ANNALS OF ECONOMICS AND FINANCE **6**, 105–127 (2005)

Semiparametric Efficient Estimation of Partially Linear Quantile Regression Models

Yiguo Sun

Department of Economics, University of Guelph, Guelph, ON, Canada N1G 2W1
E-mail: yisun@uoguelph.ca

Lee (2003) develops a $\sqrt{n}$-consistent estimator of the parametric component of a partially linear quantile regression model, which is used to obtain his one-step semiparametric efficient estimator. As a result, how well the efficient estimator performs depends on the quality of the initial $\sqrt{n}$-consistent estimator. In this paper, we aim to improve the small sample performance of the one-step efficient estimator by proposing a new $\sqrt{n}$-consistent initial estimator, which does not require any trimming procedure and is less sensitive to data outliers and the choice of bandwidth than Lee's (2003) initial consistent estimator. Monte Carlo simulation results confirm that the proposed estimator and the one-step efficient estimator derived from it have more desirable empirical features than Lee's estimators.

Key Words: Partially linear quantile regression; local polynomial regression.
JEL Classification Numbers: C13, C14.

1. INTRODUCTION

Conditional mean regression models can be used to study the average responses of a dependent variable to changes in explanatory variables. However, when residual heteroscedasticity arises, conditional quantile curves at different probability masses have different shapes which lead to different implications. As a result, looking at the average responses exclusively will ignore idiosyncratic features of data and may give improper inferences in empirical problems where we are interested in exploring responses other the average one. Much of contribution in estimation of quantile regression models is motivated by the seminal work of Koenker and Bassett (1978, 1982), Newey and Powell (1990), and Powell (1984). Buchinsky (1994) and a special issue of Empirical Economics (2001, vol. 26) provide good empirical applications in the framework of parametric regression models.

1529-7373/2005

For pure nonparametric estimation of conditional quantile curves, related work includes the kernel and k-nearest neighbor estimators of Bhattacharya and Gangopadtyay (1990), spline smoothing estimator of Koenker, Ng, and Portnoy (1994), weighted Nadaraya-Watson estimator of Hall, Wolff, and Yao (1999) and Cai (2002), the local linear regression approach of Fan, Hu, and Truong (1994), the double kernel method of Yu and Jones (1998), and others. Extension to semiparametric quantile regression models includes the quantile index model of Khan (2001), the partially linear quantile regression model of He and Shi (1996) and Lee (2003), the censored regression model of Chen and Khan (2001).

In this paper, we are interested in estimating the partially linear quantile regression model. Lee (2003) introduces a $\sqrt{n}$-consistent average quantile regression (AQR) estimator and a one-step semiparametric efficient estimator. How well the one-step semiparametric efficient estimator performs empirically depends on the quality of the initial $\sqrt{n}$-consistent estimator. It is well-known that sample mean is more sensitive to outliers than median estimator, especially when the sample size is small. Therefore, we suspect that Lee's (2003) AQR estimator may not be a good candidate in such cases and naturally the performance of his one-step efficient estimator may be negatively influenced. Actually, this conjecture is confirmed by our Monte Carlo experiments in Section 3. Motivated by this, we propose a new initial $\sqrt{n}$-consistent estimator—quantile-based quantile regression (QQR) estimator, which requires no trimming procedure and is less sensitive to data outliers and the choice of bandwidth than AQR estimator for small samples; thus the one-step efficient estimator derived from it enjoys more desirable stability and accuracy than that derived from AQR estimator.

The rest of the paper is organized as follows. Section 2 describes estimation methodology of the proposed estimator. Both consistency and asymptotic normality results are derived. Monte Carlo simulations are conducted to evaluate the small sample performance of the proposed estimator and Lee's estimators in Section 3. Section 4 concludes. All the proofs are left to the Appendix.

2. ESTIMATION METHODOLOGY

Assume that we have a partially linear quantile regression model defined at $\alpha \in (0,1)$ as follows

$$Y_i = X_i'\beta_{\alpha,0} + \theta_\alpha(Z_i) + V_i, E_\alpha(V_i|X_i, Z_i) = 0, \tag{1}$$

where $X_i \in R^k$, β_0 is a $k \times 1$ vector to be estimated and $Z_i \in R^d$, the functional form of $\theta_\alpha(\cdot)$ is unknown, $i = 1, \cdots, n$. $E_\alpha(V|x,z)$, the α^{th}-

quantile of V conditional on $(X, Z) = (x, z)$, is defined as

$$E_\alpha (V|x, z) = \inf \{v : \Pr (V \le v|X = x, Z = z) \ge \alpha\}. \quad (2)$$

We assume that Z does not contain any elements of X, or essentially, we assume that knowing $Z = z$ can not be used to calculate the exact value of x. For the sake of identification, we also assume that X and Z do not contain 1, and any constant number will be absorbed into $\theta_\alpha (\cdot)$.

Let $\tau_i = \left(\tau_i^{1'}, \cdots, \tau_i^{p'}\right)'$ contain all partial derivatives of $\theta_\alpha (z)$ at point $z_i = (z_{i1}, \cdots, z_{id})'$ up to order p, with, in particular $\tau_i^1 = \bigtriangledown_{z'} \theta_\alpha (z_i)$. In total, there are m_r distinct r^{th} order partial derivatives so that τ_i is an $m = \sum_{r=1}^p m_r$ by 1 vector. For any $\|Z_j - z_i\| = o(1)$, taking Taylor expansion of $\theta_\alpha (z_j)$ at z_i yields

$$\theta_\alpha (Z_j) = \theta_\alpha (z_i) + \tau_i' w_{ij} + o(\|Z_j - z_i\|^p), \quad (3)$$

where $w_{ij} = \left(w_{ij}^{1'}, \cdots, w_{ij}^{p'}\right)'$ contains the corresponding Taylor coefficients with w_{ij}^r being $m_r \times 1$ vector, in particular, $w_{ij}^1 = Z_j - z_i$. This suggests estimating $\theta_\alpha (Z_j)$ by a local p^{th}-order polynomial regression approach, which extends the local linear regression approach of Fan, Hu, and Truong (1994) to the multivariate case. In particular, a local constant regression approach is applied when $p = 0$; a local linear regression approach is applied when $p = 1$; and a local quadratic regression approach is applied when $p = 2$.

Under the framework of the partially linear quantile regression model, a two-step semiparametric estimator is developed and called quantile-based quantile regression (QQR) estimator. In the first step, the unknown function $\theta_\alpha (z)$ is recovered from data by nonparametric smoothing method at each data point. That is, the optimization problem solved at each point i is as follows

$$\min_{a_{i0}, a_{i1}, \beta_\alpha} \sum_{j=1}^n \rho_\alpha \left(Y_j - X_j' \beta_\alpha - a_{i0} - a_{i1}' w_{ij}\right) K_i \left(H^{-1} (Z_j - z_i)\right), \quad (4)$$

where a_{i1} is an $m \times 1$ vector, $H = diag(h_1, h_2, \cdots, h_d)$ is a smoothing parameter matrix, $K\left(H^{-1}u\right) = \prod_{j=1}^d k\left(\frac{u_j}{h_j}\right)$ for any $u \in R^d$; $\rho_\alpha (u) = u(\alpha - I(u \le 0))$ is called the check function. The leave-one-out technique is used here with $K_i \left(H^{-1} (Z_j - z_i)\right) = 0$ for $j = i$. Let $\left(\widetilde{a}_{i0}, \widetilde{a}_{i1}, \widetilde{\beta}_{\alpha,i}\right)$ be the solution of (4) at the i^{th} point. In the Appendix, we show that for any $i = 1, 2, \cdots, n$, $\widetilde{\beta}_{\alpha,i} \xrightarrow{p} \beta_{\alpha,0}$ at nonparametric rate, and $\widetilde{\theta}_\alpha (z_i) = \widetilde{a}_{i0}$

is the consistent estimator of $\theta_\alpha(z_i)$. This step aims to obtain a consistent estimator of $\theta(z)$.

In the second step, a $\sqrt{n}$-consistent estimator of $\beta_{\alpha,0}$ will be derived by solving the following minimization problem

$$\widehat{\beta}_{\alpha,QQR} = \arg\min_{\beta_\alpha \in \Theta} \sum_{i=1}^{n} \rho_\alpha \left(Y_i - X_i' \beta_\alpha - \widetilde{\theta}_\alpha(Z_i) \right), \tag{5}$$

where $\left\{ \widetilde{\theta}_\alpha(Z_i) \right\}_{i=1}^{n}$ are calculated from the first step. In the Appendix, we show that $\widehat{\beta}_{\alpha,QQR} - \beta_{\alpha,0} = O_p\left(n^{-1/2}\right)$.

Finally, $\{\theta_\alpha(z_i)\}_{i=1}^{n}$ can be estimated by the local linear regression approach

$$\min_{b_i} \sum_{j=1}^{n} \rho_\alpha \left(Y_j - X_j' \widehat{\beta}_{\alpha,QQR} - b_{i0} - b_{i1}'(Z_j - z_i) \right) K_i\left(H^{-1}(Z_j - z_i)\right), \tag{6}$$

where $b_i = (b_{i0}, b_{i1}')'$ is a $(d+1) \times 1$ vector, and

$$\widehat{\theta}_\alpha(z_i) = \widehat{E}_\alpha\left(Y_i - X_i' \widehat{\beta}_{\alpha,QQR} | Z_i = z_i \right) = \widehat{b}_{i0}, i = 1, \cdots, n.$$

For any matrix A, let $\|A\|$ denote its Euclidean norm, i.e. $\|A\| = [tr(AA')]^{1/2}$, and let $|A|$ denote its determinant if A is a square matrix. Let $\mathcal{S}_l(\mathcal{Z})$ be a functional space such that $g \in \mathcal{S}_l(\mathcal{Z})$ if $g : \mathcal{Z} \to R$, and $g(\cdot)$ has the order of smoothness l, and its $[l]^{th}$ partial derivatives satisfies the Lipschitz condition

$$\left\| D^{[l]}g(z_1) - D^{[l]}g(z_2) \right\| \le M \|z_1 - z_2\|^\gamma, \text{for any } z_1, z_2 \in \mathcal{Z}, \tag{7}$$

where $\mathcal{Z}$ is a convex set of R^d, $l = [l] + \gamma$ with $[l]$ being a nonnegative integer and $0 < \gamma < 1$, and $D^s g(z) = \partial^s g(z) / \partial z_1^{s_1} \cdots \partial z_d^{s_d}$ with $s = \sum_{i=1}^{d} s_i$ and $\{s_i\}_{i=1}^{d}$ being nonnegative integer between 0 and $[l]$.

Some regular assumptions required by the main theorems are listed below.

Assumption 1. *$\{(Y_i, X_i, Z_i)\}_{i=1}^{n}$ is an i.i.d. random sequence from a joint probability distribution $F(y, x, z)$ on $R \times R^k \times R^d$. Both $Y|X, Z$ and Z are continuously distributed.*

Assumption 2. *Let $f_v(v|x, z)$ be the conditional p.d.f of V given $(X, Z) = (x, z)$, and let $F_v(v|x, z)$ be its corresponding conditional c.d.f.. $f_v(0|x, z) >$*

0, *and* $F_v(0|x,z) = \alpha$ *for all* x *and* z. *Let* $f(z|x)$ *be the conditional probability function of* z *given* $X = x$. $f_v(v|x,z)$, $f(z|x)$, *and their first- and second-order partial derivatives are continuous and uniformly bounded (almost surely).*

Assumption 3. $\theta_\alpha(z) \in \mathcal{S}_{p+1}(\mathcal{Z})$ *and* $\Pr\{\|D^s\theta_\alpha(z)\| \leq M\} = 1$ *(a.s.) for all* $0 \leq s \leq p+1$ *or* $\theta_\alpha(z) \in \mathcal{S}_{p+\gamma}(\mathcal{Z})$ *for some* $\gamma \in (0,1)$.

Assumption 4. *The product kernel is* $K(u) = \prod_{j=1}^{d} k(u_j)$ *for any* $u \in R^d$, *where* $k(\cdot) \geqslant 0$ *is a symmetric function taking values on a bounded support* $[-1,1]$ *and*

$$\int_{-1}^{1} k(u_1)\,du_1 = 1, \int_{-1}^{1} K^2(u)\,du = R(K) < \infty. \tag{8}$$

Let $\omega = \left(\omega^{1'}, \cdots, \omega^{p'}\right)'$ *be an* $m \times d$ *nonnegative integer matrix corresponding to the polynomial order of* $(Z_{j1} - z_{i1}, \cdots, Z_{jd} - z_{id})$ *in* w_{ij}. *In particular,* ω^r *is* $m_r \times d$ *matrix with* $\sum_{i=1}^{d} \omega_{ij}^r = r$ *for* $r = 1, 2, \cdots, p$, *and* $j = 1, 2, \cdots, m_r$. *Correspondingly, denote* $\lambda(K) = \left(\lambda(K)^1, \cdots, \lambda(K)^p\right)$ *to be an* $m \times 1$ *vector with* $\lambda(K)^r$ *being* $m_r \times 1$ *vector, and*

$$\mu(K) = \begin{pmatrix} \mu(K)^{11} & \cdots & \mu(K)^{1p} \\ \vdots & \ddots & \vdots \\ \mu(K)^{p1} & \cdots & \mu(K)^{pp} \end{pmatrix}$$

to be an $m \times m$ *symmetric matrix with* $\mu(K)^{rt}$ *being* $m_r \times m_t$ *matrix, where*

$$\lambda(K)_j^r = \int_{-1}^{1} K(u) \prod_{i=1}^{d} u_i^{\omega_{ij}^r}\,du, \tag{9}$$

$$\mu(K)_{ll'}^{rt} = \int_{-1}^{1} K(u) \prod_{i=1}^{d} u_i^{\omega_{il}^r + \omega_{il'}^t}\,du, \tag{10}$$

where $t, r = 1, 2, \cdots, p; j = 1, 2, \cdots, m_r; l = 1, 2, \cdots, m_r; l' = 1, 2, \cdots, m_t$. *Obviously,* $\mu(K)$ *and* $\lambda(K)$ *are both finite.* $\lambda(K)_j^r = 0$ *if some* ω_{ij}^r *is odd and* $\mu(K)_{ll'}^{rt} = 0$ *if some* $\omega_{il}^r + \omega_{il'}^t$ *is odd.*

Finally, for any $z \in \mathcal{Z} \subset R^d$, define

$$S(z) = E\left(f_v(0|X,Z)\, f(Z|X) \begin{pmatrix} 1 & \lambda(K)' & X' \\ \lambda(K) & \mu(K) & \lambda(K)X' \\ X & X\lambda(K)' & XX' \end{pmatrix} \middle| Z = z \right). \tag{11}$$

Assumption 5. *$0 < var(V) < \infty$, $E\left(\|X\|^4\right) < \infty$. $E\left[XX'f_v(0|X,Z)\right]$ is nonsingular and finite matrix, and $S(z)$ is nonsingular and finite $(m+k+1) \times (m+k+1)$ matrix at each point $z \in \mathcal{Z} \subset R^d$.*

Assumption 6. *$\|H\| \to 0$, $n^v|H| \to \infty$, and $\sqrt{n}\|H\|^{p+1} \to 0$ as $n \to \infty$, where $p+1 > \frac{d}{2v}$ and $0 < v < \frac{1}{2} - \frac{1}{2+\delta}$ with $E\left(\|X\|^{2+\delta}\right) < \infty$ for $\delta > 0$.*

The proposed estimator can be extended to the case that Z contains discrete variables as well as continuously distributed variables. In Assumption 4, the kernel function $k(\cdot)$ having bounded support is not essential and it is assumed merely to simplify the assumptions and proofs. The main theorems of this paper are as follows and the proofs are left to the Appendix.

Theorem 1. *Suppose that Assumptions 1 - 5 hold. Then at each point i, if $\|H\| \to 0, n|H| \to \infty, \sqrt{n|H|}\|H\|^{p+1} = O(1)$, we have*

$$\widetilde{\beta}_{\alpha,i} = \beta_{\alpha,0} + O_p\left(n^{-1/2}|H|^{-1/2}\right), \tag{12}$$

$$\widetilde{\theta}_{\alpha}(z_i) = \theta_{\alpha}(z_i) + O_p\left(n^{-1/2}|H|^{-1/2}\right), \tag{13}$$

$$\frac{\partial^r \widetilde{\theta}_{\alpha}(z_i)}{\partial z_1^{\omega_{ij}^r} \cdots \partial z_d^{\omega_{dj}^r}} = \frac{\partial^r \theta_{\alpha}(z_i)}{\partial z_1^{\omega_{ij}^r} \cdots \partial z_d^{\omega_{dj}^r}} + O_p\left(n^{-1/2}\prod_{i=1}^{d} h_i^{-\frac{1}{2}-\omega_{ij}^r}\right), \tag{14}$$

for any $j = 1, 2, \cdots, m_r$; $r = 1, 2, \cdots, p$.

Remark 2.1. Theorem 1 indicates that the optimal bandwidth H is applicable. If the local $p^{th}-$order polynomial regression method is applied here, the optimal bandwidth is $h_i\tilde{}n^{-\frac{1}{2(p+1)+d}}$, $i = 1, 2, \cdots, d$.

THEOREM 2. *Under Assumptions 1-6, we have*

$$\sqrt{n}\left(\widehat{\beta}_{\alpha,QQR}-\beta_{\alpha,0}\right)\xrightarrow{d}N\left(0,\Omega_0\right), \tag{15}$$

where $\Omega_0=C^{-1}DC^{-1}$ *with* $C=E\left[XX'f_v\left(0|X,Z\right)\right]$ *and*

$$\begin{aligned} D &= \alpha\left(1-\alpha\right)E\left\{\left[X_1-E\left(\widetilde{X}_2|X_1,Z_1\right)\right]\left[X_1-E\left(\widetilde{X}_2|X_1,Z_1\right)\right]'\right\}, \\ \widetilde{X}_2 &= X_2 f_v\left(0|Z_1,X_2\right)f\left(Z_1|X_2\right)e_1'S^{-1}\left(Z_1\right)\begin{pmatrix} 1 & \lambda\left(K\right)' & X_1' \end{pmatrix}'. \end{aligned}$$

Remark 2.2. Theorem 2 requires that $h_i\sim n^{-\alpha}$ with $\frac{1}{2(p+1)}<\alpha<\frac{v}{d}$ for $i=1,2,\cdots,n$. It implies that $p+1>\frac{d}{2v}$. It implies that the optimal bandwidth $h_i\sim n^{-\frac{1}{2(p+1)+d}}$ does not satisfy this condition. Essentially, this theorem requires undersmoothing in the first step estimation of the unknown function $\theta_\alpha\left(z\right)$. This undersmoothing condition $\sqrt{n}\left\|H\right\|^{p+1}\to 0$ as $n\to\infty$, is required to remove the bias term when $\theta_\alpha\left(z\right)$ is replaced by $\widetilde{\theta}_\alpha\left(z\right)$ in the second step.

One can obtain a consistent estimator of Ω_0 by replacing the components of Ω_0 with its empirical counterparts. Thus a consistent estimator can be calculated as follows

$$\begin{aligned} \widehat{\Omega} &= \widehat{C}^{-1}\widehat{D}\widehat{C}^{-1} \\ \widehat{C} &= \frac{1}{n}\sum_{i=1}^{n}X_iX_i'\widehat{f}_v\left(0|X_i,Z_i\right) \\ \widehat{D} &= \frac{\alpha\left(1-\alpha\right)}{n}\sum_{i=1}^{n}\left[X_{1i}-\widehat{E}\left(\widetilde{X}_{2i}|X_{1i},Z_{1i}\right)\right]\left[X_{1i}-\widehat{E}\left(\widetilde{X}_{2i}|X_{1i},Z_{1i}\right)\right]', \end{aligned}$$

where $\widehat{E}\left(\widetilde{X}_{2i}|X_{1i},Z_{1i}\right)$ has to be estimated by the sample splitting method. However, we will not continue this line of discussion since the above proposed estimator is mainly introduced as an initial $\sqrt{n}$-consistent estimator for the one-step semiparametric efficient estimator defined below. The advantage or contribution of this paper is that the QQR estimator does not require any trimming procedure and is less sensitive to data outliers and the bandwidth choice of the above first step nonparametric estimation. Such stable property of the QQR estimator is expected to be passed on to the one-step efficient estimator derived from it.

2.1. Semiparametric efficient estimator

The estimator derived above is not semiparametric efficient in the sense of Newey (1990, 1994). Lee (2003) shows that the semiparametric efficient bound is

$$\begin{aligned} V_B &= \alpha(1-\alpha)\left\{E\left[f_v^2(0|X,Z)XX'\right]\right. \\ &\left. -E\left[\frac{E\left[f_v^2(0|X,Z)X|Z\right]E\left[f_v^2(0|X,Z)X'|Z\right]}{E\left[f_v^2(0|X,Z)|Z\right]}\right]\right\}^{-1} \end{aligned} \tag{16}$$

provided that $E[S_\alpha S_\alpha']$ is nonsingular, where S_α is the efficient score for β_α

$$S_\alpha(Y,X,Z,\theta_a,\beta_\alpha) = \frac{f_v(0|X,Z)}{\alpha(1-\alpha)}\left[\alpha - I(Y-\theta_a(Z)-X'\beta_\alpha \le 0)\right]\left[X-T(Z)\right], \tag{17}$$

and

$$T(Z) = \frac{E\left[f_v^2(0|X,Z)X|Z\right]}{E\left[f_v^2(0|X,Z)|Z\right]}.$$

To obtain a semiparametric efficient estimator, Lee (2003) proposes a two-step estimator. In the first step, a $\sqrt{n}-$consistent estimator, AQR estimator, is calculated as follows

$$\widehat{\beta}_{\alpha,AQR} = \frac{\sum_{i=1}^n \tau_z(Z_i)\widetilde{\beta}_{\alpha,i}}{\sum_{i=1}^n \tau_z(Z_i)} \tag{18}$$

where $\tau_z(z) = I(z \in \mathcal{Z})$ is a trimming function over a compact subset $\mathcal{Z} \in R^d$. This trimming procedure is applied to remove the negative tail effects of Z. In the second step, a one-step semiparametric efficient estimator is constructed as follows

$$\widehat{\beta}^*_{\alpha,AQR} = \widehat{\beta}_{\alpha,AQR} + \left[\sum_{i=1}^n \partial\widehat{S}_\alpha\left(\widehat{\beta}_{\alpha,AQR}\right)\Big/\partial\beta_\alpha\right]^{-1}\sum_{i=1}^n \widehat{S}_\alpha\left(\widehat{\beta}_{\alpha,AQR}\right), \tag{19}$$

where

$$\widehat{S}_\alpha(\beta_\alpha) = \frac{\widehat{f}_v(0|X,Z)}{\alpha(1-\alpha)}\left[\alpha - 1 + J\left(\frac{Y-\widehat{\theta}_a(Z)-X'\widehat{\beta}_\alpha}{j_n}\right)\right]\left[X-\widehat{T}(Z)\right] \tag{20}$$

is the smoothed estimator of $S_\alpha(\beta)$ in (17) and $\widehat{f}_v(0|X,Z)$ and $\widehat{T}(Z)$ are the kernel estimates of $f_v(0|X,Z)$ and $T(Z)$, respectively; and

$$J(x) = \begin{cases} 0 & if \;\; v < -1 \\ 0.5 + \frac{15}{16}\left(x - \frac{2}{3}x^3 + \frac{1}{5}x^5\right) & if \;\; |v| \le 1 \\ 1 & if \;\; v > 1 \end{cases}. \tag{21}$$

It is reasonable to expect that the performance of the second step efficient estimator depends on the quality of the first step estimator, $\widehat{\beta}_{\alpha,AQR}$, an estimator constructed from empirical averages of $\left\{\widetilde{\beta}_{\alpha,i}\right\}$. It is well-known that sample averages are more sensitive to data outliers than sample medians. This argument can be paralleled to our case. The trimming function $\tau_z(z)$ in (18) can remove some side effects of extreme estimates $\widetilde{\beta}_{\alpha,i}$ at some points; however, there is no theoretical guidances in choosing the trimming function. Therefore, we construct the one-step efficient estimator, $\widehat{\beta}^*_{\alpha,QQR}$, by using $\widehat{\beta}_{\alpha,QQR}$ as the initial consistent estimator in (19).

The asymptotic results of $\widehat{\beta}^*_{\alpha,QQR}$ will be the same as in Theorem 5 of Lee's (2003); however, the proofs are omitted to make the paper dense and we refer readers to Lee (2003) for assumptions required and proofs.

3. MONTE CARLO SIMULATIONS

The data is generated by the following model

$$y_i = X_i\beta_0 + g(Z_i) + \sigma_0(X_i, Z_i)(u_i - E_p(u_i)), \tag{22}$$

where $u_i\tilde{\ } iidN(0,1)$ and $E_p(u_i)$ is the pth-quantile of u_i, $i = 1, 2, \cdots, n$. $\{(X_i, Z_i)\}_{i=1}^n$ is an i.i.d. random sequence from a joint normal distribution with zero mean, unit variance, and a correlation of 0.5. We use the same data generating designs as Lee (2003). Let $DGP(i_0, j_0)$ be the true data generating mechanism with

$$\sigma_0(x,z) = \begin{cases} \frac{1}{3} & \text{if } i_0 = 1, \\ C\exp(0.25(x+z)) & \text{if } i_0 = 2, \end{cases} \tag{23}$$

where C is a constant used to normalize the mean of $\sigma_0(x,z)$ to be $\frac{1}{3}$, and

$$g(z) = \begin{cases} 1+z & \text{if } j_0 = 1, \\ z + 4exp(-2z^2)/\sqrt{2\pi} & \text{if } j_0 = 2, \\ \sin(\pi z) & \text{if } j_0 = 3. \end{cases} \tag{24}$$

Clearly, the value of i_0 describes whether disturbance heteroskedasticity exists and j_0 is used to define the function form of $g(\cdot)$.

The quantiles of interest are $\alpha = 0.15, 0.25, 0.50, 0.75,$and 0.85. The sample size is $n = 100$. The number of Monte Carlo replications is $m = 1000$. The Epanechnikov kernel function is used here $k(u) = \frac{3}{4}\left(1-u^2\right) I(|u| \leq 1)$. In the first step estimation, we choose $p = 2$ and the bandwidth $h = cn^{-2/9}\widehat{\sigma}_z$, where $\widehat{\sigma}_z$ is the sample standard deviation of Z. In order to measure sensitivity of the proposed estimator to the choices of bandwidth, we divide the interval $[0.5, 5.0]$ with equal increment 0.15 and let c be any grid value. The algorithm of Koenker and d'Orey (1987) is used to solve the optimization problems (4) and (5).

For each set of Monte Carlo experiments, we then calculate the squared root of mean squared errors and mean absolute bias of each estimator

$$RMSE_\alpha = \sqrt{\frac{1}{m}\sum_{j=1}^{m}\left(\widehat{\beta}_{\alpha,j} - \beta_{\alpha,0}\right)^2}; ABIAS_\alpha = \frac{1}{m}\sum_{j=1}^{m}\left|\widehat{\beta}_{\alpha,j} - \beta_{\alpha,0}\right|,$$

where $\widehat{\beta}_\alpha$ will be $\widehat{\beta}_{\alpha,QQR}, \widehat{\beta}_{\alpha,AQR}, \widehat{\beta}^*_{\alpha,QQR}$, and $\widehat{\beta}^*_{\alpha,AQR}$. Tables 1 and 2 present the minimum, median and standard deviations of those $RMSE_\alpha$ and $ABIAS_\alpha$. $\tau(z) = I[Q_{0.01}(z) \le z \le Q_{0.99}(z)]$ and $\tau(x,z) = I[Q_{0.01}(z) \le z \le Q_{0.99}(z), Q_{0.01}(x) \le x \le Q_{0.99}(x)]$, where $Q_p(z)$ and $Q_p(x)$ are the empirical quantiles of z and x at probability mass $p \in (0,1)$, respectively.

In Table 1, when $i_0 = 1$ (homoscedastic case), $\widehat{\beta}_{\alpha,AQR}$ attains the semiparametric efficient bounds since $var(Z|X) = 0.5$ is a constant number. In this case, we find that $\widehat{\beta}_{\alpha,AQR}$ has the smallest minimum $RMSE$ values but standard deviations are large and that $\widehat{\beta}^*_{\alpha,QQR}$ has the smallest median values and standard deviations. The standard deviations of RMSEs from $\widehat{\beta}^*_{\alpha,AQR}$ are very wild and ignored here.

The parameter j_n in Eq. (20) is set to 0.5. Our experiences agree with Lee's comment on the choice of $j_n \to 0$ as $n \to \infty$—j_n can not be too small to regulate the behavior of $\widehat{\beta}^*_{\alpha,AQR}$, which is of less concern to $\widehat{\beta}^*_{\alpha,QQR}$. The reason behind this is that small j_n tends to give zero values to $\sum_{i=1}^{n} \partial\widehat{S}_\alpha\left(\widehat{\beta}_{\alpha,AQR}\right)\Big/\partial\beta_\alpha$, which actually occurs $\widehat{\beta}_{\alpha,AQR}$ is far away from the true value as a result of inappropriate choice of the bandwidth h. Or the ill performance of $\widehat{\beta}_{\alpha,QQR}$ then $\widehat{\beta}^*_{\alpha,QQR}$ lies in its sensitivity to the choice of bandwidth h in the first step nonparametric estimation of $\widehat{\beta}_{\alpha,AQR}$. In addition, although $\widehat{\beta}_{\alpha,AQR}$ and $\widehat{\beta}^*_{\alpha,AQR}$ are both efficient theoretically in the presence of homoscedastic residuals, due to the requirement of extra nonparametric estimations the one-step efficient estimators $\widehat{\beta}^*_{\alpha,QQR}$ and $\widehat{\beta}^*_{\alpha,AQR}$ underperforms $\widehat{\beta}_{\alpha,AQR}$ in terms of minimum RMSE values.

When $i_0 = 1$, i.e. heteroskedastic case, both $\widehat{\beta}^*_{\alpha,AQR}$ and $\widehat{\beta}^*_{\alpha,QQR}$ outperform their respective initial estimators $\widehat{\beta}_{\alpha,AQR}$ and $\widehat{\beta}_{\alpha,QQR}$ and have similar minimum RMSE values; however, $\widehat{\beta}^*_{\alpha,QQR}$ provides the most promising results in terms of stability across different bandwidth choices. Therefore, in light of RMSEs, $\widehat{\beta}^*_{\alpha,QQR}$ may be recommended to empirical econometricians because of its stability with respect to the choice of h in the first step estimation and the smallest mean squared errors.

TABLE 1.

Summary statistics of squared root of mean squared errors

DGP	Prob.	QQR			AQR			QQR*			AQR*	
(i_0, j_0)		Min.	Median	Stdev.	Min.	Median	Stdev.	Min.	Median	Stdev.	Min.	Median
	0.15	0.065	0.070	0.008	0.060	0.078	0.801	0.066	0.069	0.007	0.065	0.073
	0.25	0.058	0.062	0.009	0.053	0.074	1.054	0.058	0.060	0.007	0.057	0.083
(1,1)	0.5	0.052	0.054	0.005	0.049	0.050	0.794	0.052	0.053	0.003	0.052	0.052
	0.75	0.056	0.060	0.007	0.052	0.060	0.791	0.057	0.058	0.005	0.056	0.059
	0.85	0.064	0.067	0.006	0.057	0.062	3.012	0.065	0.067	0.005	0.063	0.065
	0.15	0.065	0.070	0.009	0.060	0.077	0.802	0.066	0.069	0.007	0.065	0.073
	0.25	0.058	0.061	0.009	0.052	0.073	1.054	0.058	0.060	0.007	0.057	0.085
(1,2)	0.5	0.053	0.055	0.005	0.048	0.056	0.794	0.052	0.054	0.003	0.052	0.055
	0.75	0.059	0.061	0.006	0.052	0.060	0.792	0.058	0.060	0.005	0.057	0.059
	0.85	0.066	0.069	0.005	0.056	0.063	3.013	0.066	0.068	0.004	0.064	0.069
	0.15	0.067	0.073	0.005	0.059	0.077	0.802	0.068	0.072	0.005	0.065	0.073
	0.25	0.060	0.064	0.006	0.053	0.074	1.055	0.058	0.062	0.006	0.057	0.083
(1,3)	0.5	0.055	0.063	0.015	0.050	0.070	0.792	0.054	0.057	0.006	0.053	0.068
	0.75	0.061	0.067	0.008	0.053	0.067	0.791	0.059	0.063	0.005	0.058	0.070
	0.85	0.068	0.074	0.007	0.059	0.072	3.023	0.067	0.072	0.007	0.065	0.078
	0.15	0.058	0.066	0.017	0.058	0.089	0.381	0.054	0.059	0.011	0.054	0.078
	0.25	0.050	0.056	0.014	0.051	0.085	0.507	0.048	0.051	0.009	0.049	0.071
(2,1)	0.5	0.044	0.046	0.005	0.046	0.049	0.375	0.044	0.045	0.003	0.045	0.045
	0.75	0.048	0.051	0.006	0.050	0.065	0.374	0.047	0.049	0.005	0.048	0.052
	0.85	0.056	0.058	0.006	0.056	0.070	4.984	0.052	0.054	0.006	0.053	0.056
	0.15	0.061	0.065	0.017	0.060	0.089	0.381	0.056	0.059	0.011	0.056	0.077
	0.25	0.052	0.056	0.014	0.053	0.085	0.508	0.050	0.052	0.009	0.050	0.070
(2,2)	0.5	0.045	0.048	0.005	0.048	0.056	0.374	0.045	0.046	0.003	0.045	0.049
	0.75	0.050	0.052	0.005	0.052	0.064	0.374	0.048	0.050	0.004	0.049	0.052
	0.85	0.058	0.062	0.005	0.058	0.069	4.984	0.054	0.056	0.005	0.054	0.058
	0.15	0.063	0.069	0.010	0.063	0.091	0.381	0.058	0.064	0.008	0.058	0.075
	0.25	0.053	0.059	0.008	0.055	0.087	0.508	0.051	0.054	0.007	0.051	0.071
(2,3)	0.5	0.047	0.058	0.018	0.051	0.076	0.373	0.046	0.051	0.008	0.047	0.063
	0.75	0.055	0.063	0.013	0.055	0.073	0.373	0.050	0.055	0.006	0.051	0.060
	0.85	0.066	0.076	0.013	0.063	0.078	4.994	0.058	0.064	0.008	0.056	0.071

TABLE 2.

Summary statistics of mean absolute biases

DGP	Prob.	QQR			AQR			QQR*			AQR*	
(i_0, j_0)		Min.	Median	Stdev.	Min.	Median	Stdev.	Min.	Median	Stdev.	Min.	Median
(1,1)	0.15	0.051	0.055	0.006	0.047	0.050	0.037	0.052	0.054	0.005	0.050	0.055
	0.25	0.046	0.048	0.007	0.042	0.046	0.044	0.045	0.047	0.005	0.045	0.050
	0.5	0.042	0.043	0.004	0.039	0.040	0.031	0.041	0.042	0.003	0.041	0.042
	0.75	0.045	0.047	0.005	0.042	0.044	0.033	0.045	0.046	0.004	0.045	0.046
	0.85	0.051	0.053	0.005	0.046	0.047	0.102	0.052	0.053	0.004	0.050	0.051
(1,2)	0.15	0.051	0.055	0.007	0.048	0.050	0.036	0.051	0.053	0.006	0.050	0.055
	0.25	0.046	0.048	0.007	0.042	0.046	0.044	0.045	0.047	0.006	0.045	0.050
	0.5	0.043	0.044	0.004	0.038	0.042	0.031	0.042	0.043	0.002	0.041	0.043
	0.75	0.047	0.048	0.005	0.041	0.044	0.033	0.046	0.047	0.004	0.045	0.047
	0.85	0.052	0.055	0.004	0.044	0.048	0.102	0.052	0.054	0.003	0.051	0.052
(1,3)	0.15	0.053	0.057	0.004	0.046	0.050	0.037	0.053	0.056	0.004	0.051	0.056
	0.25	0.047	0.051	0.005	0.041	0.045	0.044	0.046	0.049	0.005	0.045	0.050
	0.5	0.044	0.050	0.012	0.040	0.047	0.030	0.043	0.046	0.004	0.042	0.049
	0.75	0.048	0.053	0.006	0.042	0.047	0.032	0.047	0.050	0.004	0.046	0.050
	0.85	0.053	0.058	0.006	0.046	0.052	0.102	0.053	0.057	0.006	0.051	0.055
(2,1)	0.15	0.046	0.051	0.015	0.045	0.054	0.033	0.043	0.046	0.009	0.042	0.052
	0.25	0.039	0.044	0.011	0.041	0.052	0.036	0.038	0.041	0.007	0.038	0.046
	0.5	0.035	0.037	0.004	0.036	0.039	0.023	0.035	0.035	0.002	0.035	0.036
	0.75	0.037	0.040	0.004	0.040	0.045	0.026	0.037	0.038	0.004	0.038	0.041
	0.85	0.044	0.046	0.004	0.045	0.049	0.166	0.041	0.042	0.004	0.041	0.043
(2,2)	0.15	0.048	0.051	0.014	0.048	0.054	0.032	0.044	0.046	0.009	0.043	0.052
	0.25	0.041	0.044	0.011	0.042	0.052	0.036	0.039	0.041	0.007	0.039	0.046
	0.5	0.036	0.038	0.004	0.038	0.042	0.022	0.036	0.037	0.002	0.036	0.039
	0.75	0.039	0.041	0.004	0.041	0.045	0.026	0.038	0.039	0.003	0.039	0.041
	0.85	0.045	0.049	0.004	0.045	0.050	0.166	0.042	0.044	0.004	0.042	0.045
(2,3)	0.15	0.049	0.055	0.009	0.049	0.055	0.030	0.046	0.050	0.006	0.046	0.055
	0.25	0.042	0.046	0.007	0.044	0.051	0.035	0.040	0.043	0.005	0.040	0.047
	0.5	0.037	0.046	0.014	0.041	0.050	0.020	0.037	0.040	0.006	0.037	0.046
	0.75	0.043	0.049	0.011	0.044	0.051	0.025	0.040	0.044	0.005	0.041	0.046
	0.85	0.051	0.060	0.011	0.049	0.056	0.166	0.045	0.050	0.007	0.044	0.051

In Table 2, when $i_0 = 0$, $\widehat{\beta}_{\alpha,AQR}$ has the smallest minimum and median mean absolute bias but large variations across different choices of bandwidth. When $i_0 = 1$, $\widehat{\beta}^*_{\alpha,QQR}$ has the smallest minimum, median and standard deviations of mean absolute bias. Therefore, in terms of mean absolute bias, $\widehat{\beta}^*_{\alpha,QQR}$ is preferred in the presence of disturbance heteroscedasticity and $\widehat{\beta}^*_{\alpha,AQR}$ is preferred in the presence of disturbance homoscedasticity if appropriate bandwidth h is used.

Since quantile regression models are of more interest for heteroscedastic case than for homoscedastic case, where $\widehat{\beta}^*_{\alpha,QQR}$ overperforms the other three estimators, $\widehat{\beta}^*_{\alpha,QQR}$ will be our selling product to empirical econometricians who are interested in applying the partially quantile regression models to their data.

4. CONCLUSION

In this paper, we improve the small sample performance of Lee's (2003) efficient estimator of the parametric component of the partially linear regression model by proposing a different initial $\sqrt{n}$-consistent estimator from his AQR estimator. Monte Carlo simulations indicate that the one-step efficient estimator calculated from our QQR estimator has desirable performance in small samples: No subjective trimming parameter to be required and stabile performances with respect to the choice of bandwidth make our QQR estimator is more user-friendly.

APPENDIX

Two frequently used formulas in the following proofs are

$$\rho_\alpha (x - y) - \rho_\alpha (x) = y [\alpha - I (x < 0)] + \int_0^y [I (x \le t) - I (x \le 0)] \, dt \quad \text{(A.1)}$$

for any $x \neq 0$; and

$$\Pr \left\{ \sum_{i=1}^n I (i : V_i = 0) = O (1) \right\} = 1 \text{ (a.s.)}, \quad \text{(A.2)}$$

where $I (i : V_i = 0) = 1$ if $V_i = 0$; 0 otherwise. Equation (A.2) is the result of the continuity assumption of $V|X, Z$ and the LLN. And in the following context, we denote $E (\cdot|I_i) = E (\cdot| X_i, Z_i)$.

Proof of Theorem 1. Suppose that we are estimating β_α and $\theta_\alpha (z)$ at a particular point (x_0, z_0, y_0). Corresponding to w_{ij} and ω defined in Section 2, we define $q = \left(q^{1'}, \cdots, q^{p'} \right)'$ where $q^r_j = \prod_{i=1}^d h_i^{\omega^r_{ij}}$ for $r =$

$1, 2, \cdots, p$; $j = 1, 2, \cdots, m_r$. For example, $q^1_{ij} = (h_1, \cdots, h_d)'$ and $q^2 = \left(h_1^2, h_1 h_2, \cdots, h_d^2\right)$. Finally, let $Q = diag\,(q)$ be an $m \times m$ diagonal matrix and

$$\begin{aligned} Y_i^* &= Y_i - X_i'\beta_{\alpha,0} - \theta_\alpha\,(z_0) - \tau_0' w_{0i}, \\ K_i &= K\left(H^{-1}\,(Z_i - z_0)\right), \text{ and } K_i = 0 \text{ if } Z_i = z_0, \\ W_i &= \left(\begin{array}{ccc} 1 & w_{0i}'Q^{-1} & X_i' \end{array}\right)', \end{aligned}$$

then $\widetilde{\delta}_n = \sqrt{n\,|H|}\begin{pmatrix} \widetilde{a}_{i0} - \theta_\alpha\,(z_0) \\ Q\,(\widetilde{a}_{i1} - \tau_0) \\ \widehat{\beta}_\alpha - \beta_{\alpha,0} \end{pmatrix}$ will minimize

$$\begin{aligned} \widehat{G}_n\,(\delta) &= \sum_{i=1}^n \left[\rho_\alpha\,(Y_i - X_i'\beta_\alpha - a_0 - a_1' w_{0i}) - \rho_\alpha\,(Y_i^*)\right] K_i \\ &= \sum_{i=1}^n \left[\rho_\alpha\left(Y_i^* - \frac{W_i'\delta}{\sqrt{n\,|H|}}\right) - \rho_\alpha\,(Y_i^*)\right] K_i, \end{aligned}$$

which is convex in δ. The rest of proofs will follow the proofs in Fan, Hu, and Truong (1994). First, rewrite above equation as

$$\widehat{G}_n\,(\delta) = E\left[\widehat{G}_n\,(\delta)\,|X, Z\right] + \frac{\delta'}{\sqrt{n\,|H|}}\sum_{i=1}^n K_i W_i\left[\rho_\alpha'\,(Y_i^*) - E\left(\rho_\alpha'\,(Y_i^*)\,|I_i\right)\right] + R_n\,(\delta), \tag{A.3}$$

where at any given $\delta \in \Theta \subset R^{m+k+1}$, we obtain $R_n\,(\delta) = o_p\,(1)$, since $E\left[R_n\,(\delta)\right] = 0$ and Condition A *(iv)* in Fan, Hu, and Truong (1994) is verified as follows

$$\begin{aligned} & E\left\{\left[\rho_\alpha\,(Y_i^* - t) - \rho_\alpha\,(Y_i^*) - \rho_\alpha'\,(Y_i^*)\,t\right]^2\right\} \\ = \ & E\left\{\int_0^t \left[I\,(Y_i^* \le s) - I\,(Y_i^* \le 0)\,ds\right]\right\}^2 + o\,(1) \\ = \ & \int_0^t\int_0^t E\left[\left(I\,(Y_i^* < s_1) - I\,(Y_i^* < 0)\right)\left(I\,(Y_i^* < s_2) - I\,(Y_i^* < 0)\right)\right] ds_1 ds_2 \\ \le \ & \int_0^{|t|}\int_0^{|t|} \Pr\left(0 < |Y_i^*| < |t|\right) ds_1 ds_2 = o\left(t^2\right) \text{ as } t \to 0, \end{aligned}$$

where the second equation follows from (A.1) and (A.2).

Denote

$$\zeta_i = \theta_\alpha\,(z_0) + \tau_0' w_{0i} - \theta_\alpha\,(Z_i) = -\frac{1}{(p+1)!}\overline{\tau}_0^{(p+1)\prime} w_{0i}^{p+1}, \tag{A.4}$$

where $\overline{\tau}_0^{(p+1)} = \tau_0^{(p+1)}(\overline{z}_i)$ and $\overline{z}_i$ lies between z_0 and Z_i. Because the kernel function $K(\cdot)$ is zero outside the unit circle, only those points that are close to z_0 will be used, that is, only Z_i such that $\left\|H^{-1}(Z_i - z_0)\right\| \leq 1$ has effect on the summation. Hence,

$$|\zeta_i| = O\left(\|H\|^{p+1}\right) \text{ uniformly over such } i \text{ that } \left\|H^{-1}(Z_i - z_0)\right\| \leq 1. \tag{A.5}$$

Since $Y|X, Z$ is continuously distributed, we have $\Pr\left\{\sum_{i=1}^{n} I(Y_i^* = 0) = O(1)\right\} = 1$ almost surely. Then applying formula (A.1), we have

$$\begin{aligned}
& E\left[\widehat{G}_n(\delta)\,|X, Z\right] \\
= & \sum_{i=1}^{n} K_i E\left\{\frac{W_i'\delta}{\sqrt{n|H|}}\rho_\alpha'(Y_i^*) + \int_0^{\frac{W_i'\delta}{\sqrt{n|H|}}} \left[I(Y_i^* < t) - I(Y_i^* < 0)\right] dt \,\middle|\, I_i\right\} + o_p(1) \\
= & \sum_{i=1}^{n} K_i \left\{\frac{W_i'\delta}{\sqrt{n|H|}} E\left[\rho_\alpha'(Y_i^*)\,|I_i\right] + \int_0^{\frac{W_i'\delta}{\sqrt{n|H|}}} \left[F_v(\zeta_i + t|I_i) - F_v(\zeta_i|I_i)\right] dt\right\} + o_p(1) \\
= & \frac{\delta'}{\sqrt{n|H|}} \sum_{i=1}^{n} K_i W_i E\left[\rho_\alpha'(Y_i^*)\,|I_i\right] + \frac{\delta' S_n \delta}{2} + o_p(1),
\end{aligned} \tag{A.6}$$

where

$$\begin{aligned}
& \left|\sum_{i=1}^{n} K_i \int_0^{\frac{W_i'\delta}{\sqrt{n|H|}}} \left[\left(f_v(\overline{\zeta}_i|I_i) - f_v(0|I_i)\right)(\zeta_i + t) - \left(f_v(\overline{\eta}_i|I_i) - f_v(0|I_i)\right)\zeta_i\right] dt\right| \\
= & \left|\sum_{i=1}^{n} K_i \int_0^{\frac{W_i'\delta}{\sqrt{n|H|}}} \left[\left(f_v(\overline{\zeta}_i|I_i) - f_v(\overline{\eta}_i|I_i)\right)\zeta_i - \left(f_v(\overline{\zeta}_i|I_i) - f_v(0|I_i)\right)t\right] dt\right| \\
\leq & M \sum_{i=1}^{n} K_i \left(\zeta_i^2 \left|\frac{W_i'\delta}{\sqrt{n|H|}}\right| + |\zeta_i| \frac{\delta W_i W_i'\delta}{n|H|}\right) \\
= & O\left(\|H\|^{2(p+1)} n^{1/2-v}\right) + O\left(\|H\|^{p+1}\right) = o_p(1),
\end{aligned} \tag{A.7}$$

where $\overline{\zeta}_i$ lies between $\zeta_i + t$ and 0, and $\overline{\eta}_i$ lies between ζ_i and 0. (A.7) holds under Assumption 6, since for any given small value $\varepsilon > 0$ and $\left\| H^{-1} (Z_i - z_0) \right\| \leq 1$, I have

$$\begin{aligned} & \Pr\left(\max_i \left| \frac{\delta' W_i}{\sqrt{n}} \right| > \varepsilon n^{-v} \right) \\ \leq \; & n \Pr\left(\left| \frac{\delta' W_i}{\sqrt{n}} \right| > \varepsilon n^{-v} \right) \\ \leq \; & nM \frac{E\left(\|X_i\|^r \right)}{\left(n^{1/2-v} \varepsilon \right)^r} \to 0, \end{aligned} \tag{A.8}$$

as $n \to \infty$ if $E\left(\|X_i\|^r \right) < \infty$ for $r > 2$ and $v < \frac{1}{2} - \frac{1}{r}$.

Lemma 1 below shows that $S_n = \frac{1}{n|H|} \sum_{i=1}^n f_v (0|I_i) K_i W_i W_i' = S(z_0) + o_p(1)$ if $\|H\| \to 0$, $n|H| \to \infty$, as $n \to \infty$, where $S(z_0)$ is defined in equation (11). Therefore, for any given δ, I have

$$\widehat{G}_n(\delta) = \delta' \widetilde{W}_n + \frac{\delta' S(z_0) \delta}{2} + r_n(\delta), \tag{A.9}$$

where $r_n(\delta) = o_p(1)$, and $\widetilde{W}_n = \frac{1}{\sqrt{n|H|}} \sum_{i=1}^n K_i W_i \rho_\alpha' (Y_i^*)$ with $\rho_\alpha' (Y_i^*) = \alpha - I(Y_i^* < 0)$. Then the convexity Lemma of Pollard (1991) yields

$$\sup_{\delta \in \Theta} |r_n(\delta)| = o_p(1). \tag{A.10}$$

And simple calculations show that

$$\widetilde{\delta}_n = -S(z_0)^{-1} \widetilde{W}_n + o_p(1). \tag{A.11}$$

Lemma 1 below shows that $\widetilde{W}_n = O_p(1)$ if $\|H\| \to 0$, $n|H| \to \infty$, $\sqrt{n|H|}\, \|H\|^{p+1} = O(1)$ as $n \to \infty$. Hence, $\widetilde{\delta}_n = O_p(1)$, which completes the proof of this theorem. ▌

Lemma 1. *If $\|H\| \to 0$, $n|H| \to \infty$, and $\sqrt{n|H|}\, \|H\|^{p+1} = O(1)$ as $n \to \infty$, then*

$$\begin{aligned} S_n &= S(z_0) + o_p(1), && \text{(A.12)} \\ \widetilde{W}_n &= O_p(1). && \text{(A.13)} \end{aligned}$$

Proof. Since

$$W_i W_i' = \begin{pmatrix} 1 & w_{0i}' Q^{-1} & X_i' \\ Q^{-1} w_{0i} & Q^{-1} w_{0i} w_{0i}' Q^{-1} & Q^{-1} w_{0i} X_i' \\ X_i & X_i w_{0i}' Q^{-1} & X_i X_i' \end{pmatrix},$$

then simple calculations yield

$$\begin{aligned}
& E\left(\frac{1}{n\,|H|}\ \sum_{i=1}^{n} K_i f_v\left(0|X_i,Z_i\right)\right) \\
=\ & E\left[f_v\left(0|X,z_0\right) f\left(z_0|X\right)\right] + O\left(\|H\|^2\right), \\
& E\left(\frac{1}{n\,|H|} \sum_{i=1}^{n} K_i w_{0i}' Q^{-1} f_v\left(0|X_i,Z_i\right)\right) \\
=\ & E\left[f_v\left(0|X,z_0\right) f\left(z_0|X\right)\right] \lambda\left(K\right)' + O\left(\|H\|^2\right), \\
& E\left(\frac{1}{n\,|H|} \sum_{i=1}^{n} K_i Q^{-1} w_{0i} w_{0i}' Q^{-1} f_v\left(0|X_i,Z_i\right)\right) \\
=\ & E\left[f_v\left(0|X,z_0\right) f\left(z_0|X\right)\right] \mu\left(K\right) + O\left(\|H\|^2\right),
\end{aligned}$$

and

$$\begin{aligned}
var\left(\frac{1}{n\,|H|} \sum_{i=1}^{n} K_i f_v\left(0|X_i,Z_i\right)\right) &= O\left(\frac{1}{n\,|H|}\right), \\
var\left(\frac{1}{n\,|H|} \sum_{i=1}^{n} K_i w_{0i}' Q^{-1} f_v\left(0|X_i,Z_i\right)\right) &= O\left(\frac{1}{n\,|H|}\right), \\
var\left(\frac{1}{n\,|H|} \sum_{i=1}^{n} K_i Q^{-1} w_{0i} w_{0i}' Q^{-1} f_v\left(0|X_i,Z_i\right)\right) &= O\left(\frac{1}{n\,|H|}\right).
\end{aligned}$$

Hence, if $E\left(\|X_i\|^4\right) < \infty$ holds, and $\|H\| \to 0, n\,|H| \to \infty$ as $n \to \infty$, then $S_n = S\left(z_0\right) + o_p\left(1\right)$.

In addition, we obtain

$$E\left(\widetilde{W}_n\right) = \sqrt{n\,|H|}\,\|H\|^{p+1} \begin{bmatrix} O\left(1\right) \\ M\lambda\left(K\right) + O\left(\|H\|^2\right) \\ O\left(1\right) \end{bmatrix},$$

and

$$\begin{aligned}
& var\left(\widetilde{W}_n\right) \\
=\ & E\left[var\left(\widetilde{W}_n|X,Z\right)\right] + var\left[E\left(\widetilde{W}_n|X,Z\right)\right] \\
=\ & \left[\alpha\left(1-\alpha\right) + O\left(\|H\|^{p+1}\right)\right] |H|^{-1} E\left(K_i^2 W_i W_i'\right) + O\left(\|H\|^{2(p+1)}\right),
\end{aligned}$$

if $\|H\| \to 0$, as $n \to \infty$, which will complete the proof of this Lemma. ∎

Proof of Theorem 2. Let $\widetilde{Y}_i^* = Y_i - \widetilde{\theta}_\alpha(Z_i) - X_i'\beta_{\alpha,0}$. Then $\widehat{\delta}_n = \sqrt{n}\left(\widehat{\beta}_\alpha - \beta_{\alpha,0}\right)$ will minimize

$$\begin{aligned}
&\widehat{Q}_n^*(\delta) \\
=& \sum_{i=1}^{n} \left[\rho_\alpha\left(Y_i - \widetilde{\theta}_\alpha(Z_i) - X_i'\beta_\alpha\right) - \rho_\alpha\left(\widetilde{Y}_i^*\right)\right] \\
=& \sum_{i=1}^{n} \left[\rho_\alpha\left(V_i + \theta_\alpha(Z_i) - \widetilde{\theta}_\alpha(Z_i) - \frac{\delta' X_i}{\sqrt{n}}\right) - \rho_\alpha\left(V_i + \theta_\alpha(Z_i) - \widetilde{\theta}_\alpha(Z_i)\right)\right]
\end{aligned}$$

which is convex in δ. Rewriting $\widehat{Q}_n^*(\delta)$ yields

$$\widehat{Q}_n^*(\delta) = E\left[\widehat{Q}_n^*(\delta)\,|X, Z\right] + \frac{\delta'}{\sqrt{n}} \sum_{i=1}^{n} X_i\left[\alpha - I(V_i < 0)\right] + R_n(\delta). \quad \text{(A.14)}$$

Applying formula (A.1), if $V_i \neq 0$, we have

$$\begin{aligned}
&\rho_\alpha\left(Y_i - \widetilde{\theta}_\alpha(Z_i) - X_i'\beta_\alpha\right) - \rho_\alpha(V_i) \\
=& \left[\widetilde{\theta}_\alpha(Z_i) - \theta_\alpha(Z_i) + \frac{\delta' X_i}{\sqrt{n}}\right]\left[\alpha - I(V_i < 0)\right] \\
&+ \int_0^{\widetilde{\theta}_\alpha(Z_i) - \theta_\alpha(Z_i) + \frac{\delta' X_i}{\sqrt{n}}} \left[I(V_i \le t) - I(V_i \le 0)\right] dt
\end{aligned}$$

and

$$\begin{aligned}
&\rho_\alpha\left(\widetilde{Y}_i^*\right) - \rho_\alpha(V_i) \\
=& \left[\widetilde{\theta}_\alpha(Z_i) - \theta_\alpha(Z_i)\right]\left[\alpha - I(V_i < 0)\right] \\
&+ \int_0^{\widetilde{\theta}_\alpha(Z_i) - \theta_\alpha(Z_i)} \left[I(V_i \le t) - I(V_i \le 0)\right] dt.
\end{aligned}$$

Hence, if $V_i \neq 0$

$$\begin{aligned}
&\rho_\alpha\left(Y_i - \widetilde{\theta}_\alpha(Z_i) - X_i'\beta_\alpha\right) - \rho_\alpha\left(\widetilde{Y}_i^*\right) \quad \text{(A.15)} \\
=& \frac{\delta' X_i}{\sqrt{n}}\left[\alpha - I(V_i < 0)\right] + \int_{\widetilde{\theta}_\alpha(Z_i) - \theta_\alpha(Z_i)}^{\widetilde{\theta}_\alpha(Z_i) - \theta_\alpha(Z_i) + \frac{\delta' X_i}{\sqrt{n}}} \left[I(V_i \le t) - I(V_i \le 0)\right] dt.
\end{aligned}$$

By (A.2), we have

$$
\begin{aligned}
& E\left(\widehat{Q}_n^*(\delta)\,|X, Z\right) \\
= & \sum_{i=1}^{n} \int_{\widetilde{\theta}_\alpha(Z_i)-\theta_\alpha(Z_i)}^{\widetilde{\theta}_\alpha(Z_i)-\theta_\alpha(Z_i)+\frac{\delta' X_i}{\sqrt{n}}} \left[F_v(t|I_i) - F_v(0|I_i)\right] dt + o_p(1) \\
= & \frac{\delta'}{2n}\left(\sum_{i=1}^{n} f_v(0|I_i)\, X_i X_i'\right)\delta + \frac{\delta'}{\sqrt{n}}\sum_{i=1}^{n} f_v(0|I_i)\, X_i\left(\widetilde{\theta}_\alpha(Z_i) - \theta_\alpha(Z_i)\right) \\
& + \sum_{i=1}^{n} \int_{\widetilde{\theta}_\alpha(Z_i)-\theta_\alpha(Z_i)}^{\widetilde{\theta}_\alpha(Z_i)-\theta_\alpha(Z_i)+\frac{\delta' X_i}{\sqrt{n}}} \frac{f_v'\left(\bar{t}|I_i\right)}{2} t^2 dt + o_p(1) \\
= & \frac{\delta'}{\sqrt{n}}\sum_{i=1}^{n} f_v(0|I_i)\, X_i\left(\widetilde{\theta}_\alpha(Z_i) - \theta_\alpha(Z_i)\right) \\
& + \frac{\delta'}{2n}\left(\sum_{i=1}^{n} f_v(0|I_i)\, X_i X_i'\right)\delta + o_p(1),
\end{aligned}
\tag{A.16}
$$

where $\bar{t}$ lies between t and 0, and

$$
\begin{aligned}
& \left|\sum_{i=1}^{n} \int_{\widetilde{\theta}_\alpha(Z_i)-\theta_\alpha(Z_i)}^{\widetilde{\theta}_\alpha(Z_i)-\theta_\alpha(Z_i)+\frac{\delta' X_i}{\sqrt{n}}} f_v'\left(\bar{t}|I_i\right) t^2 dt\right| \\
\leq\; & M \sum_{i=1}^{n}\left(\frac{\delta' X_i X_i' \delta}{n} + \left(\widetilde{\theta}_\alpha(Z_i) - \theta_\alpha(Z_i)\right)^2\right)\left|\frac{\delta' X_i}{\sqrt{n}}\right| \\
=\; & o_p(1) + O_p\left(\|H\|^{2(p+1)} + \frac{1}{n\,|H|}\right) n^{1-v} \\
=\; & o_p(1),
\end{aligned}
$$

if $n^{1-v}\|H\|^{2(p+1)} \to 0$, $\frac{1}{n^v|H|} \to 0$ as $n \to \infty$, since $\max_i\left|\frac{\delta' X_i}{\sqrt{n}}\right| = o_p\left(n^{-v}\right)$ for $0 < v < \frac{1}{2}$ by (A.8) and $\frac{1}{n}\sum_{i=1}^{n}\left(\widetilde{\theta}_\alpha(Z_i) - \theta_\alpha(Z_i)\right)^2 = O_p\left(\|H\|^{2(p+1)} + \frac{1}{n|H|}\right)$.

Next, we need to show that $R_n(\delta) = o_p(1)$. We have $E[R_n(\delta)] = 0$, and by (A.1) and (A.2), $R_n(\delta)$ can be rewritten as

$$R_n(\delta) = \sum_{i=1}^{n} [T_i - E(T_i|X,Z)] + o_p(1), \tag{A.17}$$

$$T_i = \int_{\widetilde{\theta}_\alpha(Z_i)-\theta_\alpha(Z_i)}^{\widetilde{\theta}_\alpha(Z_i)-\theta_\alpha(Z_i)+\frac{\delta' X_i}{\sqrt{n}}} [I(V_i \le t) - I(V_i \le 0)]\,dt, \tag{A.18}$$

then

$$\begin{aligned} & E\left(R_n^2(\delta)\right) \\ = & \sum_{i=1}^{n} E[T_i - E(T_i|X,Z)]^2 + o(1) \le nE\left(T_i^2\right) \\ \le & nE\left\{ \iint_{\widetilde{\theta}_\alpha(Z_i)-\theta_\alpha(Z_i)}^{\widetilde{\theta}_\alpha(Z_i)-\theta_\alpha(Z_i)+\frac{\delta' X_i}{\sqrt{n}}} \Pr(0 < |V_i| < \max(|t|,|s|)\,|X,Z)\,dtds \right\} \\ \le & nE\left[\Pr\left(0 < |V_i| < \left|\widetilde{\theta}_\alpha(Z_i) - \theta_\alpha(Z_i)\right| + \left|\frac{\delta' X_i}{\sqrt{n}}\right| \middle| X,Z\right) \frac{\delta' X_i X_i' \delta}{n}\right] \\ = & o(1), \end{aligned}$$

since $\left|\widetilde{\theta}_\alpha(Z_i) - \theta_\alpha(Z_i)\right| = O_p\left(\|H\|^{p+1}\right) + O_p\left(\frac{1}{\sqrt{n|H|}}\right)$, and $\left|\frac{\delta' X_i}{\sqrt{n}}\right| = o_p(n^{-v})$ again by (A.8).

Next, since

$$\begin{aligned} & \frac{\delta'}{\sqrt{n}} \sum_{i=1}^{n} X_i f_v(0|X_i,Z_i)\, E\left(\widetilde{\theta}_\alpha(Z_i) - \theta_\alpha(Z_i)\,|X,Z\right) \\ = & O_p\left(\sqrt{n}\,\|H\|^{p+1}\right) = o_p(1), \end{aligned} \tag{A.19}$$

if $\sqrt{n}\,\|H\|^{p+1} \to 0$ as $n \to \infty$, then

$$\widehat{Q}_n^*(\delta) = \delta' D_n + \delta' C_n \delta/2 + r_n(\delta), \tag{A.20}$$

where $r_n(\delta) = o_p(1)$, and

$$\begin{aligned}
D_n &= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} X_i \left\{ \alpha - I(V_i < 0) + f_v(0|I_i) \left[\widetilde{\theta}_\alpha(Z_i) - \theta_\alpha(Z_i) \right.\right. \\
&\quad \left.\left. - E\left(\widetilde{\theta}_\alpha(Z_i) - \theta_\alpha(Z_i) | X, Z \right) \right] \right\}, \\
C_n &= \frac{1}{n} \sum_{i=1}^{n} f_v(0|I_i) X_i X_i' \xrightarrow{p} C = E\left(f_v(0|X,Z) XX'\right).
\end{aligned}$$

Then $\sup_\delta r_n(\delta) = o_p(1)$ holds by the convexity lemma of Pollard (1991). It follows that $\widehat{\delta}_n = \sqrt{n}\left(\widehat{\beta}_\alpha - \beta_{\alpha,0}\right) = -C^{-1}D_n + o_p(1)$ holds. It implies that $\widehat{\beta}_\alpha \xrightarrow{p} \beta_{\alpha,0}$, and

$$\sqrt{n}\left(\widehat{\beta}_\alpha - \beta_{\alpha,0}\right) \xrightarrow{d} N(0, \Omega_0), \tag{A.21}$$

where $\Omega_0 = C^{-1}DC^{-1}$, with $D = \lim_{n\to\infty} E[D_n D_n']$ if $D_n \xrightarrow{d} N(0,D)$ as $n \to \infty$.

Next, we are going to show that $D_n \xrightarrow{d} N(0,D)$. First, by the results of Theorem 1, we have

$$\begin{aligned}
M_i &= \widetilde{\theta}_\alpha(Z_i) - \theta_\alpha(Z_i) - E\left(\widetilde{\theta}_\alpha(Z_i) - \theta_\alpha(Z_i) | X, Z\right) \\
&= \frac{e_1' S^{-1}(Z_i)}{n|H|} \sum_{m \neq i} K_{mi} W_{mi} \left[I(V_m < \xi_{mi}) - F_v(\xi_{mi}|I_m, I_i) \right], \\
\sup_{m,i} |\xi_{mi}| &= O\left(\|H\|^{p+1} \right),
\end{aligned}$$

where

$$\begin{aligned}
K_{mi} &= K\left(H^{-1}(Z_m - Z_i)\right); \text{ and } K_{mi} = 0 \text{ if } m = i, \\
W_{mi} &= \begin{pmatrix} 1 & w_{im}' Q^{-1} & X_m' \end{pmatrix}', \\
\xi_{mi} &= \theta_\alpha(Z_i) + \tau_i' w_{im} - \theta_\alpha(Z_m), \\
e_1 &= \begin{pmatrix} 1 & 0 & \cdots & 0 \end{pmatrix}' \text{ a } (m+k+1) \times 1 \text{ vector.}
\end{aligned}$$

Then

$$
\begin{aligned}
\frac{2D_n}{\sqrt{n}} &= \frac{1}{n(n-1)} \sum_{i=1}^{n} \sum_{m \neq i} H_n(A_i, A_m) \\
H_n(A_i, A_m) &= \widetilde{H}_n(A_i, A_m) + \widetilde{H}_n(A_m, A_i) \\
\widetilde{H}_n(A_i, A_m) &= X_i \left\{ \alpha - I(V_i < 0) + f_v(0|I_i) \frac{e_1' S^{-1}(Z_i)(n-1)}{n|H|} \right. \\
&\quad \left. \times K_{mi} W_{mi} \left[I(V_m < \xi_{mi}) - F_v(\xi_{mi}|I_m, I_i) \right] \right\},
\end{aligned}
$$

where $A_i = (X_i, Z_i, V_i)$. Thus $H_n(A_i, A_m)$ is a symmetric function and $E[H_n(A_i, A_m)] = 0$. Since

$$
E\left[\|H_n(A_i, A_m)\|^2\right] = O\left(|H|^{-1}\right) = o(n),
$$

if $n|H| \to \infty$ as $n \to \infty$, then, by the Lemma 3.1 in Powell et al. (1989), we obtain

$$
\begin{aligned}
\frac{2D_n}{\sqrt{n}} &= Er_n(A_i) + \frac{2}{n} \sum_{i=1}^{n} \left[r_n(A_i) - Er_n(A_i) \right] + o_p\left(\frac{1}{\sqrt{n}}\right) \\
&= \frac{2}{n} \sum_{i=1}^{n} r_n(A_i) + o_p\left(\frac{1}{\sqrt{n}}\right),
\end{aligned}
$$

where $r_n(A_i) = E[H_n(A_i, A_m)|A_i] \neq 0$ but $Er_n(A_i) = 0$. Simple calculations show that

$$
\begin{aligned}
& E\left[r_n(A_i) r_n(A_i)'\right] \\
&= \alpha(1-\alpha) E\left\{ \left[X_i - E\left(\widetilde{X}_m|I_i\right)\right] \left[X_i - E\left(\widetilde{X}_m|I_i\right)\right]' \right\} + o(1) \\
&= D + o_p(1),
\end{aligned}
$$

where $\widetilde{X}_m = X_m f_v(0|Z_i, X_m) f(Z_i|X_m) e_1' S^{-1}(Z_i) \begin{pmatrix} 1 & \lambda(K)' & X_i' \end{pmatrix}'$. Finally, the multivariate CLT for i.i.d. samples yields $D_n = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} r_n(A_i) \xrightarrow{d} N(0, D)$, as $n \to \infty$, which will complete the proof of this theorem. ▌

REFERENCES

Bhattacharya, P.K. and A.K. Gangopadtyay, 1990, Kernel and nearest-neighbor estimation of a conditional quantile. *The Annals of Statistics* **18(3)**, 1400-1415.

Buchinsky, M., 1994, Changes in the U.S. wage structure 1963-1987: Application of quantile regression. *Econometrica* **62**, 405-458.

Cai, Zongwu, 2002, Regression quantiles for time series. *Econometric Theory* **18**, 169-192.

Chen, Songnian and Shakeeb Khan, 2001, Semiparametric estimation of a partially linear censored regression model. *Econometric Theory* **17**, 567-590.

Fan, J., T. Hu, and Y.K. Truong, 1994, Robust non-parametric function estimation. *Scandinavian Journal of Statistics* **21**, 433-446.

Hall, Peter, Rodney C. Wolff, and Qiwei Yao, 1999, Method for estimating conditional distribution function. *Journal of the American Statistical Association* **94**, 154-163.

He, Xuming and Peide Shi, 1996, Bivariate tensor-product B-splines in a partially linear regression. *Journal of Multivariate Analysis* **58**, 162-181.

Khan, Shakeeb, 2001, Two-stage rank estimation of quantile index models. *Journal of Econometrics* **100**, 319-355.

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

Koenker, R. and G. Bassett, 1982, Robust test for heteroskedasticity based on regression quantiles. *Econometrica* **50**, 43-61.

Koenker, R. and V. d'Orey, 1987, Computing regression quantiles. *Applied Statistics* **36**, 383-393.

Koenker, R., P. Ng, and S. Portnoy, 1994, Quantile smoothing spline. *Biometrika* **81(4)**, 673-680.

Lee, Sokbae, 2003, Efficient semiparametric estimation of a partially linear quantile regression model. *Econometric theory* **19**, 1-31.

Newey, Whitney K., 1990, Semiparametric efficient bounds. *Journal of Applied Econometrics* **5(2)**, 99-135.

Newey, Whitney K., 1994, The asymptotic variance of semiparametric estimators. *Econometrica* **62(6)**, 1349-1382.

Newey, Whitney K. and J.L. Powell, 1990, Efficient estimation of linear and type I censored regression models under conditional quantile restrictions. *Econometric Theory* **6(2)**, 295 - 317.

Pollard, David, 1991, Asymptotic for least absolute deviation regression estimators. *Econometric Theory* **7**, 186-199.

Powell, J.L., 1984, Least absolute deviations estimation for the censored regression model. *Journal of Econometrics* **25**, 303-325.

Powell, J.L., J.H. Stock, and T.M. Stoker, 1989, Semiparametric estimation of index coefficients. *Econometrica* **57**, 1403-1430.

Yu, Keming and M.C. Jones, 1998. Local linear quantile regression. *Journal of the American Statistical Association* **93**, 228-237.