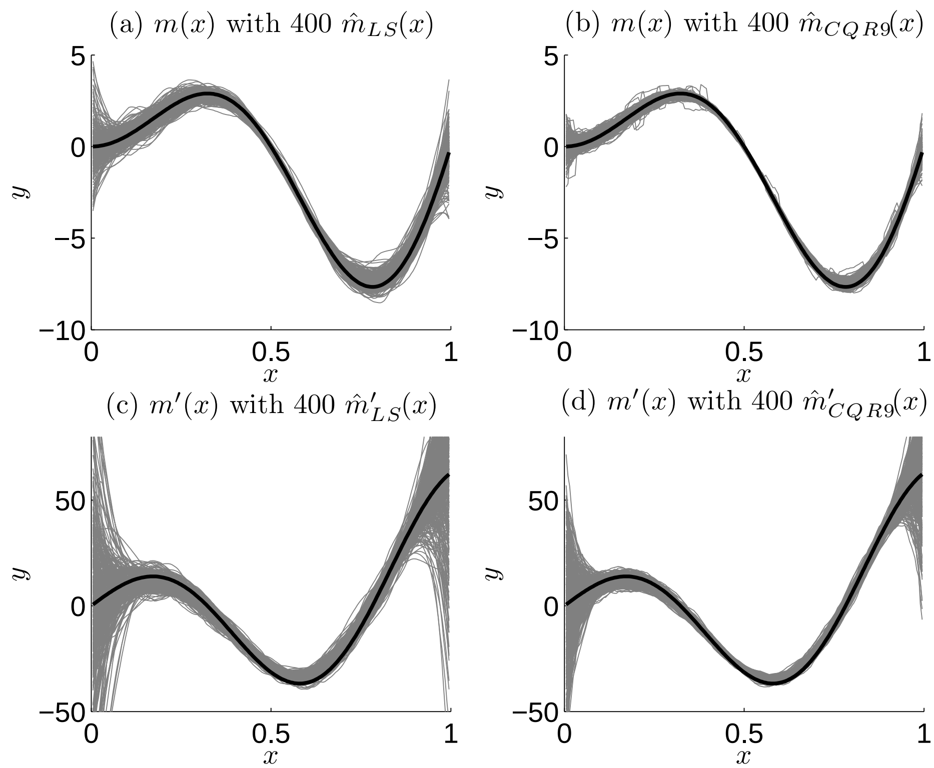

Figure 1: (a) and (c) are plots of 400 local least squares estimators of $m(\cdot)$ and $m'(\cdot)$ over 400 simulation, respectively. (b) and (d) are plots of 400 local CQR estimators of $m(\cdot)$ and $m'(\cdot)$, respectively.

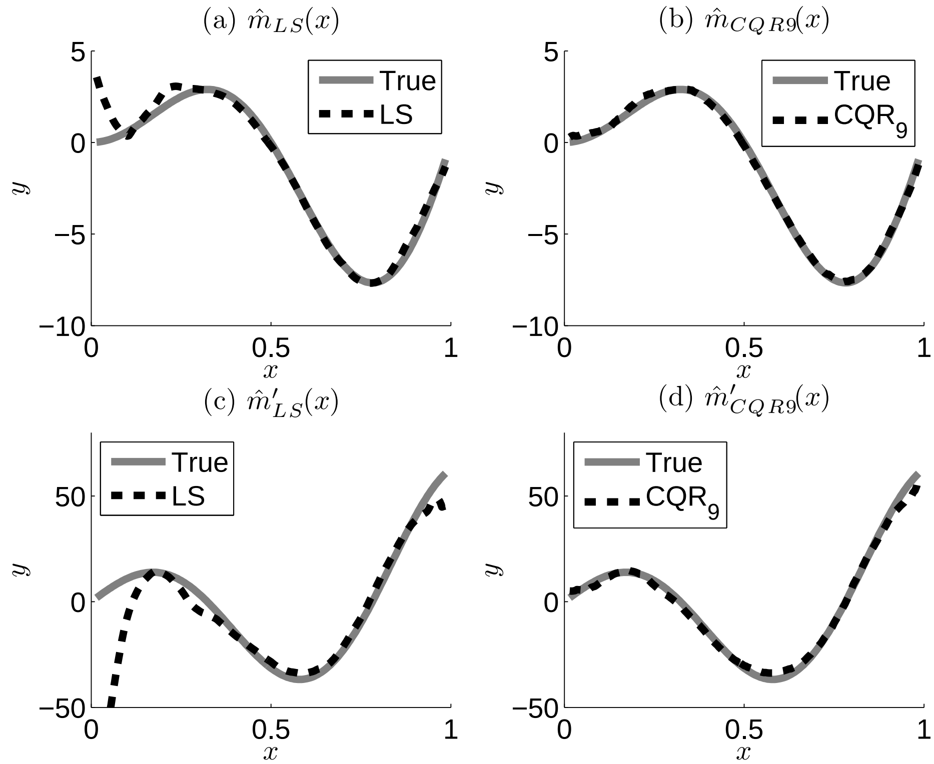

Figure 2: (a) and (c) are plots of a typical local least squares estimators of $m(\cdot)$ and $m'(\cdot)$, respectively. (b) and (d) are plots of a typical local CQR estimators of $m(\cdot)$ and $m'(\cdot)$, respectively.