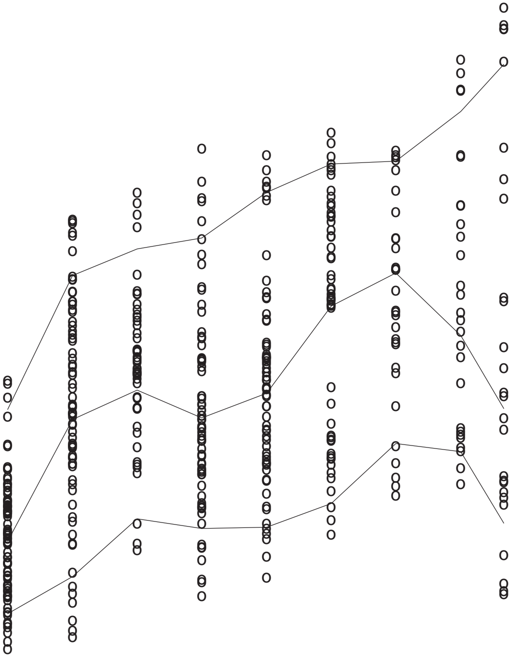
Partial Linear Quantile Regression and Bootstrap Confidence Bands

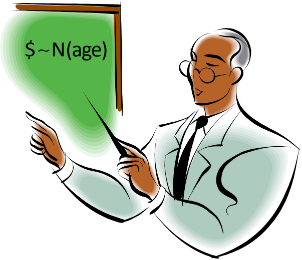
Wolfgang K. Härdle
Ya'acov Ritov
Song Song

CASE - Center for Applied Statistics and Economics
Humboldt-Universität zu Berlin

Department of Statistics
The Hebrew University of Jerusalem
http://www.case.hu-berlin.de
http://www.stat.huji.ac.il

Risky to see only part(s) of the truth!





- log(Salary) ∼ Years
- “The rich got richer and the poor got poorer!”
- Yu et al. (2003)

Quantile Regression

- QR: conditional behavior of a response Y
- Median regression $=$ mean regression (symmetric)
- "Gradually developing into a comprehensive strategy for completing the regression prediction", Koenker & Hallock (2001)

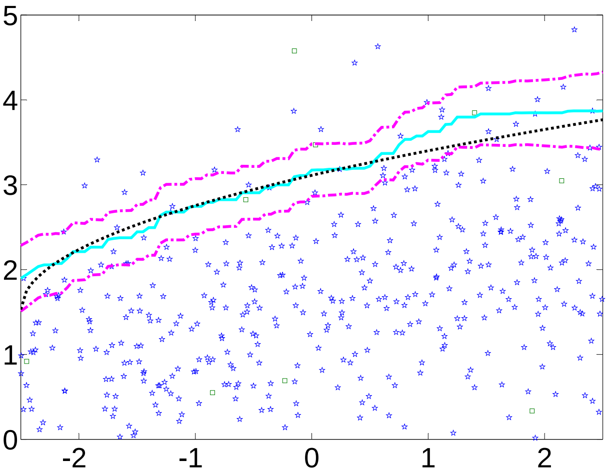


Figure 1: The 0.9-quantile curve, the 0.9-quantile smoother with $h_{0.9} = 1.25$ and 95% confidence bands. QR1

Example

- Financial Market & Econometrics
  - VaR (Value at Risk) tool to measure risk, Lauridsen (2000)
  - Detect conditional heteroscedasticity, Koenker & Bassett (1982)
- Labor Market
  - Analyse income of football players w.r.t. different ages, years, and countries, etc
  - Investigate discrimination effects, Buchinsky (1995)

$$\begin{aligned}\log(\text{Income}) &= A(\text{year, age, etc}) \\ &\quad +\beta\, B(\text{education, gender, nationality, union status, etc}) + \varepsilon\end{aligned}$$

  - Inequality analysis
  - ...

Quantile Regression

- $l(x) = F^{-1}_{Y|x}(p)$ p-quantile regression curve
- $l(x) =$ linear (parametric) form, Koenker & Bassett (1978)
- $l_h(x)$ quantile-smoother

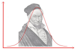
How to decide between functional forms? (global variability of the estimate, peak or valley really a feature?)

Theorem (Härdle and Song (2009))

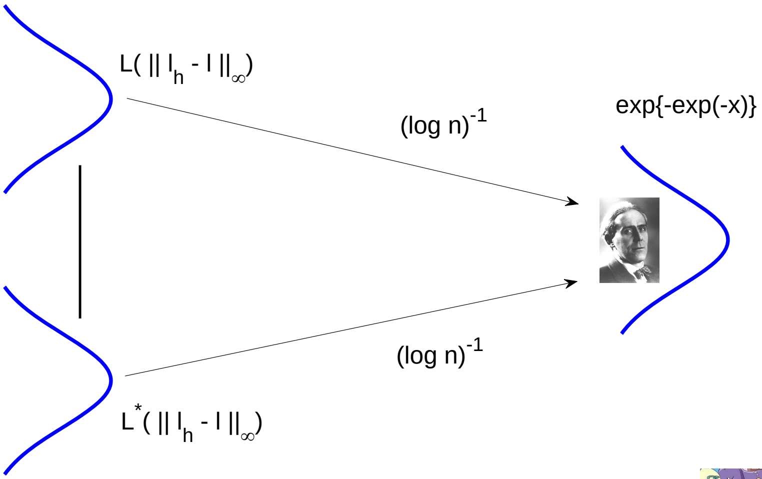
An approximate $(1-\alpha)\times 100\%$ *confidence band over* $[0,1]$ *is*

$$\begin{aligned} l_h(t) \quad \pm \quad & (nh)^{-1/2}\{p(1-p)/\hat{f}_X(t)\}^{1/2}\hat{f}^{-1}\{l(t)|t\} \\ & \times\{d_n + c(\alpha)(2\delta\log n)^{-1/2}\}\cdot\{\lambda(K)\}^{1/2}, \end{aligned} \qquad (1)$$

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

Emil Julius Gumbel on BBI: 

Challenges



Opportunities

- "Hungarian machine gun", $x \in \mathbb{R}^1$ (KMT)
  Tool to prove asymptotic bands
- Extend this to $x \in \mathbb{R}^d$ and improve band precision?
  - Hall (1991): bootstrap can beat it! (density)
  - Hahn (1995): consistency of bootstraping CDF
  - Horowitz (1998): bootstrap (pointwise) for median
  - PLM: Green & Yandell (1985), Denby (1986), Speckman (1988) and Robinson (1988)
  - Variable selection for QR: Liang and Li (2009)

Outline

1. Motivation ✓
2. Bootstrap Confidence Bands
3. Bootstrap Confidence Bands in PLMs
4. Monte Carlo Study
5. Labour Market Applications

Quantile Regression

- $\{(X_i, Y_i)\}_{i=1}^n$ i.i.d. rv's, $x \in J^* = (a, b)$ for some $0 < a < b < 1$, $y \in \mathbb{R}$
- Suppose $Y_i = l(X_i) + \varepsilon_i$, $\varepsilon_i \sim F(\cdot|X_i)$ with $F(0|X_i) = p$. Both l & F are smooth.
- Estimator $l_h(\cdot)$: the solution of

$$\frac{\sum_{i=1}^n K_h(x - X_i)\mathbf{1}\{Y_i < l_h(x)\}}{\sum_{i=1}^n K_h(x - X_i)} < p \leq \frac{\sum_{i=1}^n K_h(x - X_i)\mathbf{1}\{Y_i \leq l_h(x)\}}{\sum_{i=1}^n K_h(x - X_i)}$$

- S_n: any slowly varying function (e.g., $S_n^2 = S_n$ is valid...).

- Local rate of convergence of l_h
  $\delta_n = h^2 + (nh)^{-1/2} = \mathcal{O}(n^{-2/5})$ with $h_n = \mathcal{O}(n^{-1/5})$
- Auxiliary estimate l_g with larger bandwidth $g_n = h_n n^{\zeta}$ (ζ: $4/45$)
- $\hat{F}(\cdot|X_i) = \frac{\sum_{j=1}^{n} K_h(X_j - X_i)\mathbf{1}\{Y_j - l_h(X_i) \le \cdot\}}{\sum_{j=1}^{n} K_h(X_j - X_i)}$

Check Function

$$\rho_p(u) = pu\mathbf{1}\{u \in (0,\infty)\} - (1-p)u\mathbf{1}\{u \in (-\infty,0)\}$$

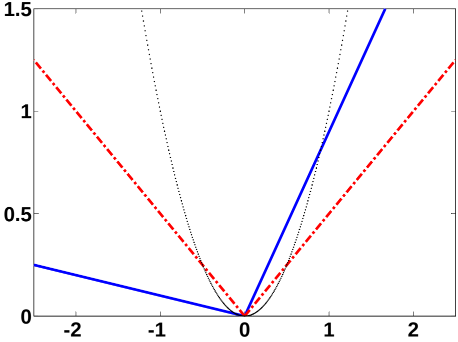


Figure 2: Check function for p=0.9, p=0.5 and weight function in conditional mean regression

The Quantile Curve

$$\rho_p(u) = pu\mathbf{1}\{u \in (0,\infty)\} - (1-p)u\mathbf{1}\{u \in (-\infty,0)\}$$

- $$l(x) = \arg\min_{\theta} \mathsf{E}\{\rho_p(Y-\theta)|X=x\}$$

- $$l_h(x) = \arg\min_{\theta} n^{-1} \sum_{i=1}^{n} \rho_p(Y_i-\theta)K_h(x-X_i)$$

  where $K_h(u) = h^{-1}K(u/h)$ is a kernel (symmetric density function with compact support) with bandwidth h

Weight Function

$$\psi(u) = p - \mathbf{1}\{u \in (-\infty, 0)\}$$

$l_h(x)$ and $l(x)$: treated as a zero of $\widetilde{H}_n\{l_h(x), x\}$ and $\widetilde{H}\{l(x), x\}$

where:

$$\widetilde{H}_n\{l_h(x), x\} = 0 : \qquad \widetilde{H}_n(\theta, x) \stackrel{\text{def}}{=} n^{-1} \sum_{i=1}^{n} K_h(x - X_i)\psi(Y_i - \theta)$$

$$\widetilde{H}\{l(x), x\} = 0 : \qquad \widetilde{H}(\theta, x) \stackrel{\text{def}}{=} \int_{\mathbb{R}} f(x, y)\psi(y - \theta)dy$$

$\hat{F}$ Approximation Performance around 0

Lemma

[Franke and Mwita (2003), p14] If assumptions (A1, A2, A4) hold, then for any small enough (positive) $\varepsilon \longrightarrow 0$,

$$\sup_{|t|<\varepsilon, i=1,\dots,n, X_i \in J^*} |\hat{F}_i(t) - F(t|X_i)| = \mathcal{O}_p(S_n \delta_n \varepsilon^{1/2} + \varepsilon^2). \qquad (2)$$

The Bootstrap Couple

- $U_1, \ldots, U_n$: i.i.d. uniform $[0, 1]$ rv's
- Bootstrap sample

$$Y_i^* = l_g(X_i) + \hat{F}_i^{-1}(U_i), \quad i = 1, \ldots, n$$

- Couple with the true conditional distribution:

$$Y_i^{\#} = l(X_i) + F^{-1}(U_i|X_i), \quad i = 1, \ldots, n.$$

Given $X_1, \ldots, X_n$: $Y_1, \ldots, Y_n$ and $Y_1^{\#}, \ldots, Y_n^{\#}$ are equally distributed.

A Very Close Couple

$$Y_i^* = l_g(X_i) + \hat{F}_i^{-1}(U_i), \quad i = 1, \ldots, n$$

$$Y_i^{\#} = l(X_i) + F^{-1}(U_i|X_i), \quad i = 1, \ldots, n.$$

Values of $Y_i^{\#}$ and Y_i^* are meaningful only if $|U_i - p| < S_n\delta_n$.
By the inverse function theorem around p, we have

$$\max_{i:|Y_i^{\#} - l(X_i)| < S_n\delta_n} |Y_i^{\#} - l(X_i) - Y_i^* + l_g(X_i)| = \mathcal{O}_p\{S_n\delta_n^{3/2}\}.$$

How Close?

- $q_{hi}(Y_1, \ldots, Y_n) \stackrel{\text{def}}{=} l_h(X_i)$ for data set $\{(X_i, Y_i)\}_{i=1}^n$
- Assumption A3 gives:

$$\max_{|X_i - x_j| < ch} |l_g(X_i) - l_g(X_j) - l(X_i) + l(X_j)| = \mathcal{O}_p(\delta_n)$$

- l_h^* and $l_h^\#$: local bootstrap quantile and its coupled sample analogue. Then

$$l_h^*(X_i) - l_g(X_i) = q_{hi}[\{Y_j^* - l_g(X_j) + l_g(X_j) - l_g(X_i)\})_{j=1}^n]$$

$$l_h^\#(X_i) - l(X_i) = q_{hi}[\{Y_j^\# - l(X_j) + l(X_j) - l(X_i)\}_{j=1}^n]$$

Thus

$$\max_i |l_h^*(X_i) - l_g(X_i) - l_h^\#(X_i) - l(X_i)| = \mathcal{O}_p(\delta_n).$$

Bootstrapping Approximation Rate

Theorem

If assumptions (A1)–(A3) hold, then

$$\sup_{x \in J^*} |l_h^*(x) - l_g(x) - l_h^{\#}(x) - l(x)| = \mathcal{O}_p(\delta_n) = \mathcal{O}_p(n^{-2/5}).$$

- Bootstrap improves the rate of convergence.

Why Oversmoothing?

- To handle the bias (closer). Tuning parameter: g
- Härdle and Marron (1991), let

$$
\begin{aligned}
b_h(x) &\stackrel{\text{def}}{=} \mathsf{E}\, l_h^{\#}(x) - l(x) \\
\hat{b}_{h,g}(x) &\stackrel{\text{def}}{=} \mathsf{E}^*\, l_h^*(x) - l_g(x)
\end{aligned}
$$

- Investigate MSE $\mathsf{E}\left[\left\{\hat{b}_{h,g}(x) - b_h(x)\right\}^2 | X_1, \ldots, X_n\right]$.
  How fast it converges to 0?

Oversmoothing

Theorem

Under some assumptions, for any $x \in J^*$

$$\mathsf{E}\left[\left\{\hat{b}_{h,g}(x) - b_h(x)\right\}^2 |X_1, \ldots, X_n\right] \sim h^4\{\mathcal{O}_p(g^4) + \mathcal{O}_p(n^{-1}g^{-5})\}$$

in the sense that the ratio between the RHS and the LHS tends in probability to 1 for some constants $C_1,\ C_2$.

To minimize MSE, $g = \mathcal{O}(n^{-1/9})$, $g \gg h$, where $h = \mathcal{O}(n^{-1/5})$

The Multivariate Case

- $x = (u, v)^\top \in \mathbb{R}^d$, $v \in \mathbb{R}$:

$$\tilde{l}(x) = u^\top \beta + l(v)$$

- Estimation idea: ANOVA, approximately linear form (locally)
- Partition $[0, 1]$ (for v) in a_n intervals I_{ni} & regard $l(v)$ as a constant item inside I_{ni}.

Two Stage Estimation Procedure

- Linear quantile regression inside each I_{ni} + Weighted mean yields $\hat{\beta}$:

$$\hat{\beta} = \arg\min_{\beta} \min_{l_1,\dots,l_{a_n}} \sum_{i=1}^{n} \psi\{Y_i - \beta^\top U_i - \sum_{j=1}^{a_n} l_j \mathbf{1}(V_i \in I_{ni})\}$$

- Smooth quantile estimate $\hat{l}_h(v)$ from $(V_i, Y_i - U_i^\top \hat{\beta})_{i=1}^{n}$.

Theorem

$\exists$ *positive definite matrices* $D_n,\ C_n$, *s.t.*

$$\sqrt{n}(\hat{\beta} - \beta) \xrightarrow{\mathcal{L}} N\{0, p(1-p)D_n^{-1} C_n D_n^{-1}\} \text{ as } n \to \infty.$$

Uniform Consistency of $\hat{l}_h(v)$

Lemma

Under assumptions (A7) & (A8), we have a.s. as $n \to \infty$

$$\sup_{v \in J^*} |\hat{l}_h(v) - l(v)| \leq C_5 \max\{(nh/\log n)^{-1/2}, h^{\tilde{\alpha}}\} \qquad (3)$$

with another constant C_5 not depending on n. If additionally $\tilde{\alpha} \geqslant \{\log(\sqrt{\log n}) - \log(\sqrt{nh})\}/\log h$, (3) can be further simplified to:

$$\sup_{v \in J^*} |\hat{l}_h(v) - l(v)| \leq C_5\{(nh/\log n)^{-1/2}\}.$$

Multidimensional Uniform Confidence Bands

- Estimation error for parametric part: $\mathcal{O}_p(n^{-1/2})$.
- Bootstrapping approximation error for nonparametric part: $\mathcal{O}_p(n^{-2/5})$, dominating!

Corollary

Under the assumptions (A1) - (A8), an approximate $(1-\alpha)\times 100\%$ confidence band over $\mathbb{R}^{d-1}\times[0,1]$ is

$$u^\top \hat{\beta} + l_h(v) \quad \pm \quad \left[\hat{f}\{l(x)|x\}\sqrt{\hat{f}_X(x)}\right]^{-1} d_\alpha^*,$$

where d_α^* is based on the bootstrap sample (specify later).

How to Bootstrap?

1) Simulate $\{(X_i, Y_i)\}_{i=1}^n, n = 1000$ w.r.t. $f(x, y)$.

$$f(x,y) = f_{y|x}(y - \sin x)\mathbf{1}(x \in [0,1]), \tag{4}$$

where $f_{y|x}(x)$ is the pdf of $\mathrm{N}(0, x)$.

2) Compute $l_h(x)$ of $Y_1, \ldots, Y_n$ and residuals $\hat{\varepsilon}_i = Y_i - l_h(X_i),\ i = 1, \ldots, n$.
If we choose $p = 0.9$, then $\Phi^{-1}(p) = 1.2816$,
$l(x) = \sin(x) + 1.2816\sqrt{x}$ and the bandwidth is $h = 0.05$.

3) Compute the conditional edf $F_{n|x}$:

$$F_n(t|x) = \frac{\sum_{i=1}^n K_h(x - X_i)\mathbf{1}\{\hat{\varepsilon}_i \leqslant t\}}{\sum_{i=1}^n K_h(x - X_i)}$$

with the quartic kernel

$$K(u) = \frac{15}{16}(1 - u^2)^2, \quad (|u| \leqslant 1).$$

4) Generate rv $\varepsilon^*_{i,b} \sim F_{n|x}$, $b = 1, \dots, B$ and construct the bootstrap sample $Y^*_{i,b}, i = 1, \dots, n, b = 1, \dots, B$ as follows:

$$Y^*_{i,b} = l_g(X_i) + \varepsilon^*_{i,b},$$

with $g = 0.2$.

5) For each bootstrap sample $\{(X_i, Y^*_{i,b})\}_{i=1}^n$, compute l^*_h and the random variable

$$d_b \stackrel{\text{def}}{=} \sup_{X \in J^*} \left[\hat{f}\{l(x)|x\} \sqrt{\hat{f}_X(x)} |l^*_h(x) - l_g(x)| \right]. \qquad (5)$$

6) Calculate the $(1-\alpha)$ quantile d^*_α of $d_1, \dots, d_B$.

7) Construct the bootstrap uniform confidence band centered around $l_h(x)$, i.e. $l_h(x) \pm \left[\hat{f}\{l(x)|x\} \sqrt{\hat{f}_X(x)} \right]^{-1} d^*_\alpha$.

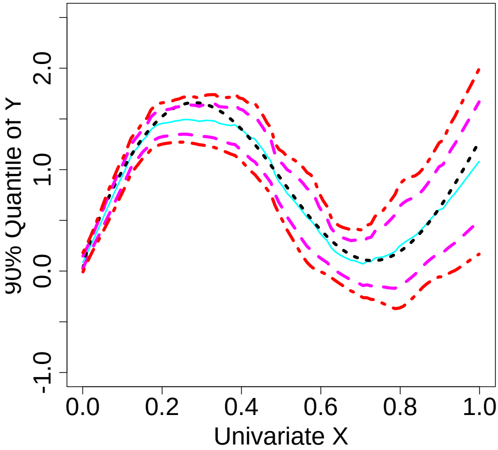


Figure 3: The real 0.9 quantile curve, 0.9 quantile estimate with corresponding 95% uniform confidence band from asymptotic theory and confidence band from bootstrapping.

Convergence Rate (n small)

n	Cov. Prob.	Area
50	0.144 (0.642)	0.58 (1.01)
100	0.178 (0.742)	0.42 (0.58)
200	0.244 (0.862)	0.31 (0.36)

Table 1: Simulated coverage probabilities & areas of nominal asymptotic (bootstrap) 95% confidence bands with 500 repetition.

- For small n, bootstrap's » asymptotic's & not sacrifice much on the band's width
- To achieve same cov. prob., quantile regression usually need more observations than mean regression
- Use larger bandwidth on both X & Y ($\hat{f}^{-1}\{l(x)|x\}$)

PLM QR

- Bivariate data $\{(U_i, V_i, Y_i)\}_{i=1}^{n}, n = 8000$ with:

$$y = 2u + v^2 + \varepsilon - \Phi(p) \tag{6}$$

  where $u \in [0, 2],\ v \in [0, 1]$ and ε is the standard normal rv.

- The real 0.9-quantile curve $\tilde{l}(x) = 2u + v^2$.
- $h = 0.2$ & $g = 0.7$. For the following specific set of random variables, $a_n = 20,\ \hat{\beta} = 2.016758$

of Partitions?

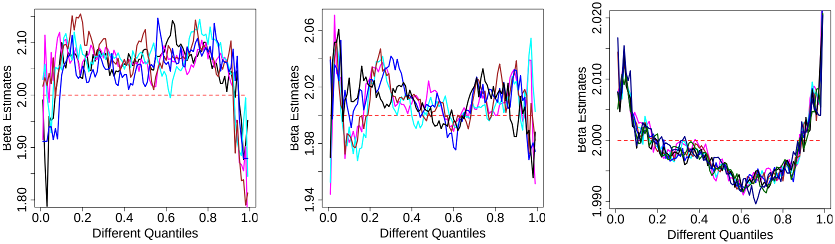


Figure 4: $\hat{\beta}$ with respect to different p for different # of observations, i.e. $n = 1000$, $n = 8000$, $n = 261148$.

a_n	$n = 1000$	$n = 8000$	$n = 261148$
$n^{1/3}/8$			$3.6 * 10^{-3}$
$n^{1/3}/4$	$5.4 * 10^{-1}$	$4.0 * 10^{-2}$	$3.3 * 10^{-3}$
$n^{1/3}/2$	$6.1 * 10^{-1}$	$3.5 * 10^{-2}$	$3.2 * 10^{-3}$
$n^{1/3}$	$6.2 * 10^{-1}$	$3.6 * 10^{-2}$	$3.1 * 10^{-3}$
$n^{1/3} \cdot 2$	$8.0 * 10^{-1}$	$3.9 * 10^{-2}$	$2.9 * 10^{-3}$
$n^{1/3} \cdot 4$	$4.9 * 10^{-1}$	$3.6 * 10^{-2}$	$2.8 * 10^{-3}$
$n^{1/3} \cdot 8$			$3.4 * 10^{-3}$

Table 2: SSE of $\hat{\beta}$ with respect to a_n for different numbers of observations.

- Suggest $a_n = n^{1/3}$ (cost / performance)

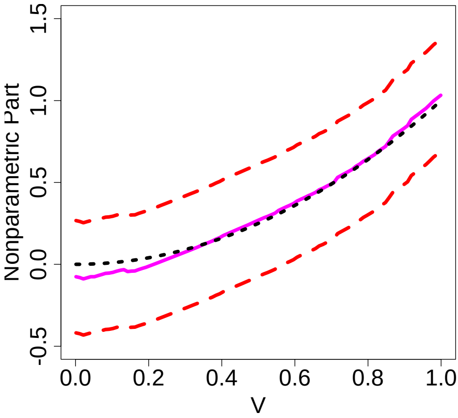


Figure 5: Nonparametric part smoothing, real 0.9 quantile curve with respect to v, 0.9 quantile smoother with corresponding 95% bootstrap uniform confidence band.

Labor Market Application

- How income depends on age w.r.t. different education levels?
- Relation: $\log(\text{Wage}) \sim \beta \cdot \text{Education} + l(\text{Age})$
- Administrative data from the German National Pension Office
- Male, born 1939 $\sim$ 1942, sample 25 - 59, full-time, begin receiving a pension in 2004 $\sim$ 2005

- Education categories: “no answer", “low education”, “apprenticeship” and “university”
- Normal impression:

$$
\begin{aligned}
& \mathsf{E}(y|v, u = \text{Low education}) \\
< \; & \mathsf{E}(y|v, u = \text{Apprenticeship}) \\
< \; & \mathsf{E}(y|v, u = \text{University})
\end{aligned}
$$

Box Plot

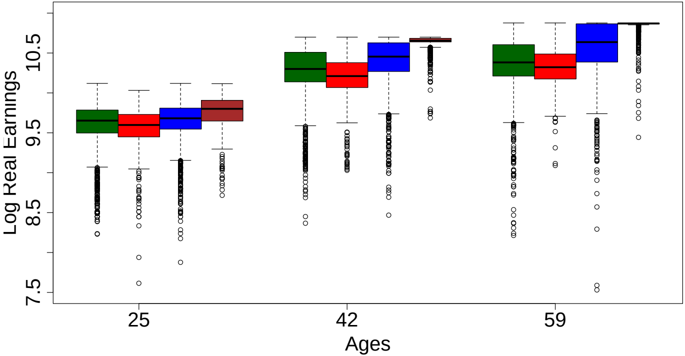


Figure 6: Boxplots for "no answer", "low education", "apprenticeship" & "university" groups corresponding to $v = 0, 0.5, 1$.

- Drop "no answer" group & $n = 175760$ observations
- 1 "low education", 2 "apprenticeship" and 3 "university"
- $175760^{1/3}/2 = 28$ partitions
- Quartic kernel, $h = 0.018$ (after rescaling)

β Estimates

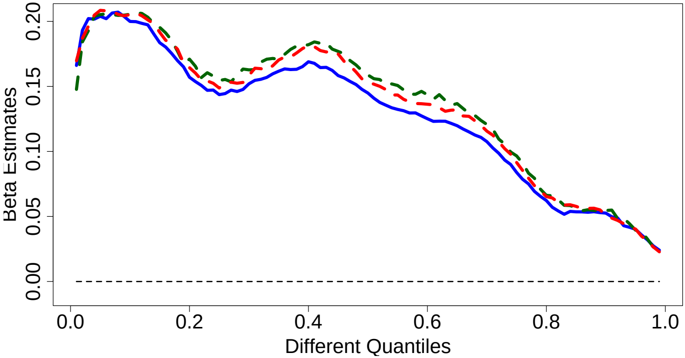


Figure 7: $\hat{\beta}$ corresponding to 8, 16, 32 partitions.

Low Income - Significant

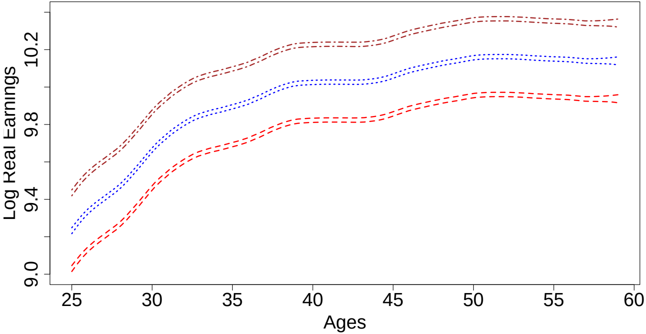


Figure 8: 95% uniform confidence bands for 0.05-quantile smoothers with 3 different education levels "low education", "apprenticeship" & "university".

Median Income - Significant

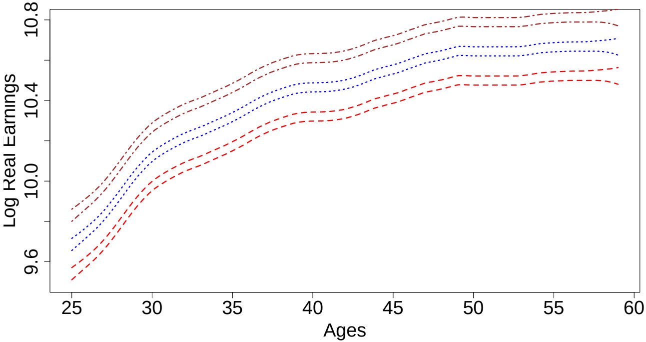


Figure 9: 95% uniform confidence bands for 0.50-quantile smoothers with 3 different education levels "low education", "apprenticeship" & "university".

High Income - Not Significant

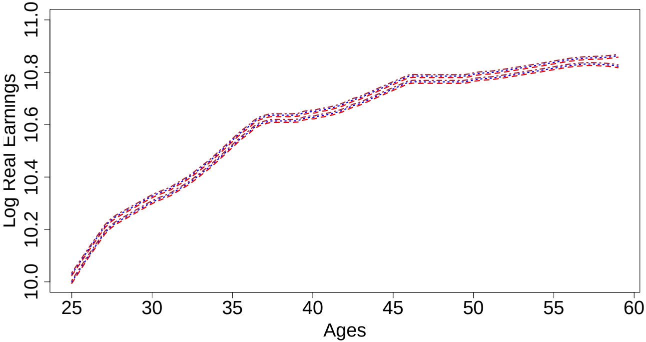


Figure 10: 95% uniform confidence bands for 0.99-quantile smoothers with 3 different education levels "low education", "apprenticeship" & "university"

Real effect of education for income?

- High educations offers a safety line!
- For high end (income) labour, high education no significant effect
    - Smart, no need go to school
    - Be scientist after Ph.D. graduation
    - Poor, not continue school, but hard working & know a lot from practice
    - Education may make people less creative
    - ...
- Causality test, Jeong, Härdle and Song (2009)

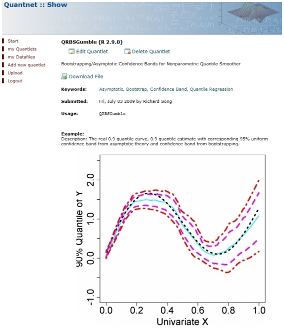
Drawbacks

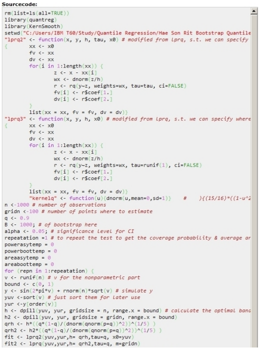
- Very rich people maybe not recorded in the pension system
- Maybe not same retirement time
- Panel data, not exactly i.i.d. (furthur research)
- . . .

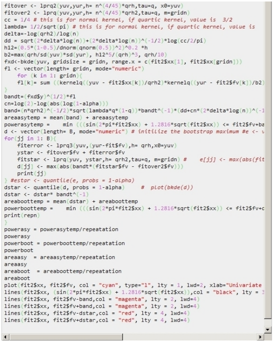
Sth must keep in mind!

- You are dealing with 70-year old people now!
- Time flies (technology level $\uparrow$), more and more high income jobs require high educated people. Time variation of the $\hat{\beta}$? further research.

Quantnet - Open for sharing







References

J. Franke and P. Mwita
Nonparametric Estimates for conditional quantiles of time series
Report in Wirtschaftsmathematik 87, University of Kaiserslautern, 2003.

P. Hall
On convergence rates of suprema
Probab. Th. Rel. Fields, 89:447-455, 1991.

Jinyong Hahn
Bootstrapping Quantile Regression Estimators
Econometric Theory, 11(1):105-121, 1995.

References

W. Härdle and J. S. Marron
Bootstrap simultaneous error bars for nonparametric regression
Ann. Statist., 19:778-796 1991.

W. Härdle and S. Song
Confidence bands in quantile regression
Econometric Theory, forthcoming, 2009.

J. Horowitz
Bootstrap methods for median regression models
Econometrica, 66:1327-1351, 1998.

References

K. Jeong, W. Härdle and S. Song
A Consistent Nonparametric Test for Causality in Quantile
Revise and resubmit to Econometric Theory, 2009.

R. Koenker and K. F. Hallock
Quantile regression
Journal of Econometric Perspectives, 15(4):143¨C156, 2001.

J. Komlós, P. Major and G. Tusnády
An Approximation of Partial Sums of Independent RV'-s, and the Sample DF. I
Z. Wahrscheinlichkeitstheorie verw., 32:111- 131, 1975.

References

K. Yu, Z. Lu and J. Stander
Quantile regression: applications and current research areas
J. Roy. Statistical Society: Series D (The Statistician),
52:331-350, 2003.

Appendix - Assumptions

λ_i and C_i: generic constants.

A1. $X_1, \ldots, X_n$ are an i.i.d. sample, and $f_X(x) \geq \lambda_0$. The quantile function satisfies: $|l'(\cdot)| \leq \lambda_1,\ |l''(\cdot)| \leq \lambda_2$.

A2. $F(t|x)$ have a density, $f(t|x) \geq \lambda_3 > 0$, continuous in x, and in t in the neighborhood of 0. That is, for some $A(\cdot)$ and $f_0(\cdot)$

$$F(t|x') = p + f_0(x)t + A(x)(x' - x) + R(t, x'; x),$$

where $\sup_{t,x,x'} \frac{|R(t,x';x)|}{t^2+|x'-x|^2} < \infty$.

Note that by Assumption A1, $l_h(x)$ is the quantile of a discrete distribution.
This distribution is equivalent to a sample size of $\mathcal{O}_p(nh)$ from a distribution with p-quantile whose biased is $\mathcal{O}_p(h^2)$ relative to the true value.
Let δ_n be the local rate of convergence of the function l_h, essentially $\delta_n = h_n^2 + (nh_n)^{-1/2} = \mathcal{O}(n^{-2/5})$, with $h_n = \mathcal{O}(n^{-1/5})$.

A3. The estimate l_g, satisfies:

$$\begin{aligned} \sup_{x \in J^*} |l_g''(x) - l''(x)| &= \mathcal{O}_p(1), \\ \sup_{x \in J^*} |l_g'(x) - l'(x)| &= \mathcal{O}_p(\delta_n / h) \end{aligned} \tag{7}$$

Note that there is no S_n term in (7) exactly because the bandwidth g_n used to calculate l_g is slightly larger than that used for l_h. As a result l_g has a slightly worse rate of convergence (as an estimator of the quantile function), but its derivatives converge faster.

We assume:

(A4). $f_X(x)$ is twice continuously differentiable and $f(t|x)$ is uniformly bounded in x and t by, say, λ_4.

(A7). The conditional densities $f(\cdot|y),\ y \in \mathbb{R}$, are uniformly local Lipschitz continuous of order $\tilde{\alpha}$ (ulL-$\tilde{\alpha}$) on J, uniformly in $y \in \mathbb{R}$, with $0 < \tilde{\alpha} \leqslant 1$, and $(nh)/\log n \to \infty$.

(A8). $\inf_{v \in J^*} \left| \int \psi\{y - l(v) + \varepsilon\} dF(y|v) \right| \geqslant \tilde{q}|\varepsilon|, \quad \text{for } |\varepsilon| \leqslant \delta_1,$

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