NONPARAMETRIC INTERMEDIATE ORDER REGRESSION QUANTILES

HIDEHIKO ICHIMURA, TAISUKE OTSU, AND JOSEPH G. ALTONJI

Abstract. This paper studies nonparametric estimation of d-dimensional conditional quantile functions and their derivatives in the tails. We investigate asymptotic properties of the local and global nonparametric quantile regression estimators proposed by Chaudhuri (1991a, b), respectively, under the intermediate order quantile asymptotics: as the sample size n goes to infinity, the quantile α_n and a bandwidth parameter δ_n satisfy $\alpha_n \to 0$ and $n\delta_n^d\alpha_n \to \infty$ (or $\alpha_n \to 1$ and $n\delta_n^d(1-\alpha_n) \to \infty$). We derive the pointwise convergence rate and asymptotic distribution of the local nonparametric quantile regression estimator, and the sup-norm convergence rate of the global nonparametric quantile regression estimator. Our results complement the papers by Chaudhuri (1991a, b), where the quantile α_n does not vary with n, and Chernozhukov (1998), where the quantile α_n satisfies $\alpha_n \to 0$ and $n\delta_n^d\alpha_n \to 0$.

1. Introduction

Since the seminal work of Koenker and Bassett (1978), quantile regression has become a standard tool for statistical data analysis. See Koenker (2005) for a review. Compared to the conventional mean regression that focuses on the conditional mean of a response variable given a vector of conditioning variables, quantile regression enables us to focus on the conditional quantiles at different quantiles. This paper studies nonparametric estimation of the d-dimensional conditional quantile functions and their derivatives in the tails. In particular, we investigate asymptotic properties of the local and global nonparametric quantile regression estimators proposed by Chaudhuri (1991a) and (1991b), respectively, under the intermediate order quantile asymptotics: as the sample size n goes to infinity, the quantile α_n and a bandwidth parameter δ_n satisfy $\alpha_n \to 0$ and $n\delta_n^d\alpha_n \to \infty$.[1] We derive the pointwise and uniform convergence rate and asymptotic distribution of the local polynomial nonparametric quantile regression estimator, and the L^∞-convergence rate of the global nonparametric quantile regression estimator. Our results complement the papers by Chaudhuri (1991a, b), where the quantile α_n does not vary with n, and Chernozhukov (1998), where the quantile α_n satisfies $\alpha_n \to 0$ and $n\delta_n^d\alpha_n \to 0$.

Both Chaudhuri (1991a, b) and Chernozhukov (1998) approach the asymptotic analysis exploiting the linear in parameter structure of the problem analogous to the approach taken by Koenker and Bassett (1978). Our method of proof follows the standard approach to examining non-linear extremum estimators. The approach allows us to establish the asymptotic results allowing for general kernel functions rather than the uniform kernel function used by Chaudhuri (1991a, b) and Chernozhukov (1998). The approach also allows us to exploit the result by Arcones (1997) to improve the pointwise convergence rate obtained by Chaudhuri (1991a).

Otsu gratefully acknowledges financial support from the ERC Consolidator Grant (SNP 615882) (Otsu).

[1] In this paper, we focuses on the lower tails, i.e., $\alpha_n \to 0$. The same argument applies to the upper tails i.e., $\alpha_n \to 1$.

In characterizing the asymptotic distribution, Chernozhukov (1998) parameterizes the tail index of the conditional distribution function of the response variable given the conditioning variables. Instead we show the asymptotic distribution is characterized by a more standard assumption of the conditional density function of the response variable given the conditioning variables being Hölder continuous.

Several previous papers studied asymptotic behaviors of the linear quantile regression estimator under the extreme order quantile asymptotics (i.e., $\alpha_n \to 0$ and $n\alpha_n \to 0$), such as Feigin and Resnick (1994), Smith (1994), Portnoy and Jurečková (1999), and Knight (2001). Chernozhukov (2005) derived the asymptotic distributions of the linear quantile regression estimator under the intermediate order quantile asymptotics (i.e., $\alpha_n \to 0$ and $n\alpha_n \to \infty$) and the border-case asymptotics (i.e., $\alpha_n \to 0$ and $n\alpha_n \to c > 0$). This paper extends Chernozhukov's (2005) result for the intermediate order linear quantile regression to a nonparametric setup.

This paper is organized as follows. Section 2 introduces the setup and nonparametric estimators. Section 3 presents our main results. All proofs are contained in the Appendix.

2. Setup and estimators

Consider an iid sample $\{Y_i, X_i\}_{i=1}^n$ from the distribution of $(Y, X) \in \mathbb{R} \times \mathbb{R}^d$. Let $F_Y(y|x_r)$ be the conditional distribution function of Y given $X = x_r$ and $\theta_\alpha(x_r) = \inf_{y\in\mathbb{R}}\{y : F_Y(y|x_r) > \alpha\}$ be the α-th conditional quantile function at x_r. We first define the local nonparametric quantile regression estimator for the conditional quantile function $\theta_\alpha(x_r)$ and its derivatives at x_r. We momentarily assume that the quantile point $\alpha \in (0, 1)$ is fixed. Let $K : \mathbb{R}^d \to \mathbb{R}$ be a kernel function and $\{\delta_n\}_{n\in\mathbb{N}}$ be a sequence of bandwidth parameters satisfying $\delta_n \to 0$ as $n \to \infty$. Detailed requirements on K and δ_n are given later. Let $u = (u_1, \ldots, u_d)' \in \mathbb{Z}_+^d$ be a vector of non-negative integers, $|u| = \sum_{j=1}^d u_j$, $A = \{u \in \mathbb{Z}_+^d : |u| \leq k\}$, and $s(A)$ denote the number of elements in the set A. Define $z^u = \prod_{j=1}^d z_j^{u_j}$ and $u! = \prod_{j=1}^d u_j!$ with the convention that $0! = 1$ for each $z \in \mathbb{R}^d$ and $u \in \mathbb{Z}_+^d$. We use the k-th order polynomial defined as $P_n(\beta, z) = \sum_{u\in A} \beta_u z^u/u!$, where $z \in \mathbb{R}^d$ and $\beta = \{\beta_u\}_{u\in A}$ is an $s(A) \times 1$ coefficient vector. For concreteness we order the terms $\beta_u z^u/u!$ in the increasing order of $|u|$ and within the same size of $|u|$, lexicographically in the increasing order of the index j and in the increasing order of u_j. Based on this notation, we consider the following (weighted) quantile regression problem:

$$\hat{\beta}_n^{(\alpha)} = \arg \min_{\beta\in\mathbb{R}^{s(A)}} \sum_{i=1}^n K\left(\frac{X_i - x_r}{\delta_n}\right) \rho_\alpha(Y_i - X_{in}'\beta), \tag{2.1}$$

where $X_{in} = \{(X_i - x_r)^u/\delta_n^{|u|}\}_{u\in A}$, $\rho_\alpha(v) = (\alpha - \mathbb{I}\{v \leq 0\})v$ is called the check function, and $\mathbb{I}\{\cdot\}$ is the indicator function. This minimization problem can be considered as a natural extension of the local polynomial (mean) regression problem to the quantile regression setup (see, e.g., Fan and Gijbels, 1996). For $u \in A$, define the differentiation as $D^u\theta_\alpha(x) = \partial^{|u|}\theta_\alpha(x)/\partial x_1^{u_1} \cdots \partial x_d^{u_d}$. The local nonparametric quantile regression estimator for $D^u\theta_\alpha(x_r)$ is defined as

$$\widehat{D^u\theta}_\alpha(x_r) = D^u P_n(\hat{\beta}_n^{(\alpha)}, (x - x_r)/\delta_n)\Big|_{x=x_r} = \delta_n^{-|u|}\hat{\beta}_{n,u}^{(\alpha)},$$

for each $u \in A$. If $u = (0, \ldots, 0)$, then $D^u\theta_\alpha(x_r)$ becomes the conditional quantile function $\theta_\alpha(x_r)$ and its estimator $\hat{\theta}_\alpha(x_r)$ is obtained as the constant term of the fitted polynomial $P_n(\hat{\beta}_n^{(\alpha)}, (X_i - x_r)/\delta_n)$. This paper studies asymptotic properties of $\hat{\beta}_n^{(\alpha)}$ and thus $\widehat{D^u\theta}_\alpha(x_r)$ when the quantile point α depends on the sample size n and converges to 0 or 1 as $n \to \infty$.

We next introduce the global nonparametric quantile regression estimator for the conditional quantile function and its derivatives over some compact and convex subset $\mathbb{C} \subset \mathbb{R}^d$. Hereafter, without loss of generality we assume that $\mathbb{C} = [-1, 1]^d$. Our results apply to any compact and convex subset of the support of X after adequate normalization. Let $\{J_n\}_{n\in\mathbb{N}}$ be a sequence of positive integers satisfying $J_n \to \infty$ as $n \to \infty$. We split the cube $\mathbb{C} = [-1, 1]^d$ by J_n^d equal-size subcubes $\{\mathbb{C}_{n,r}\}_{r=1}^{J_n^d}$ with centers $\{x_{n,r}\}_{r=1}^{J_n^d}$ and equal side length $2J_n^{-1}$. Let $\delta_n = (2J_n)^{-1}$ and consider the local quantile regression problem at $x_{n,r}$:

$$\hat{\beta}_{n,r}^{(\alpha)} = \arg \min_{\beta\in\mathbb{R}^{s(A)}} \sum_{i=1}^{n} K\left(\frac{X_i - x_{n,r}}{\delta_n}\right) \rho_\alpha(Y_i - P_n(\beta, (X_i - x_{n,r})/\delta_n)), \tag{2.2}$$

for $r = 1, \ldots, J_n^d$. Using the local estimators $\{\hat{\beta}_{n,r}^{(\alpha)}\}_{r=1}^{J_n^d}$, the global nonparametric quantile regression estimator for $D^u\theta_\alpha(x)$ over the set $\mathbb{C} = [-1, 1]^d$ is defined as

$$\widetilde{D^u\theta}_\alpha(x) = D^u P_n(\hat{\beta}_{n,r}^{(\alpha)}, (x - x_{n,r})/\delta_n) = D^u \sum_{\tilde{u}\in A} \hat{\beta}_{n,r,\tilde{u}}^{(\alpha)} [(x - x_{n,r})/\delta_n]^{\tilde{u}}/\tilde{u}!,$$

when x is an interior point of the subcube $\mathbb{C}_{n,r}$. For boundary points of $\mathbb{C}_{n,r}$, we may define $\widetilde{D^u\theta}_\alpha(x)$ as the average of the different values of $\widetilde{D^u\theta}_\alpha(x)$ computed from different adjacent subcubes. Intuitively, to estimate $D^u\theta_\alpha(x)$ at $x \in \mathbb{C}_{n,r}$, the global estimator invokes on the polynomial interpolation using the coefficients estimated by the local polynomial estimator in (2.2).

Chaudhuri (1991a, b) considered the case of the uniform kernel function (i.e., $K(a) = \mathbb{I}\{|a| \leq 1/2\}$) and explored the asymptotic properties of the local and global nonparametric quantile regression estimators, respectively, when the quantile $\alpha \in (0, 1)$ is fixed. Chernozhukov (1998) investigated the asymptotic properties of those estimators under the extreme order quantile asymptotics, where the quantile α_n depends on the sample size and satisfies $\alpha_n \to 0$ and $n\delta_n^d\alpha_n \to 0$ as $n \to \infty$.[2] This paper focuses on an intermediate between these two cases called the intermediate order quantile asymptotics, where the quantile α_n satisfies $\alpha_n \to 0$ and $n\delta_n^d\alpha_n \to \infty$ as $n \to \infty$, i.e., α_n converges to zero at a slower rate than the extreme order quantile asymptotics. Let $\hat{\beta}_n^{(\alpha_n)} = \hat{\beta}_n$ and $\hat{\beta}_{n,r}^{(\alpha_n)} = \hat{\beta}_{n,r}$ in (2.1) and (2.2), respectively. The local and global nonparametric quantile regression estimators are respectively written as

$$\begin{aligned} \widehat{D^u\theta}_{\alpha_n}(x_r) &= D^u P_n(\hat{\beta}_n, (x - x_r)/\delta_n)\Big|_{x=x_r} = \delta_n^{-|u|}\hat{\beta}_{n,u}, \\ \widetilde{D^u\theta}_{\alpha_n}(x) &= D^u P_n(\hat{\beta}_{n,r}, (x - x_{n,r})/\delta_n) = D^u \sum_{\tilde{u}\in A} \hat{\beta}_{n,r,\tilde{u}}[(x - x_{n,r})/\delta_n]^{\tilde{u}}/\tilde{u}!. \end{aligned} \tag{2.3}$$

[2]It is known that under the extreme order quantile asymptotics, the quantile regression problem in (2.1) (and (2.2) as well) is asymptotically equivalent to the following linear programming problem:

$$\max_{\beta\in\mathbb{R}^{s(A)}} \sum_{i=1}^{n} K\left(\frac{X_i - x_r}{\delta_n}\right) P_n(\beta, (X_i - x_r)/\delta_n), \quad \text{s.t. } P_n(\beta, (X_i - x_r)/\delta_n) \leq Y_i \text{ for all } i = 1, \ldots, n.$$

The main purpose of this paper is to derive the convergence rate and asymptotic distribution of the local estimator $\widehat{D^u\theta}_{\alpha_n}(x_r)$ at x_r, and the L^∞-convergence rate of the global estimator $\widetilde{D^u\theta}_{\alpha_n}(x)$ over $x \in \mathbb{C}$.

3. Asymptotic properties of estimators

We split into two cases: when the support of the conditional distribution $Y|X = x_r$ is bounded (Section 3.1) and unbounded (Section 3.2), and derive the asymptotic properties of the nonparametric estimators $\widehat{D^u\theta}_{\alpha_n}(x_r)$ and $\widetilde{D^u\theta}_{\alpha_n}(x)$.

3.1. **Bounded support case .** We first introduce some notation. Define

$$x_n = \{(x - x_r)^u / \delta_n^{|u|}\}_{u \in A}, \qquad \beta_{n,u} = D^u\theta_{\alpha_n}(x_r)\delta_n^{|u|}/u!, \qquad \beta_n = \{\beta_{n,u}\}_{u \in A}.$$

Let $f_Y(y|x_r)$ be the conditional density function of Y given $X = x_r$, and

$$\phi_n = f_Y(\theta_{\alpha_n}(x_r)|x_r),$$

which plays a key role to derive the convergence rate of the nonparametric quantile regression estimators. Intuitively, ϕ_n captures the amount of the conditional density of Y given $X = x_r$ in the tail. When the quantile point is fixed (i.e., $\alpha_n = \alpha$), this tail parameter ϕ_n is typically a positive constant. On the other hand, if the quantile point α_n drifts to zero as n increases, ϕ_n may decay to zero or diverge depending on the limiting behavior of $\theta_{\alpha_n}(x_r)$ and the shape of the density function $f_Y(y|x_r)$ at the tail.

We derive the asymptotic properties of our estimators under following assumptions:

Assumption D. $\{Y_i, X_i\}_{i=1}^n$ *is an iid sample from* $(Y, X) \in \mathbb{R} \times \mathbb{R}^d$, X *is absolutely continuous in a neighborhood* $\mathbb{B}(x_r)$ *around* x_r, *and the density function* f_X *of* X *is positive and continuous on* $\mathbb{B}(x_r)$.

Assumption K. *The kernel function* K *is a continuous density function with support* $[-1, 1]^d$, *and* $Q = f_X(x_r) \int_{[-1,1]^d} K(w)w(A)w(A)'dw$ *is positive definite, where* $w(A) = \{w^u\}_{u \in A}$ *is an* $s(A) \times 1$ *vector for* $w \in \mathbb{R}^d$.

These assumptions are common for both bounded and unbounded support cases. Assumption D is a standard requirement on the distribution of data and is imposed by Chaudhuri (1991a) for example. Although the convergence rate will be different, it may be possible to allow time series data. To this end, we need different versions of exponential type inequalities for dependent data (see, e.g., Bosq, 1998). The first condition in Assumption K is on the form of the kernel function K and is satisfied by e.g., the uniform and quadratic (or Epanechnikov) kernels. The second condition corresponds to the exclusion of multicollinearity for linear regression in this setting.

In this section, we consider the case where the support of the conditional distribution $Y|X = x_r$ is bounded and impose the following assumptions. Let $\mathbb{C}_n = \prod_{j=1}^d [x_{r,j} - \delta_n, x_{r,j} + \delta_n]^d$ be a d-dimensional cube with center x_r and side length $2\delta_n$

Assumption C1. *(i) The lower endpoint* $\theta_0(x) = \max_{y \in \mathbb{R}}\{y : F_Y(y|x) = 0\}$ *exists for all* $x \in \mathbb{B}(x_r)$. *(ii) For all* $u \in A$ *and all* n *large enough,* $D^u\theta_{\alpha_n}(x)$ *exists and is continuous at* *each* $x \in \mathbb{C}_n$. *(iii) As* $n \to \infty$, *it holds* $\alpha_n \to 0$, $\delta_n \to 0$, $n\delta_n^d \to \infty$, *and* $\sup_{x\in\mathbb{C}_n} |\{\theta_{\alpha_n}(x) - x_n'\beta_n\}/\{\theta_{\alpha_n}(x) - \theta_0(x)\}| \to 0$.

Assumption C1 (i) means that we are concerned with the bounded support case. Assumption C1 (ii), which is on the smoothness of the conditional quantile function $\theta_{\alpha_n}(x)$, is natural when we are interested in estimating the derivatives $D^u\theta_{\alpha_n}(x)$. Assumption C1 (iii) says that the remainder of the Taylor expansion of $\theta_{\alpha_n}(x)$ around $x = x_r$ should be smaller order than $\theta_{\alpha_n}(x) - \theta_0(x)$.

Assumption F1. *As* $\alpha \searrow 0$,

$$\sup_{|x-x_r|\le\eta_\alpha} \left| \frac{f_Y(\theta_\alpha(x) + \Delta_\alpha(x)|x)}{f_Y(\theta_\alpha(x_r)|x_r)} - 1 \right| \to 0,$$

for all $\eta_\alpha > 0$ *and* $\Delta_\alpha(x)$ *such that* $\eta_\alpha \to 0$ *and* $\sup_{|x-x_r|\le\eta_\alpha} |\Delta_\alpha(x)/\{\theta_\alpha(x) - \theta_0(x)\}| \to 0$.

This is a differentiability condition of $f_Y(y|x)$ in the (left) tail. This assumption needs to be verified for each setup. See below for a specific example.

To derive the convergence rate of the estimator, we add the following assumption.

Assumption R1. *As* $n \to \infty$, *it holds* $\frac{n\delta_n^d\phi_n^2}{\alpha_n \log\log n} \to \infty$, $\sqrt{\frac{n\delta_n^d\phi_n^2}{\alpha_n}} \sup_{x\in\mathbb{C}_n} |\theta_{\alpha_n}(x) - x_n'\beta_n| \to 0$, $\frac{\alpha_n \log\log n}{n\delta_n^d\phi_n^2} \sup_{x\in\mathbb{C}_n} \left|\frac{1}{\theta_{\alpha_n}(x)-\theta_0(x)}\right|^2 \to 0$, *and* $n\delta_n^d\alpha_n \log\log n$ *is bounded away from zero for all* n *large enough.*

Under these assumptions, the asymptotic properties of the local polynomial quantile regression estimator is obtained as follows. Let $Q_1 = f_X(x_r)\int_{w\in[-1,1]^d} K(w)^2 w(A)w(A)'dw$, and ι_u be a $s(A) \times 1$ vector which takes one for the element corresponding to $u \in A$ and takes zero for other elements.

Theorem 1 (Bounded case: Local estimator)**.** *Suppose that Assumptions D, K, C1, and F1 hold.*

- **(i):** ***Consistency:*** $|\hat{\beta}_n - \beta_n| = o_{as}(1)$.
- **(ii):** ***Convergence rate:*** *If Assumption R1 additionally holds, then*

$$|\widehat{D^u\theta}_{\alpha_n}(x_r) - D^u\theta_{\alpha_n}(x_r)| = O_{as}\left(\sqrt{\frac{\alpha_n \log\log n}{n\delta_n^{d+2|u|}\phi_n^2}}\right),$$

for each $u \in A$.

- **(iii):** ***Asymptotic distribution:*** *If Assumption R1 additionally holds, then*

$$\sqrt{\frac{n\delta_n^d\phi_n^2}{\alpha_n}}(\hat{\beta}_n - \beta_n) \xrightarrow{d} N(0, Q^{-1}Q_1Q^{-1}),$$

and

$$\sqrt{\frac{n\delta_n^{d+2|u|}\phi_n^2}{\alpha_n}}\left(\widehat{D^u\theta}_{\alpha_n}(x_r) - D^u\theta_0(x_r)\right) \xrightarrow{d} N(0, \iota_u'Q^{-1}Q_1Q^{-1}\iota_u).$$

Theorem 1 (i) is on the (strong) consistency of $\hat{\beta}_n$ to β_n. Since $\widehat{D^u\theta}_{\alpha_n}(x_r) = \delta_n^{-|u|}\hat{\beta}_{n,u}$ for $u \in A$, this consistency does not necessarily guarantee the consistency of $\widehat{D^u\theta}_{\alpha_n}(x_r)$ to $D^u\theta_{\alpha_n}(x_r)$.

Theorem 1 (ii) derives the almost sure convergence rate of $\widehat{D^u\theta}_{\alpha_n}(x_r)$ to $D^u\theta_{\alpha_n}(x_r)$. The convergence rate depends on the dimension of covariates (d), the total order of derivatives ($|u|$), the behavior of the tail parameter (ϕ_n), and the convergence rates of the quantile point (α_n) and bandwidth parameter (δ_n). If d is higher, the convergence rate is slower, i.e., the curse of dimensionality emerges. If $|u|$ is higher, the convergence rate also becomes slower, i.e., higher order derivatives are more difficult to estimate. The term ϕ_n^2/α_n appearing the convergence rate is distinctive to the intermediate order quantile asymptotics. This term reveals how the thickness of the tail ϕ_n (relative to $\sqrt{\alpha_n}$) affects the convergence rate of the estimator $\widehat{D^u\theta}_{\alpha_n}(x_r)$ to $D^u\theta_{\alpha_n}(x_r)$. If the quantile point is fixed, then the term ϕ_n^2/α_n is typically fixed and the conventional rate $O_{as}\left(\sqrt{\frac{\log\log n}{n\delta_n^{d+2|u|}}}\right)$ emerges.[3] On the other hand, if the conditional density $f_Y(y|x)$ has a thin tail and ϕ_n^2/α_n decays to zero, then the convergence rate becomes slower than the conventional rate. Also, if the conditional density $f_Y(y|x)$ has a thick tail and ϕ_n^2/α_n diverges to infinity (e.g., ϕ_n converges to a positive constant), then the convergence rate becomes faster than the conventional rate.

Theorem 1 (iii) says that under the intermediate order quantile asymptotics, the estimator $\widehat{D^u\theta}_{\alpha_n}(x_r)$ is asymptotically normal at the $\sqrt{\frac{n\delta_n^{d+2|u|}\phi_n^2}{\alpha_n}}$ rate. It is easy to estimate the asymptotic variance $Q^{-1}Q_1Q^{-1}$, which depends only on the density $f_X(x_r)$ of the covariates and known constant matrices Q and Q_1. However, in order to construct a confidence interval for $D^u\theta_{\alpha_n}(x_r)$, we also have to estimate the tail parameter ϕ_n.

A key condition to the above theorem is Assumption F1. As an example, consider the following model

$$Y = \theta_0(X) + \sigma U, \tag{3.1}$$

where X and U are independent, σ is a positive constant, and $\max_{u\in\mathbb{R}}\{u : F_U(u) = 0\} = 0$ (thus Assumption C1 (i) is satisfied). In this case, since $\theta_\alpha(x) - \theta_0(x) = \sigma F_U^{-1}(\alpha)$, Assumption F1 is written as

$$\sup_{|x-x_r|\le\eta_\alpha}\left|\frac{f_U\left(F_U^{-1}(\alpha)\left(1+\frac{\Delta_\alpha(x)}{\sigma F_U^{-1}(\alpha)}\right)\right)}{f_U(F_U^{-1}(\alpha))} - 1\right| \to 0, \tag{3.2}$$

as $\alpha \searrow 0$ for all η_α and $\Delta_\alpha(x)$ such that $\sup_{|x-x_r|\le\eta_\alpha}\left|\frac{\Delta_\alpha(x)}{F_U^{-1}(\alpha)}\right| \to 0$. Suppose $\frac{\partial F_U^{-1}(\alpha)}{\partial\alpha} = \frac{1}{f_U(F_U^{-1}(\alpha))}$ is regularly varying at 0 with exponent $-\xi-1$ for $\xi > 0$. Then by Resnick (1984, Propositions 0.5 and 0.8), we can see that (3.2) is satisfied.

Finally, the L_∞-convergence rate of the global estimator $\widetilde{D^u\theta}_{\alpha_n}(x)$ over the set $\mathbb{C}$ is obtained as follows.

[3]Since we applied a law of iterated logarithm in the proof, this convergence rate is slightly different from that of Chaudhuri (1991a, Theorem 3.2) which obtained $O_{as}\left(\sqrt{\frac{\log n}{n\delta_n^{d+2|u|}}}\right)$.

Theorem 2 (Bounded case: Global estimator)**.** *Suppose that Assumptions D, K, C1, and R1 hold by replacing $\mathbb{B}(x_r)$ with $\mathbb{C}$, and Assumption F1 holds for each $x_r \in \mathbb{C}$. Then*

$$\sup_{x\in\mathbb{C}} \left| \widetilde{D^u\theta}_{\alpha_n}(x) - D^u\theta_{\alpha_n}(x) \right| = O_{as}\left(\sqrt{\frac{\alpha_n \log n}{n\delta_n^{d+2|u|}\phi_n^2}} \right).$$

The convergence rate is similar to that of the pointwise rate in Theorem 1 (ii) except for the $\log n$ term. Basically same comments to Theorem 1 (ii) apply.

3.2. **Unbounded Support Case .** We now consider the unbounded support case. In this case, although the boundary function $\theta_0(x)$ does not exist, we can still analyze the asymptotic behaviors of the estimation error $\widehat{D^u\theta}_{\alpha_n}(x_r) - D^u\theta_{\alpha_n}(x_r)$. In addition to Assumptions D and K, we impose the following assumptions.

Assumption C2. *(i) For each $x \in \mathbb{B}(x_r)$, the support of the conditional distribution $Y|X = x$ is unbounded from below. (ii) For all $u \in A$ and all n large enough, $D^u\theta_{\alpha_n}(x)$ exists and is continuous at each $x \in \mathbb{B}(x_r)$. (iii) $\sup_{x\in\mathbb{B}(x_r)} \left|\frac{x_n'\beta_n - \theta_{\alpha_n}(x)}{\theta_{\alpha_n}(x)}\right| \to 0$ and $\sup_{x\in\mathbb{B}(x_r)} \theta_{\alpha_n}(x) \to -\infty$ as $n \to \infty$. (iv) $\sqrt{\frac{\alpha_n \log\log n}{n\delta_n^d\phi_n^2}} \sup_{x\in\mathbb{B}(x_r)} \left|\frac{1}{\theta_{\alpha_n}(x)}\right| \to 0$.*

Assumption C2 (i) explicitly states unboundedness of the support of $Y|X = x$ over $x \in \mathbb{B}(x_r)$. Assumptions C2 (ii) is same as Assumption C1 (ii) for the bounded case, and Assumption C2 (iii) is an analog of Assumptions C1 (iii) for the unbounded case.

Assumption F2. *As $\alpha \searrow 0$,*

$$\sup_{|x-x_r|\le\eta_\alpha} \left| \frac{f_Y(\theta_\alpha(x)\{1 + \Delta_\alpha(x)/\theta_\alpha(x)\}|x)}{f_Y(\theta_\alpha(x_r)|x_r)} - 1 \right| \to 0,$$

for $\eta_\alpha > 0$ and $\Delta_\alpha(x)$ such that $\eta_\alpha \to 0$ and $\sup_{|x-x_r|\le\eta_\alpha} |\Delta_\alpha(x)/\theta_\alpha(x)| \to 0$.

Assumption F2, which is an analog of Assumption F1, is on the differentiability of $F_Y(y|x)$ and $f_Y(y|x)$ at $y = \theta_{\alpha_n}(x)$ in a neighborhood of x_r.

To derive the convergence rate of the estimator, we add the following assumption.

Assumption R2. *(i) As $n \to \infty$, $\alpha_n \to 0$, $\delta_n \to 0$, and $\frac{\phi_n}{\alpha_n} \sup_{x\in\mathbb{B}(x_r)} |\theta_{\alpha_n}(x) - x_n'\beta_n| \to 0$. Also $n\delta_n^d\phi_n$ and $\frac{n\delta_n^d\phi_n^2}{\alpha_n}$ are bounded away from zero for all n large enough. (ii) $n\delta_n^d\alpha_n \log\log n$ is bounded away from zero for all n large enough.*

Based on these assumptions, the asymptotic properties of the local nonparametric quantile regression estimator $\widehat{D^u\theta}_{\alpha_n}(x_r)$ for the unbounded support case are obtained as follows.

Theorem 3 (Unbounded case: Local estimator)**.** *Suppose that Assumptions D, K, C2, and F2 hold. Then $|\hat\beta_n - \beta_n| = o_{as}(1)$. Furthermore, if Assumption R2 additionally holds, then the convergence rate and asymptotic distribution in Theorem 1 (ii)-(iii) hold true.*

Similar comments to Theorem Theorem 1 apply.

Finally, the L_∞-convergence rate of the global estimator $\widetilde{D^u\theta}_{\alpha_n}(x)$ over the set $\mathbb{C}$ is obtained as follows.

Theorem 4 (Unbounded case: Global estimator)**.** *Suppose that Assumptions D, K, C2, and R2 hold by replacing $\mathbb{B}(x_r)$ with $\mathbb{C}$, and Assumption F2 holds for each $x_r \in \mathbb{C}$. Then the same convergence rate in Theorem 2 holds true.*

Appendix A. Proofs

Since the proofs are similar, we only show the results for the bounded case in Section 3.1. Hereafter, let c and c_l for $l = 1, 2, \ldots$ be generic positive constants, $K_{in} = K\left(\frac{X_i - x_r}{\delta_n}\right)$. Also the relation "$a_n \sim b_n$" means that $a_n/b_n \to 1$ as $n \to \infty$.

A.1. **Proof of Theorem 1 (i).** The proof is split into three steps.

Step 1: Overall sketch. Using the reparametrization $\Delta = \beta - \beta_n$, the recentered estimator $\hat{\Delta}_n = \hat{\beta}_n - \beta_n$ is obtained as the minimizer of the objective function:

$$\begin{aligned} Q_n(\Delta) &= \sum_{i=1}^{n} K_{in}[\rho_{\alpha_n}(Y_i - X'_{in}\beta_n - X'_{in}\Delta) - \rho_{\alpha_n}(Y_i - X'_{in}\beta_n)] \\ &= -\sum_{i=1}^{n} K_{in}[\alpha_n - \mathbb{I}\{Y_i - X'_{in}\beta_n < 0\}]X'_{in}\Delta \\ &\quad + \sum_{i=1}^{n} K_{in}(\mathbb{I}\{Y_i - X'_{in}\beta_n \le 0\} - \mathbb{I}\{Y_i - X'_{in}\beta_n \le X'_{in}\Delta\})(Y_i - X'_{in}\beta_n - X'_{in}\Delta) \\ &= -Q'_{1n}\Delta + Q_{2n}(\Delta), \end{aligned}$$

where Q_{1n} and $Q_{2n}(\Delta)$ are implicitly defined. Suppose we have

$$\left|\frac{1}{n\delta_n^d} Q_{1n}\right| \overset{a.s.}{\to} 0, \tag{A.1}$$

$$\text{for each } \Delta \neq 0,\ \frac{1}{n\delta_n^d} Q_{2n}(\Delta) \text{ converges a.s. to a strictly positive constant.} \tag{A.2}$$

Then by the convexity of $Q_n(\Delta)$ and $Q_n(0) = 0$, an adapted version of Newey and McFadden (1994, Theorem 2.7) yields the conclusion $\hat{\Delta}_n = o_{as}(1)$. In the following, we show (A.1) in Step 2 and (A.2) in Step 3.

Step 2: Show (A.1). Let $Z_{1i} = K_{in}[\alpha_n - \mathbb{I}\{Y_i - X'_{in}\beta_n < 0\}]X_{in}$ (i.e., $Q_{1n} = \sum_{i=1}^{n} Z_{1i}$) and $B_{1n} = \{|\sum_{i=1}^{n} Z_{1i}| \ge M_1 n\delta_n^d\}$ for $M_1 > 0$. From the Borel-Cantelli lemma, it is sufficient to show that $\sum_{n=1}^{\infty} \Pr\{B_{1n}\} < \infty$ for each $M_1 > 0$. Since $\{Z_{1i}\}_{i=1}^{n}$ is iid (by Assumption D) and $|Z_{1i}|$ is bounded for all $i = 1, \ldots, n$ (by Assumption K), we have

$$\begin{aligned} Var\left(\sum_{i=1}^{n} Z_{1i}\right) &\le c_1 n|E[Z_{1i}Z'_{1i}]| = c_1 n|E[K_{in}^2\{\alpha_n^2 + (1 - 2\alpha_n)F_Y(X'_{in}\beta_n|X_i)\}X_{in}X'_{in}]| \\ &\le c_1 nE[K_{in}^2|X_{in}|^2] = c_1 n\delta_n^d \int_{w \in [-1,1]^d} K(w)^2|w(A)|^2 f_X(x_r + \delta_n w)dw, \end{aligned}$$

for all n large enough, where the first equality follows from the law of iterated expectation, the second inequality follows from $F_Y(X'_{in}\beta_n|X_i) \le 1$ and $\alpha_n \to 0$, and the second equality follows from the change of variables $w = \frac{x - x_r}{\delta_n}$. Similarly, we have

$$\begin{aligned} |E[Z_{1i}]| &= |E[K_{in}\{\alpha_n - F_Y(X_{in}'\beta_n|X_i)\}X_{in}]| \\ &= \delta_n^d \left| \int_{w\in[-1,1]^d} K(w) \left[\begin{array}{c} F_Y(\theta_{\alpha_n}(x_r+\delta_n w)|x_r+\delta_n w) \\ -F_Y(w(A)'\beta_n|x_r+\delta_n w) \end{array} \right] w(A) f_X(x_r+\delta_n w)dw \right| \\ &\leq \delta_n^d \left| \int_{w\in[-1,1]^d} K(w)w(A)f_X(x_r+\delta_n w)dw \right| \sup_{x\in\mathbb{B}(x_r)} |F_Y(\theta_{\alpha_n}(x)|x) - F_Y(x_n'\beta_n|x)|, \end{aligned}$$

for all n large enough. By Assumptions C1 (iii) and F1 (by setting $\Delta_\alpha(x) = (1-t(x))\{\theta_{\alpha_n}(x) - x_n'\beta_n\}$ for $t(x) \in [0,1]$), we obtain $|E[Z_{1i}]| = o(\delta_n^d)$. Therefore, from the Bernstein inequality (see, e.g., van der Vaart and Wellner, 1996, pp. 102-103), there exists a positive constant c_2 such that

$$\begin{aligned} \Pr\{B_{1n}\} &\leq \Pr\left\{ \left| \sum_{i=1}^n (Z_{1i} - E[Z_{1i}]) \right| \geq M_1 n\delta_n^d - n|E[Z_{1i}]| \right\} \\ &\leq 2\exp\left(-c_2 \frac{M_1^2 n^2 \delta_n^{2d}}{c_1 n\delta_n^d + M_1 n\delta_n^d} \right) = 2\exp\left(-c_2 \frac{M_1^2 n\delta_n^d}{c_1 + M_1} \right), \end{aligned}$$

for all $M_1 > 0$ and all n large enough. From $n\delta_n^d \to \infty$ (by Assumption C1 (iii)), we obtain (A.1).

Step 3: Show (A.2). Pick any $\Delta \in \mathbb{R}^{s(A)}$. It is sufficient to show that

$$\frac{1}{n\delta_n^d} E[Q_{2n}(\Delta)] \to C > 0, \tag{A.3}$$

$$\frac{1}{n\delta_n^d} \{Q_{2n}(\Delta) - E[Q_{2n}(\Delta)]\} \overset{a.s.}{\to} 0. \tag{A.4}$$

First, (A.3) is obtained as follows:

$$\begin{aligned} \frac{1}{n\delta_n^d} E[Q_{2n}(\Delta)] &= \frac{1}{\delta_n^d} E[K_{in}(\mathbb{I}\{Y_i - X_{in}'\beta_n \leq 0\} - \mathbb{I}\{Y_i - X_{in}'\beta_n \leq X_{in}'\Delta\})(Y_i - X_{in}'\beta_n - X_{in}'\Delta)] \\ &= \frac{1}{\delta_n^d} \int_{x\in\mathbb{C}_n} K\left(\frac{x-x_r}{\delta_n}\right) \left\{ \int_0^{x_n'\Delta} F_Y(x_n'\beta_n + s|x) - F_Y(x_n'\beta_n|x)ds \right\} dx \\ &= \frac{1}{\delta_n^d} \int_{x\in\mathbb{C}_n, x_n'\Delta>0} K\left(\frac{x-x_r}{\delta_n}\right) \left\{ \int_0^{x_n'\Delta} F_Y(x_n'\beta_n + s|x)ds \right\} dx \\ &\quad - \frac{1}{\delta_n^d} \int_{x\in\mathbb{C}_n, x_n'\Delta>0} K\left(\frac{x-x_r}{\delta_n}\right) \left\{ \int_0^{x_n'\Delta} F_Y(x_n'\beta_n|x)ds \right\} dx \\ &\quad + \frac{1}{\delta_n^d} \int_{x\in\mathbb{C}_n, x_n'\Delta\leq 0} K\left(\frac{x-x_r}{\delta_n}\right) \left\{ \int_0^{x_n'\Delta} F_Y(x_n'\beta_n + s|x) - F_Y(x_n'\beta_n|x)ds \right\} dx \\ &= Q_{2n}^a - Q_{2n}^b + Q_{2n}^c, \end{aligned}$$

where the second equality follows from Knight's equality. For Q_{2n}^b, it holds

$$\begin{aligned} |Q_{2n}^b| &= \frac{1}{\delta_n^d}\int_{x\in\mathbb{C}_n, x_n'\Delta>0} K\left(\frac{x-x_r}{\delta_n}\right) x_n'\Delta F_Y(x_n'\beta_n|x)dx \\ &\le \left(\int_{w\in[-1,1]^d} K(w)|w(A)'\Delta|dw\right)\sup_{x\in\mathbb{B}(x_r)} F_Y(x_n'\beta_n|x)\to 0, \end{aligned}$$

where the convergence follows from Assumptions C1 (i) and (iii). Similarly, Q_{2n}^c satisfies

$$|Q_{2n}^c| \le \frac{1}{\delta_n^d}\int_{x\in\mathbb{C}_n, x_n'\Delta\le 0} K\left(\frac{x-x_r}{\delta_n}\right)\left\{\int_0^{x_n'\Delta} F_Y(x_n'\beta_n|x)ds\right\}dx \to 0.$$

Therefore it is sufficient for (A.3) to show that $Q_{2n}^a \to C > 0$. Observe that

$$\begin{aligned} Q_{2n}^a &= \frac{1}{\delta_n^d}\int_{x\in\mathbb{C}_n, x_n'\Delta>0} K\left(\frac{x-x_r}{\delta_n}\right)\left\{\int_0^{x_n'\Delta} F_Y(x_n'\beta_n+s|x)ds\right\}dx \\ &= \int_{w\in[-1,1]^d, w(A)'\Delta>0} K(w)\left\{\int_0^{\frac{w(A)'\Delta}{2}} F_Y(w(A)'\beta_n+s|x_r+\delta_n w)ds\right\}dw \\ &\quad + \int_{w\in[-1,1]^d, w(A)'\Delta>0} K(w)\left\{\int_{w(A)'\Delta/2}^{w(A)'\Delta} F_Y(w(A)'\beta_n+s|x_r+\delta_n w)ds\right\}dw \\ &\ge \int_{w\in[-1,1]^d, w(A)'\Delta>0} K(w)\left\{\int_{w(A)'\Delta/2}^{w(A)'\Delta} F_Y(w(A)'\beta_n+s|x_r+\delta_n w)ds\right\}dw \\ &\ge \int_{w\in[-1,1]^d, w(A)'\Delta>0} K(w)\left\{\int_{w(A)'\Delta/2}^{w(A)'\Delta} F_Y(\theta_{\alpha_n}(x_r+\delta_n w)+s/2|x_r+\delta_n w)ds\right\}dw \\ &\ge c_3\int_{w\in[-1,1]^d, w(A)'\Delta>0} K(w)\frac{w(A)'\Delta}{2}dw > 0, \end{aligned}$$

for all n large enough, where the second equality follows from the change of variables, the second inequality follows from Assumption C1 (iii), the third inequality follows from Assumption C1 (i), and the last inequality follows from Assumption K.

We now derive (A.4). Let $Z_{2i} = K_{in}[\mathbb{I}\{Y_i - X_{in}'\beta_n \le 0\} - \mathbb{I}\{Y_i - X_{in}'\beta_n \le X_{in}'\Delta\}](Y_i - X_{in}'\beta_n - X_{in}'\Delta)$ (i.e., $Q_{2n}(\Delta) = \sum_{i=1}^n Z_{2i}$) and $B_{2n} = \left\{|\sum_{i=1}^n (Z_{2i} - E[Z_{2i}])| \ge M_2 n\delta_n^d\right\}$ for $M_2 > 0$. It is sufficient to show that $\sum_{n=1}^\infty \Pr\{B_{2n}\} < \infty$ for any $M_2 > 0$ by the Borel-Cantelli lemma. Since the component $[\mathbb{I}\{Y_i - X_{in}'\beta_n \le 0\} - \mathbb{I}\{Y_i - X_{in}'\beta_n \le X_{in}'\Delta\}](Y_i - X_{in}'\beta_n - X_{in}'\Delta)$ is bounded,

$$E[Z_{2i}^2] \le c_4 E[K_{in}^2] = c_4\delta_n^d \int_{w\in[-1,1]^d} K(w)^2 f_X(x_r+\delta_n w)dw.$$

Therefore, from the Bernstein inequality, there exists a positive constant c_5 such that

$$\Pr\{B_{2n}\} \le 2\exp\left(-c_5\frac{M_2^2 n^2\delta_n^{2d}}{c_4 n\delta_n^d + M_2 n\delta_n^d}\right) = 2\exp\left(-c_5\frac{M_1^2 n\delta_n^d}{c_4+M_1}\right),$$

for all $M_2 > 0$ and all n large enough. From Assumption C1 (iii), we obtain (A.4).

A.2. **Proof of Theorem 1 (ii).** The proof is split into three steps. Let $a_n = \sqrt{\frac{n\delta_n^d\phi_n^2}{\alpha_n \log\log n}}$.

Step 1: Overall sketch. Using the reparametrization $\Delta = a_n(\beta - \beta_n)$, the normalized estimator $\tilde{\Delta}_n = a_n(\hat{\beta}_n - \beta_n)$ is obtained as the minimizer of the objective function:

$$\begin{aligned}
Q_n(\Delta) &= \sum_{i=1}^n K\left(\frac{X_i - x_r}{\delta_n}\right)[\rho_\alpha(Y_i - X_{in}'\beta_n - a_n^{-1}X_{in}'\Delta) - \rho_\alpha(Y_i - X_{in}'\beta_n)] \\
&= -a_n^{-1}\sum_{i=1}^n K_{in}[\alpha_n - \mathbb{I}\{Y_i - X_{in}'\beta_n < 0\}]X_{in}'\Delta \\
&\quad + \sum_{i=1}^n K_{in}\int_0^{a_n^{-1}X_{in}'\Delta}(\mathbb{I}\{Y_i - X_{in}'\beta_n \le s\} - \mathbb{I}\{Y_i - X_{in}'\beta_n \le 0\})ds \\
&= -a_n^{-1}Q_{1n}'\Delta + \tilde{Q}_{2n}(\Delta),
\end{aligned}$$

where Q_{1n} and $\tilde{Q}_{2n}(\Delta)$ are implicitly defined, and the second inequality follows from Knight's equality. Since $\tilde{\Delta}_n$ minimizes $Q_n(\Delta)$ with respect to Δ and $Q_n(0) = 0$, we have

$$\begin{aligned}
0 &\ge \frac{a_n}{\sqrt{n\delta_n^d\alpha_n \log\log n}}\frac{Q_n(\tilde{\Delta}_n) - Q_n(0)}{(1 + |\tilde{\Delta}_n|)^2} \\
&= -\frac{Q_{1n}'}{\sqrt{n\delta_n^d\alpha_n \log\log n}}\frac{\tilde{\Delta}_n}{(1 + |\tilde{\Delta}_n|)^2} + \frac{\phi_n}{\alpha_n \log\log n}\frac{\tilde{Q}_{2n}(\tilde{\Delta}_n)}{(1 + |\tilde{\Delta}_n|)^2},
\end{aligned} \tag{A.5}$$

for all n large enough. In Steps 2 and 3 below, we show that

$$\frac{|Q_{1n}|}{\sqrt{n\delta_n^d\alpha_n \log\log n}} = O_{as}(1), \tag{A.6}$$

$$\sup_{\Delta\in B_n(M)}\left|\frac{\phi_n}{\alpha_n \log\log n}\frac{\tilde{Q}_{2n}(\Delta)}{(1+|\Delta|)^2} - \frac{1}{2}\frac{\Delta'}{1+|\Delta|}Q\frac{\Delta}{1+|\Delta|}\right| \overset{a.s.}{\to} 0 \quad \text{for each } M, \epsilon > 0, \tag{A.7}$$

respectively, where $B_n(M) = \{\Delta \in \mathbb{R}^{s(A)} : M \le |\Delta| \le a_n\}$. Now suppose the conclusion of Theorem 1 (ii) does not hold, i.e., $|\tilde{\Delta}_n| \to \infty$ as $n \to \infty$. Then we must have $\frac{|\tilde{\Delta}_n|}{(1+|\tilde{\Delta}_n|)^2} \to 0$, and the right hand side of (A.5) must converge a.s. to a positive constant from (A.6), (A.7) with $\tilde{\Delta}_n = o_{as}(a_n)$ (by Theorem 1 (i)), and the positive definiteness of Q. This is a contradiction. In the following, we show (A.6) in Step 2 and (A.7) in Step 3.

Step 2: Show (A.6). For this step, we apply a general law of iterated logarithm by Arcones (1997, Theorem 3.3). Let $X_{in}(t) = \{(X_i - x_r)^u/(\delta_n t)^{|u|}\}_{u\in A}$ and

$$g(X_i, \delta_n t) = K\left(\frac{X_i - x_r}{\delta_n t}\right)[\alpha(\delta_n t) - \mathbb{I}\{Y_i - X_{in}(t)'\beta_n < 0\}]X_{in}(t),$$

where $\alpha : [0, 1] \to (0, 1]$ is a monotone function satisfying $\alpha(\delta_n) = \alpha_n$. In our setup, Conditions (i), (vi), and (viii) of Arcones (1997, Theorem 3.3) are trivially satisfied. Assumptions C1 and R1 guarantee Conditions (ii) and (iii) of Arcones (1997, Theorem 3.3). Also, Conditions (iv) and (v) of Arcones (1997, Theorem 3.3) are implied by the Markov inequality, boundedness of $g(X_i, \delta_n t)$, and Assumptions C1 and R1.

To check Condition (vii) of Arcones (1997, Theorem 3.3), it is sufficient to check that conditions of Arcones (1997, Theorem 2.6) hold in our setup. Since $\{g(\cdot, t) : t \in (0, 1]\}$ is a VC subgraph class, Condition (i) of Arcones (1997, Theorem 2.6) is satisfied. Assumptions C1 and R1 guarantee Condition (ii) of Arcones (1997, Theorem 2.6). Since $g(X_i, \delta_n t)$ is bounded, Assumptions

C1 and R1 imply Condition (iii) of Arcones (1997, Theorem 2.6). Also note that

$$
\begin{aligned}
& E[|g(X_i,\delta_n)|^2] \\
\leq\ & \int_{x\in\mathbb{C}_n} K\left(\frac{x-x_r}{\delta_n}\right)^2 \{\alpha_n^2 + (1-2\alpha_n)F_Y(x_n'\beta_n|x)\}|x_n|^2 f_X(x)dx \\
\leq\ & \delta_n^d \int_{w\in[-1,1]^d} K(w)^2 F_Y(w(A)'\beta_n|x_r+\delta_n w)|w(A)|^2 f_X(x_r+\delta_n w)dw \\
& +\delta_n^d\alpha_n^2 \int_{w\in[-1,1]^d} K(w)^2|w(A)|^2 f_X(x_r+\delta_n w)dw,
\end{aligned}
$$

where the first inequality follows from the law of iterated expectation and the second inequality follows from the change of variables and $(1-2\alpha_n)\leq 1$. The first term satisfies

$$
\begin{aligned}
& \delta_n^d \int_{w\in[-1,1]^d} K(w)^2 F_Y(w(A)'\beta_n|x_r+\delta_n w)|w(A)|^2 f_X(x_r+\delta_n w)dw \\
\leq\ & \delta_n^d\alpha_n \int_{w\in[-1,1]^d} K(w)^2 \left|\frac{F_Y(w(A)'\beta_n|x_r+\delta_n w)}{\alpha_n}-1\right| |w(A)|^2 f_X(x_r+\delta_n w)dw \\
& +\delta_n^d\alpha_n \int_{w\in[-1,1]^d} K(w)^2|w(A)|^2 f_X(x_r+\delta_n w)dw \\
\leq\ & c\delta_n^d\alpha_n \int_{w\in[-1,1]^d} K(w)^2|w(A)|^2 f_X(x_r+\delta_n w)dw,
\end{aligned}
$$

for all n large enough, where the last inequality follows from

$$
\begin{aligned}
& \sup_{w\in[-1,1]^d}\left|\frac{F_Y(w(A)'\beta_n|x_r+\delta_n w)}{\alpha_n}-1\right| \\
\leq\ & \sup_{w\in[-1,1]^d}\left|\frac{f_Y(\theta_{\alpha_n}(x_r+\delta_n w)+R_n(x_r+\delta_n w)|x_r+\delta_n w)}{\phi_n}\right| \frac{\phi_n}{\alpha_n}\sup_{x\in\mathbb{C}_n}|x_n'\beta_n-\theta_{\alpha_n}(x)|\to 0,
\end{aligned}
$$

where $R_n(x)\in(0, x_n'\beta_n-\theta_{\alpha_n}(x))$ and the convergence follows from Assumptions C1 (iii), F1, and R1. Combining these results,

$$
E[|g(X_i,\delta_n)|^2]\leq c\delta_n^d\alpha_n\int_{w\in[-1,1]^d} K(w)^2|w(A)|^2 f_X(x_r+\delta_n w)dw. \tag{A.8}
$$

Thus, Condition (iv) of Arcones (1997, Theorem 2.6) is satisfied. A similar argument guarantees Condition (v) of Arcones (1997, Theorem 2.6). From (A.8) and the Cauchy-Schwarz inequality, we have $\sup_{s,t\in(0,1]} E[|g(X_i,\delta_n t)-g(X_i,\delta_n s)|^2]\leq c_6\delta_n^d\alpha_n$ and Condition (vi) of Arcones (1997, Theorem 2.6) is satisfied. Boundedness of $g(X_i,\delta_n t)$ implies Condition (vii) of Arcones (1997, Theorem 2.6). Therefore, Arcones (1997, Theorem 3.3) yields the law of iterated logarithm:

$$
\frac{1}{\sqrt{n\delta_n^d\alpha_n\log\log n}}|Q_{1n}-E[Q_{1n}]| = O_{as}(1).
$$

Step 3: Show (A.7). It is sufficient to show that

$$
\frac{\phi_n}{\alpha_n\log\log n}\frac{E[\tilde{Q}_{2n}(\Delta)]}{(1+|\Delta|)^2}\to\frac{1}{2}\frac{\Delta'}{1+|\Delta|}Q\frac{\Delta}{1+|\Delta|}, \tag{A.9}
$$

$$
\frac{\phi_n}{\alpha_n\log\log n}\frac{|\tilde{Q}_{2n}(\Delta)-E[\tilde{Q}_{2n}(\Delta)]|}{(1+|\Delta|)^2}\overset{a.s.}{\to} 0. \tag{A.10}
$$

First, (A.9) is obtained as follows:

$$
\begin{aligned}
& \frac{\phi_n}{\alpha_n \log\log n} \frac{E[\tilde{Q}_{2n}(\Delta)]}{(1+|\Delta|)^2} \\
= & \frac{1}{(1+|\Delta|)^2} \frac{1}{\delta_n^d} \int_{x\in\mathbb{C}_n} K\left(\frac{x-x_r}{\delta_n}\right) \left\{ \int_0^{x_n'\Delta} s' \left(\frac{F_Y(x_n'\beta_n + a_n^{-1}s'|x) - F_Y(x_n'\beta_n|x)}{\phi_n a_n^{-1}s'} \right) ds' \right\} f_X(x)dx \\
\sim & \frac{1}{2}\frac{\Delta'}{1+|\Delta|} \left[\int_{w\in[-1,1]^d} K(w)w(A)w(A)' f_X(x_r+\delta_n w)dw \right] \frac{\Delta}{1+|\Delta|} \to \frac{1}{2}\frac{\Delta'}{1+|\Delta|} Q \frac{\Delta}{1+|\Delta|},
\end{aligned}
$$

uniformly on $\Delta \in B_n(M)$, where the equality follows from the law of iterated expectation and the change of variables $s' = a_n s$. The tilde relation follows from (by an expansion around $s=0$ and Assumptions F1 and R1)

$$
\begin{aligned}
& \sup_{x\in\mathbb{C}_n} \sup_{s\in(0,x_n'\Delta)} \left| \frac{F_Y(x_n'\beta_n + a_n^{-1}s|x) - F_Y(x_n'\beta_n|x)}{\phi_n a_n^{-1}s} - 1 \right| \\
\leq & \sup_{x\in\mathbb{C}_n} \sup_{s\in(0,x_n'\Delta)} \left| \frac{f_Y(x_n'\beta_n + a_n^{-1}s|x)}{\phi_n} - 1 \right| \to 0.
\end{aligned}
$$

Now we show (A.10). Define $Z_{3i} = \frac{1}{(1+|\Delta|)^2} K_{in} \int_0^{X_{in}'\Delta} [\mathbb{I}\{Y_i - X_{in}'\beta_n \leq a_n^{-1}s'\} - \mathbb{I}\{Y_i - X_{in}'\beta_n \leq 0\}]ds$, and note that

$$
\begin{aligned}
E[Z_{3i}^2] \leq & c_7 E\left[K_{in}^2 [\mathbb{I}\{Y_i - X_{in}'\beta_n \leq a_n^{-1}X_{in}'\Delta\} - \mathbb{I}\{Y_i - X_{in}'\beta_n \leq 0\}]^2 \frac{(X_{in}'\Delta)^2}{(1+|\Delta|)^4} \right] \\
\leq & c_8 E[K_{in}^2 |\mathbb{I}\{Y_i - X_{in}'\beta_n \leq a_n^{-1}X_{in}'\Delta\} - \mathbb{I}\{Y_i - X_{in}'\beta_n \leq 0\}|] \\
\leq & c_8 \phi_n a_n^{-1} \int_{x\in\mathbb{C}_n} K\left(\frac{x-x_r}{\delta_n}\right)^2 |x_n'\Delta| \left| \frac{F_Y(x_n'\beta_n + a_n^{-1}x_n'\Delta|x) - F_Y(x_n'\beta_n|x)}{\phi_n a_n^{-1}x_n'\Delta} \right| f_X(x)dx \\
\sim & c_8 \sqrt{\frac{\delta_n^d \alpha_n \log\log n}{n}} \int_{w\in[-1,1]^d} K(w)^2 |w(A)'\Delta| f_X(x_r+\delta_n w)dw, \qquad \text{(A.11)}
\end{aligned}
$$

uniformly on $\Delta \in B_n(M)$, where the first inequality follows from evaluating the integration, the second inequality follows from the boundedness of $\sup_{\Delta\in B_n(M)} \frac{(X_{in}'\Delta)^2}{(1+|\Delta|)^4}$ and the property of the indicator function, and the third inequality follows from the law of iterated expectation. The tilde relation follows from (by an expansion around $\Delta = 0$ and Assumptions F1 and R1)

$$
\begin{aligned}
& \sup_{x\in\mathbb{C}_n} \left| \frac{F_Y(x_n'\beta_n + a_n^{-1}x_n'\Delta|x) - F_Y(x_n'\beta_n|x)}{\phi_n a_n^{-1}x_n'\Delta} - 1 \right| \\
\leq & \sup_{x\in\mathbb{C}_n} \left| \frac{f_Y(x_n'\beta_n + a_n^{-1}x_n'\tilde{\Delta}|x)}{\phi_n} - 1 \right| \to 0,
\end{aligned}
$$

where $\tilde{\Delta} \in (0, \Delta)$. Note that the change of variables $s' = a_n s$ implies

$$
\frac{\phi_n}{\alpha_n \log\log n} \frac{\tilde{Q}_{2n}(\Delta)}{(1+|\Delta|)^2} = \frac{1}{\sqrt{n\delta_n^d \alpha_n \log\log n}} \sum_{i=1}^n Z_{3i}.
$$

Thus, the Bernstein inequality implies that there exists a positive constant c_9 such that

$$
\begin{aligned}
&\Pr\left\{\frac{\phi_n}{\alpha_n \log\log n}\frac{|\tilde{Q}_{2n}(\Delta)-E[\tilde{Q}_{2n}(\Delta)]|}{(1+|\Delta|)^2}\geq M_3\right\} \\
=\ &\Pr\left\{\left|\sum_{i=1}^{n}(Z_{3i}-E[Z_{3i}])\right|\geq M_3\sqrt{n\delta_n^d\alpha_n\log\log n}\right\}\leq \exp\left(-c_9\frac{n\delta_n^d\alpha_n\log\log n}{nE[Z_{3i}^2]+\sqrt{n\delta_n^d\alpha_n\log\log n}}\right),
\end{aligned}
$$

for all $\Delta\in B_n(M)$ and $M_3>0$ and all n large enough. Therefore, from (A.11) and $n\delta_n^d\alpha_n\log\log n=O(1)$ (by Assumption R1), the Borel-Cantelli lemma implies (A.7).

A.3. **Proof of Theorem 1 (iii).** It is sufficient to derive the asymptotic distribution of the normalized estimator $\bar{\Delta}_n=b_n(\hat{\beta}_n-\beta_n)$ with $b_n=\sqrt{\frac{n\delta_n^d\phi_n^2}{\alpha_n}}$. Note that $\bar{\Delta}_n$ minimizes the following objective function with respect to Δ:

$$
\begin{aligned}
Q_n(\Delta) &= \sum_{i=1}^{n}K\left(\frac{X_i-x_r}{\delta_n}\right)\left[\rho_\alpha(Y_i-X_{in}'\beta_n-b_n^{-1}X_{in}'\Delta)-\rho_\alpha(Y_i-X_{in}'\beta_n)\right] \\
&= -b_n^{-1}\sum_{i=1}^{n}K_{in}[\alpha_n-\mathbb{I}\{Y_i-X_{in}'\beta_n<0\}]X_{in}'\Delta \\
&\quad +\sum_{i=1}^{n}K_{in}\int_0^{b_n^{-1}X_{in}'\Delta}\left(\mathbb{I}\{Y_i-X_{in}'\beta_n\leq s\}-\mathbb{I}\{Y_i-X_{in}'\beta_n\leq 0\}\right)ds \\
&= -b_n^{-1}Q_{1n}'\Delta+\bar{Q}_{2n}(\Delta),
\end{aligned}
$$

where Q_{1n} and $\bar{Q}_{2n}(\Delta)$ are implicitly defined, and the second inequality follows from Knight's equality. We first consider Q_{1n}. A similar argument to (A.8) yields

$$
\frac{|E[Q_{1n}]|}{\sqrt{n\delta_n^d\alpha_n}}\sim\sqrt{\frac{n\delta_n^d\phi_n^2}{\alpha_n}}\sup_{x\in\mathbb{C}_n}|\theta_{\alpha_n}(x)-x_n'\beta_n|\int_{w\in[-1,1]^d}K(w)|w(A)|f_X(x_r+\delta_n w)dw\to 0, \quad \text{(A.12)}
$$

where the convergence follows from the additional assumption for this part. Thus, we have

$$
\begin{aligned}
Var\left(\frac{1}{\sqrt{n\delta_n^d\alpha_n}}Q_{1n}\right) &= \frac{1}{n\delta_n^d\alpha_n}E[Q_{1n}Q_{1n}']+o(1) \\
= &\int_{w\in[-1,1]^d}K(w)^2w(A)w(A)'f_X(x_r+\delta_n w)dw \\
&+\frac{\phi_n}{\alpha_n}\sup_{x\in\mathbb{C}_n}|x_n'\beta_n-\theta_{\alpha_n}(x)|\int_{w\in[-1,1]^d}K(w)^2w(A)w(A)'f_X(x_r+\delta_n w)dw+o(1) \\
\to\ &Q_1,
\end{aligned}
$$

where the first equality follows from (A.12), the second equality follows from similar arguments in (A.12) and the change of variables, and the convergence follows from Assumption R1. Thus, the Lindeberg-Feller central limit theorem implies

$$
-\frac{1}{\sqrt{n\delta_n^d\alpha_n}}Q_{1n}\xrightarrow{d}Q_{1\infty}\sim N(0,Q_1).
$$

We now consider $\bar{Q}_{2n}(\Delta)$. By a similar argument to show (A.9), we obtain

$$\frac{b_n}{\sqrt{n\delta_n^d \alpha_n}} \bar{Q}_{2n}(\Delta) \xrightarrow{p} \frac{1}{2}\Delta' Q \Delta,$$

for each $\Delta \in \mathbb{R}^{s(A)}$. Therefore, the weak marginal limit of $\frac{b_n}{\sqrt{n\delta_n^d \alpha_n}} Q_n(\Delta)$ is $Q_\infty(\Delta) = Q_{1\infty}'\Delta + \frac{1}{2}\Delta' Q\Delta$ for each $\Delta \in \mathbb{R}^{s(A)}$. Since Q is positive-definite (by Assumption K), the marginal limit $Q_\infty(\Delta)$ is uniquely minimized at $\Delta_\infty = Q^{-1}Q_{1\infty} \sim N(0, Q^{-1}Q_1 Q^{-1})$. Therefore, the convexity lemma of Knight (1999) yields the conclusion.

A.4. **Proof of Theorem 2.** Let $\tilde{x}_{n,r} = \{(x - x_{n,r})^u / \delta_n^{|u|}\}_{u \in A}$, $\tilde{X}_{in,r} = \{(X_i - x_{n,r})^u / \delta_n^{|u|}\}_{u \in A}$, and $K_{in,r} = K\left(\frac{X_i - x_{n,r}}{\delta_n}\right)$. Taylor expansions of $D^u \theta_{\alpha_n}(x)$ around $\{x_{n,r}\}_{r=1}^{J_n^d}$ and Assumption R1 imply

$$\sup_{x \in \mathbb{C}} \left| \widetilde{D^u \theta}_{\alpha_n}(x) - D^u \theta_{\alpha_n}(x) \right| \leq \max_{1 \leq r \leq J_n^d} \sup_{x \in \mathbb{C}_{n,r}} |D^u \{\tilde{x}_{n,r}(\hat{\beta}_{n,r} - \beta_{n,r})\}| + O\left(\sqrt{\frac{\alpha_n \log n}{n\delta_n^d \phi_n^2}} \right).$$

Let $\dot{\Delta}_{n,r} = d_n(\hat{\beta}_{n,r} - \beta_{n,r})$ with $d_n = \sqrt{\frac{n\delta_n^d \phi_n^2}{\alpha_n \log n}}$ for $r = 1, \ldots, J_n^d$. From the boundedness of $\max_{1 \leq r \leq J_n^d} \sup_{x \in \mathbb{C}_{n,r}} |\tilde{x}_{n,r}|$, it is sufficient to show that $\max_{1 \leq r \leq J_n^d} |\dot{\Delta}_{n,r}| = O_{as}(1)$.

Note that $\dot{\Delta}_{n,r}$ minimizes the following objective function with respect to Δ:

$$\begin{aligned}
Q_{n,r}(\Delta) &= \sum_{i=1}^n K\left(\frac{X_i - x_{n,r}}{\delta_n}\right) \left[\rho_\alpha(Y_i - \tilde{X}_{in,r}'\beta_{n,r} - d_n^{-1}\tilde{X}_{in,r}'\Delta) - \rho_\alpha(Y_i - \tilde{X}_{in,r}'\beta_{n,r})\right] \\
&= -d_n^{-1} \sum_{i=1}^n K_{in,r}[\alpha_n - \mathbb{I}\{Y_i - \tilde{X}_{in,r}'\beta_{n,r} < 0\}]\tilde{X}_{in,r}'\Delta \\
&\quad + \sum_{i=1}^n K_{in,r} \int_0^{d_n^{-1}\tilde{X}_{in,r}'\Delta} \left(\mathbb{I}\{Y_i - \tilde{X}_{in,r}'\beta_{n,r} \leq s\} - \mathbb{I}\{Y_i - \tilde{X}_{in,r}'\beta_{n,r} \leq 0\} \right) ds \\
&= -d_n^{-1} Q_{1n,r}'\Delta + \dot{Q}_{2n,r}(\Delta),
\end{aligned}$$

where Q_{1n} and $\dot{Q}_{2n}(\Delta)$ are implicitly defined, and the second inequality follows from Knight's equality. Observe that there exist positive constants $c, c_1 > 0$ such that

$$\begin{aligned}
\Pr\{|\dot{\Delta}_{n,r}| \geq M\} &= \Pr\left\{ 0 \geq \frac{d_n\{Q_{n,r}(\dot{\Delta}_{n,r}) - Q_{n,r}(0)\}}{\sqrt{n\delta_n^d \alpha_n \log n}(1 + |\dot{\Delta}_{n,r}|)^2},\ |\dot{\Delta}_{n,r}| \geq M \right\} \\
&\leq \Pr\left\{ \frac{|Q_{1n,r}|}{\sqrt{n\delta_n^d \alpha_n \log n}} \geq \frac{d_n \dot{Q}_{2n,r}(\dot{\Delta}_{n,r})}{\sqrt{n\delta_n^d \alpha_n \log n}(1 + |\dot{\Delta}_{n,r}|)^2}(1 + M) \right\} \\
&\leq \Pr\left\{ |Q_{1n,r}| \geq c\sqrt{n\delta_n^d \alpha_n \log n} \right\} \leq \exp(-c_1 \log n), \qquad \text{(A.13)}
\end{aligned}$$

for all $r = 1, \ldots, J_n^d$, $M > 0$, and n large enough, where the equality follows from the fact that $\dot{\Delta}_{n,r}$ minimizes $Q_{n,r}(\Delta)$, the first inequality follows the set inclusion relation, the second inequality follows from $\sup_{\Delta \in B_n(M)} \left| \frac{d_n \dot{Q}_{2n,r}(\Delta)}{\sqrt{n\delta_n^d \alpha_n \log n}(1+|\Delta|)^2} - \frac{1}{2}\frac{\Delta'}{1+|\Delta|} Q \frac{\Delta}{1+|\Delta|} \right| \xrightarrow{a.s.} 0$ (by a similar argument to derive (A.7)), and the last inequality follows from the Bernstein inequality with a similar argument to derive (A.6). From (A.13), there exists an appropriately large positive constant M

such that

$$\sum_{n=1}^{\infty} \Pr\left\{\max_{1\leq r\leq J_n^d} |\dot{\Delta}_{n,r}| \geq M\right\} \leq \sum_{n=1}^{\infty}\sum_{r=1}^{J_n^d} \Pr\{|\dot{\Delta}_{n,r}| \geq M\} \leq \sum_{n=1}^{\infty} J_n^d \exp(-c_2 \log n) < \infty$$

Therefore, the conclusion follows from the Borel-Cantelli lemma.

References

[1] Arcones, M. A. (1997) The law of the iterated logarithm for a triangular array of empirical process, *Electronic Journal of Probability*, 2, 1-39.

[2] Bosq, D. (1998) *Nonparametric Statistics for Stochastic Processes*, Springer.

[3] Chaudhuri, P. (1991a) Nonparametric estimates of regression quantiles and their local Bahadur representation, *Annals of Statistics*, 19, 760-777.

[4] Chaudhuri, P. (1991b) Global nonparametric estimation of conditional quantile functions and their derivatives, *Journal of Multivariate Analysis*, 39, 246-269.

[5] Chernozhukov, V. (1998) Nonparametric extreme regression quantiles, Working paper.

[6] Chernozhukov, V. (2005) Extremal quantile regression, *Annals of Statistics*, 33, 806-839.

[7] Fan, J. and I. Gijbels (1996) *Local Polynomial Modelling and Its Applications*, Chapman & Hall.

[8] Feigin, P. D. and S. I. Resnick (1994) Limit distributions for linear programming time series estimators, *Stochastic Process and Its Applications*, 51, 135-165.

[9] Knight, K. (1999) Epi-convergence and stochastic equisemicontinuity, Working paper.

[10] Knight, K. (2001) Limiting distributions of linear programming estimators, *Extremes*, 4, 87-103.

[11] Koenker, R. (2005) *Quantile Regression*, Cambridge University Press.

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

[13] Newey, W. K. and D. L. McFadden (1994) Large sample estimation and hypothesis testing, in: R. F. Engle and D. L. McFadden, eds., *Handbook of Econometrics*, Vol. IV, ch. 36, Elsevier, Amsterdam.

[14] Portnoy, S. and J. Jurečková (1999) On extreme regression quantiles, *Extremes*, 2, 227-243.

[15] Smith, R. L. (1994) Nonregular regression, *Biometrika*, 81, 173-183.

[16] van der Vaart, A. W. and J. A. Wellner (1996) *Weak Convergence and Empirical Processes*, Springer-Verlag, New York.

Eller College of Management, University of Arizona.

Email address: ichimura@email.arizona.edu

Department of Economics, London School of Economics, Houghton Street, London, WC2A 2AE, UK.

Email address: t.otsu@lse.ac.uk

Department of Economics, Yale University, 37 Hillhouse Avenue, New Haven, CT 06511, USA.

Email address: joseph.altonji@yale.edu