

SFB 649 Discussion Paper 2008-027

The Stochastic Fluctuation of the Quantile Regression Curve

Wolfgang Härdle*
Song Song*



* Humboldt-Universität zu Berlin, Germany

This research was supported by the Deutsche
Forschungsgemeinschaft through the SFB 649 "Economic Risk".

<http://sfb649.wiwi.hu-berlin.de>
ISSN 1860-5664

SFB 649, Humboldt-Universität zu Berlin
Spandauer Straße 1, D-10178 Berlin



SFB 649 ECONOMIC RISK BERLIN

Confidence Bands in Quantile Regression*

Wolfgang K. Härdle [†]; Song Song [‡]

Abstract

Let $(X_1, Y_1), \dots, (X_n, Y_n)$ be i.i.d. rvs and let $l(x)$ be the unknown p -quantile regression curve of Y conditional on X . A quantile-smoother $l_n(x)$ is a localised, nonlinear estimator of $l(x)$. The strong uniform consistency rate is established under general conditions. In many applications it is necessary to know the stochastic fluctuation of the process $\{l_n(x) - l(x)\}$. Using strong approximations of the empirical process and extreme value theory, we consider the asymptotic maximal deviation $\sup_{0 \leq x \leq 1} |l_n(x) - l(x)|$. The derived result helps in the construction of a uniform confidence band for the quantile curve $l(x)$. This confidence band can be applied as a econometric model check. An economic application considers the relation between age and earnings in labour market by means of parametric model specification tests, which presents a new framework to describe trends in the entire wage distribution in a parsimonious way.

Keywords: Quantile Regression; Consistency Rate; Confidence Band; Check Function; Kernel Smoothing; Nonparametric Fitting

JEL classification: C00; C14; J01; J31

1 Introduction

In standard regression function estimation, most investigations are concerned with the conditional mean regression. However, new insights about the un-

*The financial support from the Deutsche Forschungsgemeinschaft via SFB 649 “Ökonomisches Risiko”, Humboldt-Universität zu Berlin is gratefully acknowledged. We thank the editor and two referees for concrete suggestions on improving the manuscript and restructuring the paper. Their valuable comments and suggestions are gratefully acknowledged.

[†]Professor at Humboldt-Universität zu Berlin and Director of CASE - Center for Applied Statistics and Economics, Humboldt-Universität zu Berlin, Spandauer Straße 1, 10178 Berlin, Germany.

[‡]Corresponding author. Research associate at the Institute for Statistics and Econometrics of Humboldt-Universität zu Berlin, Spandauer Straße 1, 10178 Berlin, Germany. Email: songsong@cms.hu-berlin.de.

derlying structures can be gained by considering other aspects of the conditional distribution. The quantile curves are key aspects of inference in various economic problems and are of great interest in practice. These describe the conditional behaviour of a response variable (e.g. wage of workers) given the value of an explanatory variable (e.g. education level, experience, occupation of workers), and investigate changes in both tails of the distribution, other than just the mean.

When examining labour markets, economists are concerned with whether discrimination exists, for example for different genders, nationalities, union status and so on. To study this question, we need to separate out other effects first, e.g. age, education, etc. The crucial relation between age and earnings or salaries belongs to the most carefully studied subjects in labor economics. The fundamental work in mean regression can be found in Murphy and Welch (1990). Quantile regression estimates could provide more accurate measures. Koenker and Hallock (2001) present a basket of important economic applications, including quantile Engel curves and claim that “quantile regression is gradually developing into a comprehensive strategy for completing the regression prediction”. Besides this, it is also well-known that a quantile regression model (e.g. the conditional median curve) is more robust to outliers, especially for fat-tailed distributions. For symmetric conditional distributions the quantile regression generates the nonparametric mean regression analysis since the $p = 0.5$ (median) quantile curve coincides with the mean regression.

As first introduced by Koenker and Bassett (1978), one may assume a parametric model for the p -quantile curve and estimate parameters by the interior point method discussed by Koenker and Park (1996) and Portnoy and Koenker (1997). Similarly, we can also adopt nonparametric methods to estimate conditional quantiles. The first one, a more direct approach using a check function such as a robustified local linear smoother, is provided by Fan et al. (1994) and further extended by Yu and Jones (1997, 1998). An alternative procedure is first to estimate the conditional distribution function using the double-kernel local linear technique of Fan et al. (1996) and then to invert the conditional distribution estimator to produce an estimator of a conditional quantile by Yu and Jones (1997, 1998). Beside these, Hall et al. (1999) proposed a weighted version of the Nadaraya-Watson estimator, which was further studied by Cai (2002). Recently Jeong and Härdle (2008) have developed the conditional quantile causality test. More generally, for an M -regression function which involves quantile regression as a special case, the uniform Bahadur representation and application to the additive model is studied by Kong et al. (2008). An interesting question for the parametric fitting, especially from labour economicists, would be how well these models

fit the data, when compared with the nonparametric estimation method.

Let $(X_1, Y_1), (X_2, Y_2), \dots, (X_n, Y_n)$ be a sequence of independent identically distributed bivariate random variables with joint pdf $f(x, y)$, joint cdf $F(x, y)$, conditional pdf $f(y|x), f(x|y)$, conditional cdf $F(y|x), F(x|y)$ for Y given X and X given Y respectively, and marginal pdf $f_X(x)$ for X , $f_Y(y)$ for Y where $x \in J$, and J is a possibly infinite interval in \mathbb{R}^d and $y \in \mathbb{R}$. In general, X may be a multivariate covariate, although here we restrict attention to the univariate case and $J = [0, 1]$ for convenience. Let $l(x)$ denote the p -quantile curve, i.e. $l(x) = F_{Y|x}^{-1}(p)$.

Under a ‘‘check function’’, the quantile regression curve $l(x)$ can be viewed as the minimiser of $L(\theta) \stackrel{\text{def}}{=} \mathbf{E}\{\rho_p(y - \theta)|X = x\}$ (w.r.t. θ) with $\rho_p(u) = pu\mathbf{1}\{u \in (0, \infty)\} - (1 - p)u\mathbf{1}\{u \in (-\infty, 0)\}$ which was originally motivated by an exercise in Ferguson (1967)[p.51] in the literature.

A kernel-based p -quantile curve estimator $l_n(x)$ can naturally be constructed by minimising:

$$L_n(\theta) = n^{-1} \sum_{i=1}^n \rho_p(Y_i - \theta) K_h(x - X_i) \quad (1)$$

with respect to $\theta \in I$ where I is a possibly infinite, or possibly degenerate, interval in \mathbb{R} , and $K_h(u) = h^{-1}K(u/h)$ is a kernel with bandwidth h . The numerical solution of (1) may be found iteratively as in Lejeune and Sarda (1988) and Yu et al. (2003).

In light of the concepts of M -estimation as in Huber (1981), if we define $\psi(u)$ as:

$$\begin{aligned} \psi_p(u) &= p\mathbf{1}\{u \in (0, \infty)\} - (1 - p)\mathbf{1}\{u \in (-\infty, 0)\} \\ &= p - \mathbf{1}\{u \in (-\infty, 0)\}, \end{aligned}$$

$l_n(x)$ and $l(x)$ can be treated as a zero (w.r.t. θ) of the function:

$$\tilde{H}_n(\theta, x) \stackrel{\text{def}}{=} n^{-1} \sum_{i=1}^n K_h(x - X_i) \psi(Y_i - \theta) \quad (2)$$

$$\tilde{H}(\theta, x) \stackrel{\text{def}}{=} \int_{\mathbb{R}} f(x, y) \psi(y - \theta) dy \quad (3)$$

correspondingly.

To show the uniform consistency of the quantile smoother, we shall reduce the problem of strong convergence of $l_n(x) - l(x)$, uniformly in x , to an application of the strong convergence of $\tilde{H}_n(\theta, x)$ to $\tilde{H}(\theta, x)$, uniformly in x

and θ , as given by Theorem 2.2 in Härdle et al. (1988). It is shown that under general conditions almost surely (a.s.)

$$\sup_{x \in J} |l_n(x) - l(x)| \leq B^* \max\{(nh/(\log n))^{-1/2}, h^{\tilde{\alpha}}\}, \quad \text{as } n \rightarrow \infty.$$

where B^* and $\tilde{\alpha}$ are parameters defined more precisely in Section 2.

Please note that without assuming K has compact support (as we do here) under similar assumptions Franke and Mwita (2003) get:

$$\begin{aligned} l_n(x) &= \hat{F}_{Y|x}^{-1}(p) \\ \hat{F}(y|x) &= \frac{\sum_{i=1}^n K_h(x - X_i) \mathbf{1}(Y_i < y)}{\sum_{i=1}^n K_h(x - X_i)} \\ \sup_{x \in J} |l_n(x) - l(x)| &\leq B^{**} \{(nh/(s_n \log n))^{-1/2} + h^2\}, \quad \text{as } n \rightarrow \infty. \end{aligned}$$

for α -mixing data where B^{**} is some constant and $s_n, n \geq 1$ is an increasing sequence of positive integers satisfying $1 \leq s_n \leq \frac{n}{2}$ and some other criteria. Thus $\{nh/(\log n)\}^{-1/2} \leq \{nh/(s_n \log n)\}^{-1/2}$.

By employing similar methods as those developed in Härdle (1989) it is shown in this paper that

$$\begin{aligned} &P \left((2\delta \log n)^{1/2} \left[\sup_{x \in J} r(x) |\{l_n(x) - l(x)\}| / \lambda(K)^{1/2} - d_n \right] < z \right) \\ &\longrightarrow \exp\{-2 \exp(-z)\}, \quad \text{as } n \rightarrow \infty. \end{aligned} \tag{4}$$

from the asymptotic Gumbel distribution where $r(x)$, δ , $\lambda(K)$, d_n are suitable scaling parameters. The asymptotic result (4) therefore allows the construction of (asymptotic) uniform confidence bands for $l(x)$ based on specifications of the stochastic fluctuation of $l_n(x)$. The strong approximation with Brownian bridge techniques that we use in this paper is available only for the approximation of the 2-dimensional empirical process. The extension to the multivariate covariable can be done by partial linear modelling which deserves further research.

The plan of the paper is as follows. In Section 2, the stochastic fluctuation of the process $\{l_n(x) - l(x)\}$ and the uniform confidence band are presented through the equivalence of several stochastic processes, with a strong uniform consistency rate of $\{l_n(x) - l(x)\}$ also shown. In Section 3, in a small Monte Carlo study we investigate the behaviour of $l_n(x)$ when the data is generated by fat-tailed conditional distributions of $(Y|X = x)$. In Section 4, an application considers a wage-earning relation in the labour market. All proofs are sketched in Section 5.

2 Results

The following assumptions will be convenient. To make x and X clearly distinguishable, we replace x by t sometimes, but they are essentially the same.

(A1) The kernel $K(\cdot)$ is positive, symmetric, has compact support $[-A, A]$ and is Lipschitz continuously differentiable with bounded derivatives;

(A2) $(nh)^{-1/2}(\log n)^{3/2} \rightarrow 0$, $(n \log n)^{1/2}h^{5/2} \rightarrow 0$, $(nh^3)^{-1}(\log n)^2 \leq M$, M a constant;

(A3) $h^{-3}(\log n) \int_{|y|>a_n} f_Y(y)dy = \mathcal{O}(1)$, $f_Y(y)$ the marginal density of Y , $\{a_n\}_{n=1}^\infty$ a sequence of constants tending to infinity as $n \rightarrow \infty$;

(A4) $\inf_{t \in J} |q(t)| \geq q_0 > 0$, where $q(t) = \partial \mathbb{E}\{\psi(Y - \theta)|t\} / \partial \theta|_{\theta=l(t)} \cdot f_X(t) = f\{l(t)|t\}f_X(t)$;

(A5) the quantile function $l(t)$ is Lipschitz twice continuously differentiable, for all $t \in J$.

(A6) $0 < m_1 \leq f_X(t) \leq M_1 < \infty$, $t \in J$; the conditional densities $f(\cdot|y)$, $y \in \mathbb{R}$, are uniform local Lipschitz continuous of order $\tilde{\alpha}$ (uLL- $\tilde{\alpha}$) on J , uniformly in $y \in \mathbb{R}$, with $0 < \tilde{\alpha} \leq 1$.

Define also

$$\begin{aligned} \sigma^2(t) &= \mathbb{E}[\psi^2\{Y - l(t)\}|t] = p(1 - p) \\ H_n(t) &= (nh)^{-1} \sum_{i=1}^n K\{(t - X_i)/h\} \psi\{Y_i - l(t)\} \\ D_n(t) &= \partial(nh)^{-1} \sum_{i=1}^n K\{(t - X_i)/h\} \psi\{Y_i - \theta\} / \partial \theta|_{\theta=l(t)} \end{aligned}$$

and assume that $\sigma^2(t)$ and $f_X(t)$ are differentiable.

Assumption (A1) on the compact support of the kernel could possibly be relaxed by introducing a cutoff technique as in Csörgö and Hall (1982) for density estimators. Assumption (A2) has purely technical reasons: to keep the bias at a lower rate than the variance and to ensure the vanishing of some non-linear remainder terms. Assumption (A3) appears in a somewhat modified form also in Johnston (1982). Assumptions (A5, A6) are common assumptions in robust estimation as in Huber (1981), Härdle et al. (1988) that are satisfied by exponential, and generalised hyperbolic distributions.

For the uniform strong consistency rate of $l_n(x) - l(x)$, we apply the result of Härdle et al. (1988) by taking $\beta(y) = \psi(y - \theta)$, $y \in \mathbb{R}$, for $\theta \in I = \mathbb{R}$, $q_1 = q_2 = -1$, $\gamma_1(y) = \max\{0, -\psi(y - \theta)\}$, $\gamma_2(y) = \min\{0, -\psi(y - \theta)\}$ and $\lambda = \infty$ to satisfy the representations for the parameters there. Thus from Theorem 2.2 and Remark 2.3(v) there we immediately have the following lemma.

LEMMA 2.1 Let $\tilde{H}_n(\theta, x)$ and $\tilde{H}(\theta, x)$ be given by (2) and (3). Under assumption (A6) and $(nh/\log n)^{-1/2} \rightarrow \infty$ through (A2), for some constant A^* not depending on n , we have a.s. as $n \rightarrow \infty$

$$\sup_{\theta \in I} \sup_{x \in J} |\tilde{H}_n(\theta, x) - \tilde{H}(\theta, x)| \leq A^* \max\{(nh/\log n)^{-1/2}, h^{\tilde{\alpha}}\} \quad (5)$$

For our result on $l_n(\cdot)$, we shall also require

$$\inf_{x \in J} \left| \int \psi\{y - l(x) + \varepsilon\} dF(y|x) \right| \geq \tilde{q}|\varepsilon|, \quad \text{for } |\varepsilon| \leq \delta_1, \quad (6)$$

where δ_1 and \tilde{q} are some positive constants, see also Härdle and Luckhaus (1984). This assumption is satisfied if there exists a constant \tilde{q} such that $f(l(x)|x) > \tilde{q}/p$, $x \in J$.

THEOREM 2.1 Under the conditions of Lemma 2.1 and also assuming (6), we have a.s. as $n \rightarrow \infty$

$$\sup_{x \in J} |l_n(x) - l(x)| \leq B^* \max\{(nh/\log n)^{-1/2}, h^{\tilde{\alpha}}\} \quad (7)$$

with $B^* = A^*/m_1\tilde{q}$ not depending on n and m_1 a lower bound of $f_X(t)$. If additionally $\tilde{\alpha} \geq \{\log(\sqrt{\log n}) - \log(\sqrt{nh})\}/\log h$, it can be further simplified to

$$\sup_{x \in J} |l_n(x) - l(x)| \leq B^* \{(nh/\log n)^{-1/2}\}.$$

THEOREM 2.2 Let $h = n^{-\delta}$, $\frac{1}{5} < \delta < \frac{1}{3}$, $\lambda(K) = \int_{-A}^A K^2(u) du$ and

$$d_n = (2\delta \log n)^{1/2} + (2\delta \log n)^{-1/2} [\log\{c_1(K)/\pi^{1/2}\} + \frac{1}{2}\{\log \delta + \log \log n\}],$$

$$\text{if } c_1(K) = \{K^2(A) + K^2(-A)\}/\{2\lambda(K)\} > 0$$

$$d_n = (2\delta \log n)^{1/2} + (2\delta \log n)^{-1/2} \log\{c_2(K)/2\pi\}$$

$$\text{otherwise with } c_2(K) = \int_{-A}^A \{K'(u)\}^2 du / \{2\lambda(K)\}.$$

Then (4) holds with

$$r(x) = (nh)^{1/2} f\{l(x)|x\} \{f_X(x)/p(1-p)\}^{1/2}.$$

This theorem can be used to construct uniform confidence intervals for the regression function as stated in the following corollary.

COROLLARY 2.1 *Under the assumptions of the theorem above, an approximate $(1 - \alpha) \times 100\%$ confidence band over $[0, 1]$ is*

$$l_n(t) \pm (nh)^{-1/2} \{p(1-p)\lambda(K)/\hat{f}_X(t)\}^{1/2} \hat{f}^{-1}\{l(t)|t\} \{d_n + c(\alpha)(2\delta \log n)^{-1/2}\},$$

where $c(\alpha) = \log 2 - \log |\log(1 - \alpha)|$ and $\hat{f}_X(t)$, $\hat{f}\{l(t)|t\}$ are consistent estimates for $f_X(t)$, $f\{l(t)|t\}$.

In the literature, according to Fan et al. (1994, 1996), Yu and Jones (1997, 1998), Hall et al. (1999), Cai (2002) and others, asymptotic normality at interior points for various nonparametric smoothers, e.g. local constant, local linear, reweighted NW methods, etc. has been shown:

$$\sqrt{nh}\{l_n(t) - l(t)\} \sim N(0, \tau^2(t))$$

with $\tau^2(t) = \lambda(K)p(1-p)/[f_X(t)f^2\{l(t)|t\}]$. Please note that the bias term vanishes here as we adjust h . With $\tau(t)$ introduced, we can further write Corollary 2.1 as:

$$l_n(t) \pm (nh)^{-1/2} \{d_n + c(\alpha)(2\delta \log n)^{-1/2}\} \hat{\tau}(t).$$

Through minimising the approximation of AMSE (asymptotic mean square error), the optimal bandwidth h_p can be computed. In practice, the rule-of-thumb for h_p is given by Yu and Jones (1998):

1. Use ready-made and sophisticated methods to select optimal bandwidth h_{mean} from conditional mean regression, e.g. Ruppert et al. (1995)
2. $h_p = [p(1-p)/\varphi^2\{\Phi^{-1}(p)\}]^{1/5} \cdot h_{\text{mean}}$
with φ , Φ as the pdf and cdf of a standard normal distribution

Obviously the further p lies from 0.5, the more smoothing is necessary.

The proof is essentially based on a linearisation argument after a Taylor series expansion. The leading linear term will then be approximated in a similar way as in Johnston (1982), Bickel and Rosenblatt (1973). The main idea behind the proof is a strong approximation of the empirical process of $\{(X_i, Y_i)_{i=1}^n\}$ by a sequence of Brownian bridges as proved by Tusnady (1977).

As $l_n(t)$ is the zero (w.r.t. θ) of $\tilde{H}_n(\theta, t)$, it follows by applying 2nd-order Taylor expansions to $\tilde{H}_n(\theta, t)$ around $l(t)$ that

$$l_n(t) - l(t) = \{H_n(t) - \mathbf{E} H_n(t)\}/q(t) + R_n(t) \quad (8)$$

where $\{H_n(t) - \mathbf{E} H_n(t)\}/q(t)$ is the leading linear term and

$$R_n(t) = H_n(t)\{q(t) - D_n(t)\}/\{D_n(t) \cdot q(t)\} + \mathbf{E} H_n(t)/q(t) + \frac{1}{2}\{l_n(t) - l(t)\}^2 \cdot \{D_n(t)\}^{-1} \quad (9)$$

$$\cdot (nh)^{-1} \sum_{i=1}^n K\{(x - X_i)/h\} \psi''\{Y_i - l(t) + r_n(t)\}, \quad (10)$$

$$|r_n(t)| < |l_n(t) - l(t)|.$$

is the remainder term. In Section 5 it is shown (Lemma 5.1) that $\|R_n\| = \sup_{t \in J} |R_n(t)| = \mathcal{O}_p\{(nh \log n)^{-1/2}\}$.

Furthermore, the rescaled linear part

$$Y_n(t) = (nh)^{1/2} \{\sigma^2(t) f_X(t)\}^{-1/2} \{H_n(t) - \mathbf{E} H_n(t)\}$$

is approximated by a sequence of Gaussian processes, leading finally to the Gaussian process

$$Y_{5,n}(t) = h^{-1/2} \int K\{(t - x)/h\} dW(x). \quad (11)$$

Drawing upon the result of Bickel and Rosenblatt (1973), we finally obtain asymptotically the Gumbel distribution.

We also need the Rosenblatt (1952) transformation,

$$T(x, y) = \{F_{X|y}(x|y), F_Y(y)\},$$

which transforms (X_i, Y_i) into $T(X_i, Y_i) = (X'_i, Y'_i)$ mutually independent uniform rv's. In the event that x is a d -dimension covariate, the transformation becomes:

$$T(x_1, x_2, \dots, x_d, y) = \{F_{X_1|y}(x_1|y), F_{X_2|y}(x_2|x_1, y), \dots, F_{X_k|x_{d-1}, \dots, x_1, y}(x_k|x_{d-1}, \dots, x_1, y), F_Y(y)\}. \quad (12)$$

With the aid of this transformation, Theorem 1 of Tusnady (1977) may be applied to obtain the following lemma.

LEMMA 2.2 *On a suitable probability space a sequence of Brownian bridges B_n exists that*

$$\sup_{x \in J, y \in \mathbb{R}} |Z_n(x, y) - B_n\{T(x, y)\}| = \mathcal{O}\{n^{-1/2}(\log n)^2\} \quad a.s.,$$

where $Z_n(x, y) = n^{1/2}\{F_n(x, y) - F(x, y)\}$ denotes the empirical process of $\{(X_i, Y_i)\}_{i=1}^n$.

For $d > 2$, it is still an open problem which deserves further research.

Before we define the different approximating processes, let us first rewrite (11) as a stochastic integral w.r.t. the empirical process $Z_n(x, y)$,

$$Y_n(t) = \{hg'(t)\}^{-1/2} \iint K\{(t-x)/h\}\psi\{y-l(t)\}dZ_n(x, y),$$

$$g'(t) = \sigma^2(t)f_X(t).$$

The approximating processes are now:

$$Y_{0,n}(t) = \{hg(t)\}^{-1/2} \iint_{\Gamma_n} K\{(t-x)/h\}\psi\{y-l(t)\}dZ_n(x, y) \quad (13)$$

where $\Gamma_n = \{|y| \leq a_n\}$, $g(t) = \mathbf{E}[\psi^2\{y-l(t)\} \cdot \mathbf{1}(|y| \leq a_n) | X = t] \cdot f_X(t)$

$$Y_{1,n}(t) = \{hg(t)\}^{-1/2} \iint_{\Gamma_n} K\{(t-x)/h\}\psi\{y-l(t)\}dB_n\{T(x, y)\} \quad (14)$$

$\{B_n\}$ being the sequence of Brownian bridges from Lemma 2.2.

$$Y_{2,n}(t) = \{hg(t)\}^{-1/2} \iint_{\Gamma_n} K\{(t-x)/h\}\psi\{y-l(t)\}dW_n\{T(x, y)\} \quad (15)$$

$\{W_n\}$ being the sequence of Wiener processes satisfying

$$B_n(x', y') = W_n(x', y') - x'y'W_n(1, 1)$$

$$Y_{3,n}(t) = \{hg(t)\}^{-1/2} \iint_{\Gamma_n} K\{(t-x)/h\}\psi\{y-l(x)\}dW_n\{T(x, y)\} \quad (16)$$

$$Y_{4,n}(t) = \{hg(t)\}^{-1/2} \int g(x)^{1/2} K\{(t-x)/h\}dW(x) \quad (17)$$

$$Y_{5,n}(t) = h^{-1/2} \int K\{(t-x)/h\}dW(x) \quad (18)$$

$\{W(\cdot)\}$ being the Wiener process.

Lemmas 5.2 to 5.7 ensure that all these processes have the same limit distributions. The result then follows from

LEMMA 2.3 (Theorem 3.1 in Bickel and Rosenblatt (1973)) Let $d_n, \lambda(K), \delta$ as in Theorem 2.2. Let

$$Y_{5,n}(t) = h^{-1/2} \int K\{(t-x)/h\} dW(x).$$

Then, as $n \rightarrow \infty$, the supremum of $Y_{5,n}(t)$ has a Gumbel distribution.

$$P \left\{ (2\delta \log n)^{1/2} \left[\sup_{t \in J} |Y_{5,n}(t)| / \{\lambda(K)\}^{1/2} - d_n \right] < z \right\} \rightarrow \exp\{-2 \exp(-z)\}.$$

3 A Monte Carlo Study

We generate bivariate data $\{(X_i, Y_i)\}_{i=1}^n, n = 500$ with joint pdf:

$$\begin{aligned} f(x, y) &= g(y - \sqrt{x + 2.5}) \mathbf{1}(x \in [-2.5, 2.5]) \\ g(u) &= \frac{9}{10} \varphi(u) + \frac{1}{90} \varphi(u/9). \end{aligned} \quad (19)$$

The p -quantile curve $l(x)$ can be obtained from a zero (w.r.t. θ) of:

$$9\Phi(\theta) + \Phi(\theta/9) = 10p,$$

with Φ as the cdf of a standard normal distribution. Solving it numerically gives the 0.5-quantile curve $l(x) = \sqrt{x + 2.5}$, and the 0.9-quantile curve $l(x) = 1.5296 + \sqrt{x + 2.5}$. We use the quartic kernel:

$$\begin{aligned} K(u) &= \frac{15}{16} (1 - u^2)^2, & |u| \leq 1, \\ &= 0, & |u| > 1. \end{aligned}$$

In Fig. 1 the raw data, together with the 0.5-quantile curve, are displayed. The random variables generated with probability $\frac{1}{10}$ from the fat-tailed pdf $\frac{1}{9} \varphi(u/9)$, see (19), are marked as squares whereas the standard normal rv's are shown as stars. We then compute both the Nadaraya-Watson estimator $m_n^*(x)$ and the 0.5-quantile smoother $l_n(x)$. The bandwidth is set to 1.25 which is equivalent to 0.25 after rescaling x to $[0, 1]$ and fulfills the requirements of Theorem 2.2.

In Fig. 1 $l(x), m_n^*(x)$ and $l_n(x)$ are shown as a dotted line, dashed-dot line, and solid line respectively. At first sight $m_n^*(x)$ has clearly more variation and has the expected sensitivity to the fat-tails of $f(x, y)$. A closer look reveals that $m_n^*(x)$ for $x \approx 0$ apparently even leaves the 0.5-quantile curve. It may be surprising that this happens at $x \approx 0$ where no outlier is placed, but a closer look at Fig. 1 shows that the large negative data values at both $x \approx -0.1$ and

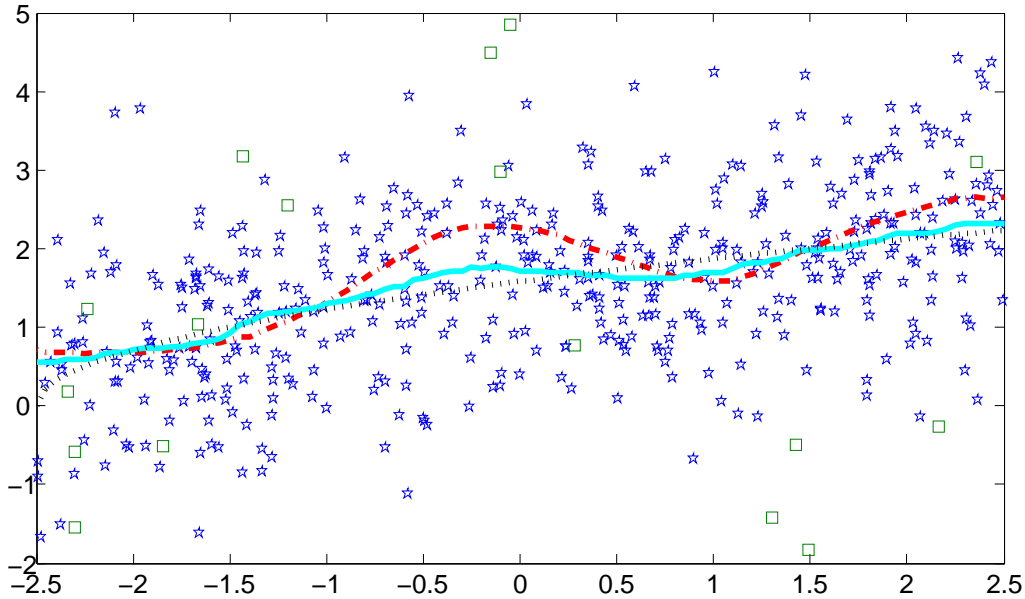


Figure 1: The 0.5-quantile curve, the Nadaraya-Watson estimator $m_n^*(x)$, and the 0.5-quantile smoother $l_n(x)$.

$x \approx 0.25$ cause the problem. This data value is inside the window ($h = 1.10$) and therefore distorts $m_n^*(x)$ for $x \approx 0$. The quantile-smoother $l_n(x)$ (solid line) is unaffected and stays fairly close to the 0.5-quantile curve. Similar results can be obtained in Fig. 2 corresponding to the 0.9 quantile ($h = 1.25$) with the 95% confidence band.

4 Application

Recently there has been great interest in finding out how the financial returns of a job depend on the age of the employee. We use the Current Population Survey (CPS) data from 2005 for the following group: male aged 25 – 59, full-time employed, and college graduate containing 16,731 observations, for the age-earning estimation. As is usual for wage data, a log transformation to hourly real wages (unit: US dollar) is carried out first. In the CPS all ages (25 ~ 59) are reported as integers. We rescaled them into $[0, 1]$ by dividing 40 by bandwidth 0.059 for nonparametric quantile-smoothers. This is equivalent to set bandwidth 2 for the original age data.

In Fig. 3 the original observations are displayed as small stars. The local 0.5 and 0.9 quantiles at the integer points of age are shown as dashed lines,

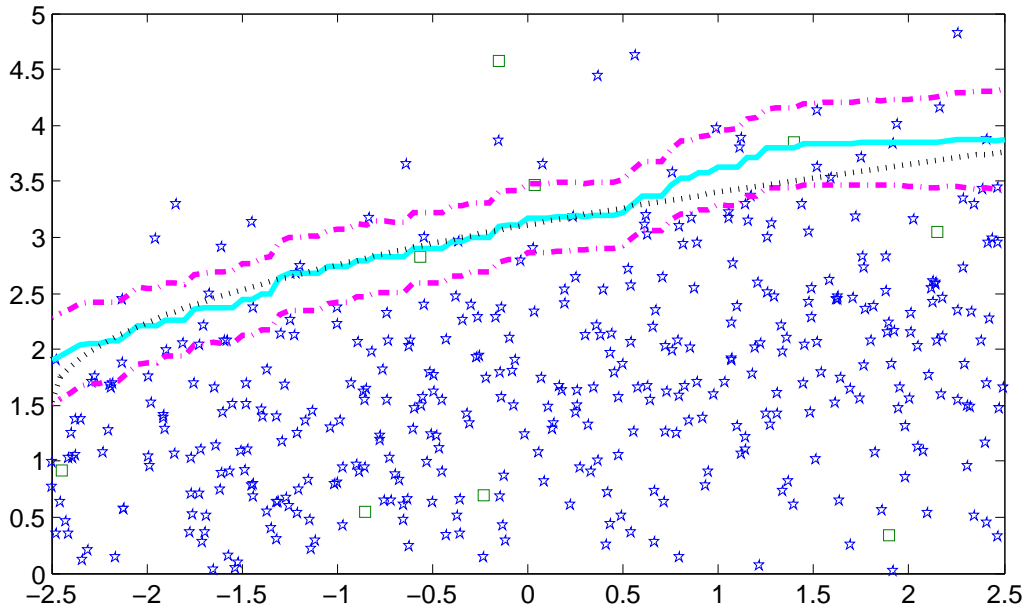


Figure 2: The 0.9-quantile curve, the 0.9-quantile smoother and 95% confidence band.

whereas the corresponding nonparametric quantile-smoothers are displayed as solid lines with corresponding 95% uniform confidence bands shown as dashed-dot lines. A closer look reveals a quadratic relation between age and logged hourly real wages. If we use several popular parametric methods to estimate the 0.5 and 0.9 conditional quantiles, e.g. quadratic, quartic and set of dummies (a dummy variable for each 5-year age group) models as in Fig. 4. With the help of the 95% uniform confidence bands, we can do the parametric model specification test. At the 5% significance level, we could not reject any model. However, when the confidence level further decreases and the uniform confidence bands get narrower, “set of dummies” parametric model will be the first one to be rejected. At the 10% significance level, the set of dummies (for age groups) model is rejected while the other two are not. As the quadratic model performs quite similar by the quartic one, for simplicity, it is suggested in practice to measure the log(wage)-earning relation which coincides with Murphy and Welch (1990) in mean regression.

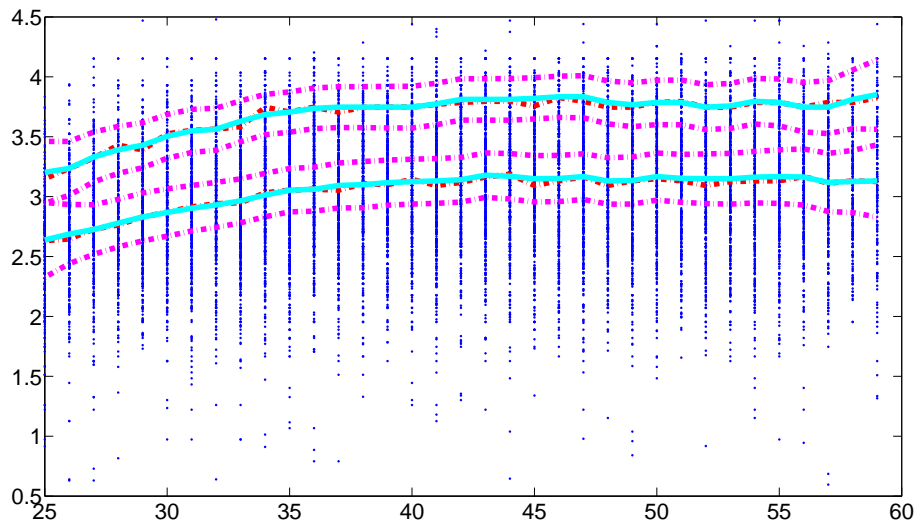


Figure 3: The original observations, local quantiles, 0.5, 0.9-quantile smoothers and corresponding 95% confidence bands.

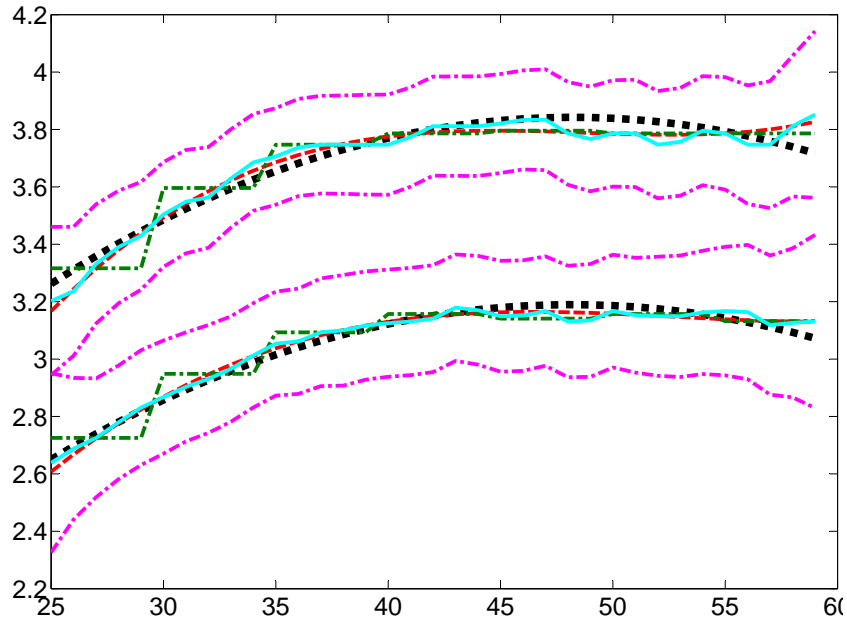


Figure 4: Quadratic, quartic, set of dummies (for age groups) estimates, 0.5, 0.9-quantile smoothers and their corresponding 95% confidence bands.

5 Appendix

Proof of Theorem 2.1. By the definition of $l_n(x)$ as a zero of (2), we have, for $\varepsilon > 0$,

$$\text{if } l_n(x) > l(x) + \varepsilon, \text{ and then } \tilde{H}_n\{l(x) + \varepsilon, x\} > 0. \quad (20)$$

Now

$$\tilde{H}_n\{l(x) + \varepsilon, x\} \leq \tilde{H}\{l(x) + \varepsilon, x\} + \sup_{\theta \in I} |\tilde{H}_n(\theta, x) - \tilde{H}(\theta, x)|. \quad (21)$$

Also, by the identity $\tilde{H}\{l(x), x\} = 0$, the function $\tilde{H}\{l(x) + \varepsilon, x\}$ is not positive and has a magnitude $\geq m_1 \tilde{q} \varepsilon$ by assumption (A6) and (6), for $0 < \varepsilon < \delta_1$. That is, for $0 < \varepsilon < \delta_1$,

$$\tilde{H}\{l(x) + \varepsilon, x\} \leq -m_1 \tilde{q} \varepsilon. \quad (22)$$

Combining (20), (21) and (22), we have, for $0 < \varepsilon < \delta_1$:

$$\text{if } l_n(x) > l(x) + \varepsilon, \text{ and then } \sup_{\theta \in I} \sup_{x \in J} |\tilde{H}_n(\theta, x) - \tilde{H}(\theta, x)| > m_1 \tilde{q} \varepsilon.$$

With a similar inequality proved for the case $l_n(x) < l(x) + \varepsilon$, we obtain, for $0 < \varepsilon < \delta_1$:

$$\text{if } \sup_{x \in J} |l_n(x) - l(x)| > \varepsilon, \text{ and then } \sup_{\theta \in I} \sup_{x \in J} |\tilde{H}_n(\theta, x) - \tilde{H}(\theta, x)| > m_1 \tilde{q} \varepsilon. \quad (23)$$

It readily follows that (23), and (5) imply (7). \blacksquare

Below we first show that $\|R_n\|_\infty = \sup_{t \in J} |R_n(t)|$ vanishes asymptotically faster than the rate $(nh \log n)^{-1/2}$; for simplicity we will just use $\|\cdot\|$ to indicate the sup-norm.

LEMMA 5.1 *For the remainder term $R_n(t)$ defined in (9) we have*

$$\|R_n\| = o_p\{(nh \log n)^{-1/2}\}. \quad (24)$$

Proof. First we have by the positivity of the kernel K ,

$$\begin{aligned} \|R_n\| &\leq \left[\inf_{0 \leq t \leq 1} \{|D_n(t)| \cdot q(t)\} \right]^{-1} \{ \|H_n\| \cdot \|q - D_n\| + \|D_n\| \cdot \|E H_n\| \} \\ &\quad + C_1 \cdot \|l_n - l\|^2 \cdot \left\{ \inf_{0 \leq t \leq 1} |D_n(t)| \right\}^{-1} \cdot \|f_n\|_\infty, \end{aligned}$$

where $f_n(x) = (nh)^{-1} \sum_{i=1}^n K\{(x - X_i)/h\}$.

The desired result (5.1) will then follow if we prove

$$\|H_n\| = \mathcal{O}_p\{(nh)^{-1/2}(\log n)^{1/2}\} \quad (25)$$

$$\|q - D_n\| = \mathcal{O}_p\{(nh)^{-1/4}(\log n)^{-1/2}\} \quad (26)$$

$$\|\mathbf{E} H_n\| = \mathcal{O}(h^2) \quad (27)$$

$$\|l_n - l\|^2 = \mathcal{O}_p\{(nh)^{-1/2}(\log n)^{-1/2}\} \quad (28)$$

Since (27) follows from the well-known bias calculation

$$\mathbf{E} H_n(t) = h^{-1} \int K\{(t-u)/h\} \mathbf{E}[\psi\{y-l(t)\}|X=u] f_X(u) du = \mathcal{O}(h^2),$$

where $\mathcal{O}(h^2)$ is independent of t in Parzen (1962), we have from assumption (A2) that $\|\mathbf{E} H_n\| = \mathcal{O}_p\{(nh)^{-1/2}(\log n)^{-1/2}\}$.

According to Lemma A.3 in Franke and Mwita (2003),

$$\sup_{t \in J} |H_n(t) - \mathbf{E} H_n(t)| = \mathcal{O}\{(nh)^{-1/2}(\log n)^{1/2}\}.$$

and the following inequality

$$\begin{aligned} \|H_n\| &\leq \|H_n - \mathbf{E} H_n\| + \|\mathbf{E} H_n\|. \\ &= \mathcal{O}\{(nh)^{-1/2}(\log n)^{1/2}\} + \mathcal{O}_p\{(nh)^{-1/2}(\log n)^{-1/2}\} \\ &= \mathcal{O}\{(nh)^{-1/2}(\log n)^{1/2}\} \end{aligned}$$

Statement (25) thus is obtained.

Statement (26) follows in the same way as (25) using assumption (A2) and the Lipschitz continuity properties of K , ψ' , l .

According to the uniform consistency of $l_n(t) - l(t)$ shown before, we have

$$\|l_n - l\| = \mathcal{O}_p\{(nh)^{-1/2}(\log n)^{1/2}\}$$

which implies (28).

Now the assertion of the lemma follows, since by tightness of $D_n(t)$, $\inf_{0 \leq t \leq 1} |D_n(t)| \geq q_0$ *a.s.* and thus

$$\|R_n\| = \mathcal{O}_p\{(nh \log n)^{-1/2}\}(1 + \|f_n\|).$$

Finally, by Theorem 3.1 of Bickel and Rosenblatt (1973), $\|f_n\| = \mathcal{O}_p(1)$; thus the desired result $\|R_n\| = \mathcal{O}_p\{(nh \log n)^{-1/2}\}$ follows. ■

We now begin with the subsequent approximations of the processes $Y_{0,n}$ to $Y_{5,n}$.

LEMMA 5.2

$$\|Y_{0,n} - Y_{1,n}\| = \mathcal{O}\{(nh)^{-1/2}(\log n)^2\} \quad a.s.$$

Proof. Let t be fixed and put $L(y) = \psi\{y - l(t)\}$ still depending on t . Using integration by parts, we obtain

$$\begin{aligned} & \iint_{\Gamma_n} L(y)K\{(t-x)/h\}dZ_n(x,y) \\ &= \int_{u=-A}^A \int_{y=-a_n}^{a_n} L(y)K(u)dZ_n(t-h \cdot u, y) \\ &= - \int_{-A}^A \int_{-a_n}^{a_n} Z_n(t-h \cdot u, y)d\{L(y)K(u)\} \\ & \quad + L(a_n)(a_n) \int_{-A}^A Z_n(t-h \cdot u, a_n)dK(u) \\ & \quad - L(-a_n)(-a_n) \int_{-A}^A Z_n(t-h \cdot u, -a_n)dK(u) \\ & \quad + K(A) \left\{ \int_{-a_n}^{a_n} Z_n(t-h \cdot A, y)dL(y) \right. \\ & \quad \left. + L(a_n)(a_n)Z_{n_a}(t-h \cdot A, a_n) - L(-a_n)(-a_n)Z_n(t-h \cdot A, -a_n) \right\} \\ & \quad - K(-A) \left\{ \int_{-a_n}^{a_n} Z_n(t+h \cdot A, y)dL(y) + L(a_n)(a_n)Z_n(t+h \cdot A, a_n) \right. \\ & \quad \left. - L(-a_n)(-a_n)Z_n(t+h \cdot A, -a_n) \right\}. \end{aligned}$$

If we apply the same operation to $Y_{1,n}$ with $B_n\{T(x, y)\}$ instead of $Z_n(x, y)$ and use Lemma 2.2, we finally obtain

$$\sup_{0 \leq t \leq 1} h^{1/2}g(t)^{1/2}|Y_{0,n}(t) - Y_{1,n}(t)| = \mathcal{O}\{n^{-1/2}(\log n)^2\} \quad a.s..$$

■

LEMMA 5.3 $\|Y_{1,n} - Y_{2,n}\| = \mathcal{O}_p(h^{1/2})$.

Proof. Note that the Jacobian of $T(x, y)$ is $f(x, y)$. Hence

$$\begin{aligned} & Y_{1,n}(t) - Y_{2,n}(t) \\ &= \left| \{g(t)h\}^{-1/2} \iint_{\Gamma_n} \psi\{y - l(t)\}K\{(t-x)/h\}f(x, y)dx dy \right| \cdot |W_n(1, 1)|. \end{aligned}$$

It follows that

$$h^{-1/2}\|Y_{1,n} - Y_{2,n}\| \leq |W_n(1, 1)| \cdot \|g^{-1/2}\| \cdot \sup_{0 \leq t \leq 1} h^{-1} \iint_{\Gamma_n} |\psi\{y - l(t)\} K\{(t - x)/h\}| f(x, y) dx dy.$$

Since $\|g^{-1/2}\|$ is bounded by assumption, we have

$$h^{-1/2}\|Y_{1,n} - Y_{2,n}\| \leq |W_n(1, 1)| \cdot C_4 \cdot h^{-1} \int K\{(t - x)/h\} dx = \mathcal{O}_p(1).$$

■

LEMMA 5.4 $\|Y_{2,n} - Y_{3,n}\| = \mathcal{O}_p(h^{1/2})$.

Proof. The difference $|Y_{2,n}(t) - Y_{3,n}(t)|$ may be written as

$$\left| \{g(t)h\}^{-1/2} \iint_{\Gamma_n} [\psi\{y - l(t)\} - \psi\{y - l(x)\}] K\{(t - x)/h\} dW_n\{T(x, y)\} \right|.$$

If we use the fact that l is uniformly continuous, this is smaller than

$$h^{-1/2}|g(t)|^{-1/2} \cdot \mathcal{O}_p(h)$$

and the lemma thus follows. ■

LEMMA 5.5 $\|Y_{4,n} - Y_{5,n}\| = \mathcal{O}_p(h^{1/2})$.

Proof.

$$\begin{aligned} |Y_{4,n}(t) - Y_{5,n}(t)| &= h^{-1/2} \left| \int \left[\left\{ \frac{g(x)}{g(t)} \right\}^{1/2} - 1 \right] K\{(t - x)/h\} dW(x) \right| \\ &\leq h^{-1/2} \left| \int_{-A}^A W(t - hu) \frac{\partial}{\partial u} \left[\left\{ \frac{g(t - hu)}{g(t)} \right\}^{1/2} - 1 \right] K(u) du \right| \\ &\quad + h^{-1/2} \left| K(A)W(t - hA) \left[\left\{ \frac{g(t - Ah)}{g(t)} \right\}^{1/2} - 1 \right] \right| \\ &\quad + h^{-1/2} \left| K(-A)W(t + hA) \left[\left\{ \frac{g(t + Ah)}{g(t)} \right\}^{1/2} - 1 \right] \right| \\ &= S_{1,n}(t) + S_{2,n}(t) + S_{3,n}(t), \quad \text{say.} \end{aligned}$$

The second term can be estimated by

$$h^{-1/2}\|S_{2,n}\| \leq K(A) \cdot \sup_{0 \leq t \leq 1} |W(t - Ah)| \cdot \sup_{0 \leq t \leq 1} h^{-1} \left| \left[\left\{ \frac{g(t - Ah)}{g(t)} \right\}^{1/2} - 1 \right] \right|;$$

by the mean value theorem it follows that

$$h^{-1/2}\|S_{2,n}\| = \mathcal{O}_p(1).$$

The first term $S_{1,n}$ is estimated as

$$\begin{aligned} h^{-1/2}S_{1,n}(t) &= \left| h^{-1} \int_{-A}^A W(t-uh)K'(u) \left[\left\{ \frac{g(t-uh)}{g(t)} \right\}^{1/2} - 1 \right] du \right. \\ &\quad \left. - \frac{1}{2} \int_{-A}^A W(t-uh)K(u) \left\{ \frac{g(t-uh)}{g(t)} \right\}^{1/2} \left\{ \frac{g'(t-uh)}{g(t)} \right\} du \right| \\ &= |T_{1,n}(t) - T_{2,n}(t)|, \quad \text{say}; \end{aligned}$$

$\|T_{2,n}\| \leq C_5 \cdot \int_{-A}^A |W(t-hu)| du = \mathcal{O}_p(1)$ by assumption on $g(t) = \sigma^2(t) \cdot f_X(t)$. To estimate $T_{1,n}$ we again use the mean value theorem to conclude that

$$\sup_{0 \leq t \leq 1} h^{-1} \left| \left\{ \frac{g(t-uh)}{g(t)} \right\}^{1/2} - 1 \right| < C_6 \cdot |u|;$$

hence

$$\|T_{1,n}\| \leq C_6 \cdot \sup_{0 \leq t \leq 1} \int_{-A}^A |W(t-hu)|K'(u)u du = \mathcal{O}_p(1).$$

Since $S_{3,n}(t)$ is estimated as $S_{2,n}(t)$, we finally obtain the desired result. \blacksquare

The next lemma shows that the truncation introduced through $\{a_n\}$ does not affect the limiting distribution.

LEMMA 5.6 $\|Y_n - Y_{0,n}\| = \mathcal{O}_p\{(\log n)^{-1/2}\}$.

Proof. We shall only show that $g'(t)^{-1/2}h^{-1/2} \iint_{\mathbb{R}-\Gamma_n} \psi\{y-l(t)\}K\{(t-x)/h\}dZ_n(x,y)$ fulfills the lemma. The replacement of $g'(t)$ by $g(t)$ may be proved as in Lemma A.4 of Johnston (1982). The quantity above is less than $h^{-1/2}\|g^{-1/2}\| \cdot \|\iint_{\{|y|>a_n\}} \psi\{y-l(\cdot)\}K\{(\cdot-x)/h\}dZ(x,y)\|$. It remains to be shown that the last factor tends to zero at a rate $\mathcal{O}_p\{(\log n)^{-1/2}\}$. We show first that

$$\begin{aligned} V_n(t) &= (\log n)^{1/2}h^{-1/2} \iint_{\{|y|>a_n\}} \psi\{y-l(t)\}K\{(t-x)/h\}dZ_n(x,y) \\ &\xrightarrow{p} 0 \quad \text{for all } t \end{aligned}$$

and then we show tightness of $V_n(t)$, the result then follows:

$$\begin{aligned} V_n(t) &= (\log n)^{1/2}(nh)^{-1/2} \sum_{i=1}^n [\psi\{Y_i - l(t)\} \mathbf{1}(|Y_i| > a_n) K\{(t - X_i)/h\} \\ &\quad - \mathbf{E} \psi\{Y_i - l(t)\} \mathbf{1}(|Y_i| > a_n) K\{(t - X_i)/h\}] \\ &= \sum_{i=1}^n X_{n,t}(t), \end{aligned}$$

where $\{X_{n,t}(t)\}_{i=1}^n$ are i.i.d. for each n with $\mathbf{E} X_{n,t}(t) = 0$ for all $t \in [0, 1]$. We then have

$$\begin{aligned} \mathbf{E} X_{n,t}^2(t) &\leq (\log n)(nh)^{-1} \mathbf{E} \psi^2\{Y_i - l(t)\} \mathbf{1}(|Y_i| > a_n) K^2\{(t - X_i)/h\} \\ &\leq \sup_{-A \leq u \leq A} K^2(u) \cdot (\log n)(nh)^{-1} \mathbf{E} \psi^2\{Y_i - l(t)\} \mathbf{1}(|Y_i| > a_n); \end{aligned}$$

hence

$$\begin{aligned} \text{Var}\{V_n(t)\} &= \mathbf{E} \left\{ \sum_{i=1}^n X_{n,t}(t) \right\}^2 = n \cdot \mathbf{E} X_{n,t}^2(t) \\ &\leq \sup_{-A \leq u \leq A} K^2(u) h^{-1} (\log n) \int_{\{|y| > a_n\}} f_y(y) dy \cdot M_\psi. \end{aligned}$$

where M_ψ denotes an upper bound for ψ^2 . This term tends to zero by assumption (A3). Thus by Markov's inequality we conclude that

$$V_n(t) \xrightarrow{p} 0 \quad \text{for all } t \in [0, 1].$$

To prove tightness of $\{V_n(t)\}$ we refer again to the following moment condition as stated in Lemma 5.1:

$$\begin{aligned} \mathbf{E}\{|V_n(t) - V_n(t_1)| \cdot |V_n(t_2) - V_n(t)|\} &\leq C' \cdot (t_2 - t_1)^2 \\ C' &\text{ denoting a constant, } t \in [t_1, t_2]. \end{aligned}$$

We again estimate the left-hand side by Schwarz's inequality and estimate each factor separately,

$$\begin{aligned} \mathbf{E}\{V_n(t) - V_n(t_1)\}^2 &= (\log n)(nh)^{-1} \mathbf{E} \left[\sum_{i=1}^n \Psi_n(t, t_1, X_i, Y_i) \cdot \mathbf{1}(|Y_i| > a_n) \right. \\ &\quad \left. - \mathbf{E}\{\Psi_n(t, t_1, X_i, Y_i) \cdot \mathbf{1}(|Y_i| > a_n)\} \right]^2, \end{aligned}$$

where $\Psi_n(t, t_1, X_i, Y_i) = \psi\{Y_i - l(t)\}K\{(t - X_i)/h\} - \psi\{Y_i - l(t_1)\}K\{(t_1 - X_i)/h\}$. Since ψ, K are Lipschitz continuous except at one point and the expectation is taken afterwards, it follows that

$$\begin{aligned} & [\mathbf{E}\{V_n(t) - V_n(t_1)\}^2]^{1/2} \\ & \leq C_7 \cdot (\log n)^{1/2} h^{-3/2} |t - t_1| \cdot \left\{ \int_{\{|y| > a_n\}} f_y(y) dy \right\}^{1/2}. \end{aligned}$$

If we apply the same estimation to $V_n(t_2) - V_n(t_1)$ we finally have

$$\begin{aligned} & \mathbf{E}\{|V_n(t) - V_n(t_1)| \cdot |V_n(t_2) - V_n(t)|\} \\ & \leq C_7^2 (\log n) h^{-3} |t - t_1| |t_2 - t| \times \int_{\{|y| > a_n\}} f_y(y) dy \\ & \leq C' \cdot |t_2 - t_1|^2 \text{ since } t \in [t_1, t_2] \text{ by (A3)}. \end{aligned}$$

■

LEMMA 5.7 *Let $\lambda(K) = \int K^2(u) du$ and let $\{d_n\}$ be as in the theorem. Then*

$$(2\delta \log n)^{1/2} [\|Y_{3,n}\| / \{\lambda(K)\}^{1/2} - d_n]$$

has the same asymptotic distribution as

$$(2\delta \log n)^{1/2} [\|Y_{4,n}\| / \{\lambda(K)\}^{1/2} - d_n].$$

Proof. $Y_{3,n}(t)$ is a Gaussian process with

$$\mathbf{E} Y_{3,n}(t) = 0$$

and covariance function

$$\begin{aligned} r_3(t_1, t_2) &= \mathbf{E} Y_{3,n}(t_1) Y_{3,n}(t_2) \\ &= \{g(t_1)g(t_2)\}^{-1/2} h^{-1} \iint_{\Gamma_n} \psi^2\{y - l(x)\} K\{(t_1 - x)/h\} \\ &\quad \times K\{(t_2 - x)/h\} f(x, y) dx dy \\ &= \{g(t_1)g(t_2)\}^{-1/2} h^{-1} \iint_{\Gamma_n} \psi^2\{y - l(x)\} f(y|x) dy K\{(t_1 - x)/h\} \\ &\quad \times K\{(t_2 - x)/h\} f_X(x) dx \\ &= \{g(t_1)g(t_2)\}^{-1/2} h^{-1} \int g(x) K\{(t_1 - x)/h\} K\{(t_2 - x)/h\} dx \\ &= r_4(t_1, t_2) \end{aligned}$$

where $r_4(t_1, t_2)$ is the covariance function of the Gaussian process $Y_{4,n}(t)$, which proves the lemma. ■

References

- Bickel, P. and Rosenblatt, M. (1973). On some global measures of the deviation of density function estimators. *Annals of Statistics*, 1:1071–1095.
- Cai, Z. W. (2002). Regression quantiles for time series. *Econometric Theory*, 18:169–192.
- Csörgö, S. and Hall, P. (1982). Upper and lower classes for triangular arrays. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 61:207–222.
- Fan, J., Hu, T. C., and Troung, Y. K. (1994). Robust nonparametric function estimation. *Scandinavian Journal of Statistics*, 21:433–446.
- Fan, J., Yao, Q., and Tong, H. (1996). Estimation of conditional densities and sensitivity measures in nonlinear dynamical systems. *Biometrika*, 83:189–206.
- Ferguson, T. S. (1967). *Mathematical Statistics: A Decision Theoretic Approach*. New York: Academic Press.
- Franke, J. and Mwita, P. (2003). Nonparametric estimates for conditional quantiles of time series. *Report in Wirtschaftsmathematik 87, University of Kaiserslautern*.
- Hall, P., Wolff, R., and Yao, Q. (1999). Methods for estimating a conditional distribution function. *Journal of the American Statistical Association*, 94:154–163.
- Härdle, W. (1989). Asymptotic maximal deviation of M-smoothers. *Journal of Multivariate Analysis*, 29:163–179.
- Härdle, W., Janssen, P., and Serfling, R. (1988). Strong uniform consistency rates for estimators of conditional functionals. *Annals of Statistics*, 16:1428–1429.
- Härdle, W. and Luckhaus, S. (1984). Uniform consistency of a class of regression function estimators. *Annals of Statistics*, 12:612–623.
- Huber, P. (1981). *Robust Statistics*. Wiley, New York.
- Jeong, K. and Härdle, W. (2008). A consistent nonparametric test for causality in quantile. *SFB 649 Discussion Paper*.

- Johnston, G. (1982). Probabilities of maximal deviations of nonparametric regression function estimates. *Journal of Multivariate Analysis*, 12:402–414.
- Koenker, R. and Bassett, G. W. (1978). Regression quantiles. *Econometrica*, 46:33–50.
- Koenker, R. and Hallock, K. F. (2001). Quantile regression. *Journal of Econometric Perspectives*, 15(4):143–156.
- Koenker, R. and Park, B. J. (1996). An interior point algorithm for nonlinear quantile regression. *Journal of Econometrics*, 71:265–283.
- Kong, E., Linton, O., and Xia, Y. (2008). Uniform Bahadur representation for local polynomial estimates of M-regression and its application to the additive model. *Econometric Theory*, *accepted*.
- Lejeune, M. G. and Sarda, P. (1988). Quantile regression: A nonparametric approach. *Computational Statistics and Data Analysis*, 6:229–239.
- Murphy, K. and Welch, F. (1990). Empirical age-earnings profiles. *Journal of Labor Economics*, 8(2):202–229.
- Parzen, M. (1962). On estimation of a probability density function and mode. *Annals of Mathematical Statistics*, 32:1065–1076.
- Portnoy, S. and Koenker, R. (1997). The Gaussian hare and the Laplacian tortoise: computability of squared-error versus absolute -error estimators (with discussion). *Statistical Sciences*, 12:279–300.
- Rosenblatt, M. (1952). Remarks on a multivariate transformation. *Annals of Mathematical Statistics*, 23:470–472.
- Ruppert, D., Sheather, S. J., and Wand, M. P. (1995). An effective bandwidth selector for local least squares regression. *Journal of the American Statistical Association*, 90:1257–1270.
- Tusnady, G. (1977). A remark on the approximation of the sample distribution function in the multidimensional case. *Periodica Mathematica Hungarica*, 8:53–55.
- Yu, K. and Jones, M. C. (1997). A comparison of local constant and local linear regression quantile estimation. *Computational Statistics and Data Analysis*, 25:159–166.

- Yu, K. and Jones, M. C. (1998). Local linear quantile regression. *Journal of the American Statistical Association*, 93:228–237.
- Yu, K., Lu, Z., and Stander, J. (2003). Quantile regression: applications and current research areas. *Journal of the Royal Statistical Society: Series D*, 52:331–350.

SFB 649 Discussion Paper Series 2008

For a complete list of Discussion Papers published by the SFB 649, please visit <http://sfb649.wiwi.hu-berlin.de>.

- 001 "Testing Monotonicity of Pricing Kernels" by Yuri Golubev, Wolfgang Härdle and Roman Timonfeev, January 2008.
- 002 "Adaptive pointwise estimation in time-inhomogeneous time-series models" by Pavel Cizek, Wolfgang Härdle and Vladimir Spokoiny, January 2008.
- 003 "The Bayesian Additive Classification Tree Applied to Credit Risk Modelling" by Junni L. Zhang and Wolfgang Härdle, January 2008.
- 004 "Independent Component Analysis Via Copula Techniques" by Ray-Bing Chen, Meihui Guo, Wolfgang Härdle and Shih-Feng Huang, January 2008.
- 005 "The Default Risk of Firms Examined with Smooth Support Vector Machines" by Wolfgang Härdle, Yuh-Jye Lee, Dorothea Schäfer and Yi-Ren Yeh, January 2008.
- 006 "Value-at-Risk and Expected Shortfall when there is long range dependence" by Wolfgang Härdle and Julius Mungo, January 2008.
- 007 "A Consistent Nonparametric Test for Causality in Quantile" by Kiho Jeong and Wolfgang Härdle, January 2008.
- 008 "Do Legal Standards Affect Ethical Concerns of Consumers?" by Dirk Engelmann and Dorothea Kübler, January 2008.
- 009 "Recursive Portfolio Selection with Decision Trees" by Anton Andriyashin, Wolfgang Härdle and Roman Timofeev, January 2008.
- 010 "Do Public Banks have a Competitive Advantage?" by Astrid Matthey, January 2008.
- 011 "Don't aim too high: the potential costs of high aspirations" by Astrid Matthey and Nadja Dwenger, January 2008.
- 012 "Visualizing exploratory factor analysis models" by Sigbert Klinke and Cornelia Wagner, January 2008.
- 013 "House Prices and Replacement Cost: A Micro-Level Analysis" by Rainer Schulz and Axel Werwatz, January 2008.
- 014 "Support Vector Regression Based GARCH Model with Application to Forecasting Volatility of Financial Returns" by Shiyi Chen, Kiho Jeong and Wolfgang Härdle, January 2008.
- 015 "Structural Constant Conditional Correlation" by Enzo Weber, January 2008.
- 016 "Estimating Investment Equations in Imperfect Capital Markets" by Silke Hüttel, Oliver Mußhoff, Martin Odening and Nataliya Zynych, January 2008.
- 017 "Adaptive Forecasting of the EURIBOR Swap Term Structure" by Oliver Blaskowitz and Helmut Herwatz, January 2008.
- 018 "Solving, Estimating and Selecting Nonlinear Dynamic Models without the Curse of Dimensionality" by Viktor Winschel and Markus Krätzig, February 2008.
- 019 "The Accuracy of Long-term Real Estate Valuations" by Rainer Schulz, Markus Staiber, Martin Wersing and Axel Werwatz, February 2008.
- 020 "The Impact of International Outsourcing on Labour Market Dynamics in Germany" by Ronald Bachmann and Sebastian Braun, February 2008.
- 021 "Preferences for Collective versus Individualised Wage Setting" by Tito Boeri and Michael C. Burda, February 2008.

SFB 649, Spandauer Straße 1, D-10178 Berlin
<http://sfb649.wiwi.hu-berlin.de>

This research was supported by the Deutsche Forschungsgemeinschaft through the SFB 649 "Economic Risk".



- 022 "Lumpy Labor Adjustment as a Propagation Mechanism of Business Cycles" by Fang Yao, February 2008.
- 023 "Family Management, Family Ownership and Downsizing: Evidence from S&P 500 Firms" by Jörn Hendrich Block, February 2008.
- 024 "Skill Specific Unemployment with Imperfect Substitution of Skills" by Runli Xie, March 2008.
- 025 "Price Adjustment to News with Uncertain Precision" by Nikolaus Hautsch, Dieter Hess and Christoph Müller, March 2008.
- 026 "Information and Beliefs in a Repeated Normal-form Game" by Dietmar Fehr, Dorothea Kübler and David Danz, March 2008.
- 027 "The Stochastic Fluctuation of the Quantile Regression Curve" by Wolfgang Härdle and Song Song, March 2008.

SFB 649, Spandauer Straße 1, D-10178 Berlin
<http://sfb649.wiwi.hu-berlin.de>

This research was supported by the Deutsche
Forschungsgemeinschaft through the SFB 649 "Economic Risk".

