EC709 Advanced Econometrics II

Lecture 3: Quantile Methods

Iván Fernández-Val

Boston University

Fall 2024

General References

- Koenker, R. (2005), *Quantile Regression*, Econometric Society Monographs, 38, Cambridge University Press.
- Koenker, R. (Ed.), Chernozhukov, V. (Ed.), He, X. (Ed.), L. Peng (Ed.) (2018), *Handbook of Quantile Regression,* New York: Chapman and Hall/CRC.

Ouline

1. Exogenous Conditional Quantile Models
2. Koenker and Bassett QR estimator
3. Uniform inference
4. Equivariance and Censored Quantile Regression
5. Models with Endogeneity
6. Treatment effects

1. Exogenous Conditional Quantile Models

- The linear regression model $\mathbb{E}(Y|X) = X'\beta$ shows how the center of the conditional distribution of Y varies with X
- In many applications it is interesting to know how X affects other parts of the distribution
  - How does education affect earnings of those who earn little?
  - How does income affect consumption at different consumption levels?
  - How does smoking affect birth weights of small babies?
- Quantile and distributional methods are designed to answer these questions
- We first consider models where X is exogenous, analog of classical regression models
- Let's start by reviewing some properties of the quantiles

Quantiles and Conditional Quantiles

- Let $F_Y(y)$ denote the CDF of Y. The uth quantile of Y is

$$\mathbb{Q}_Y(u) = \inf\{y : F_Y(y) \geq u\}, \quad u \in (0,1),$$

  i.e., the minimum value of Y such that the amount u of the probability distribution lies below this value

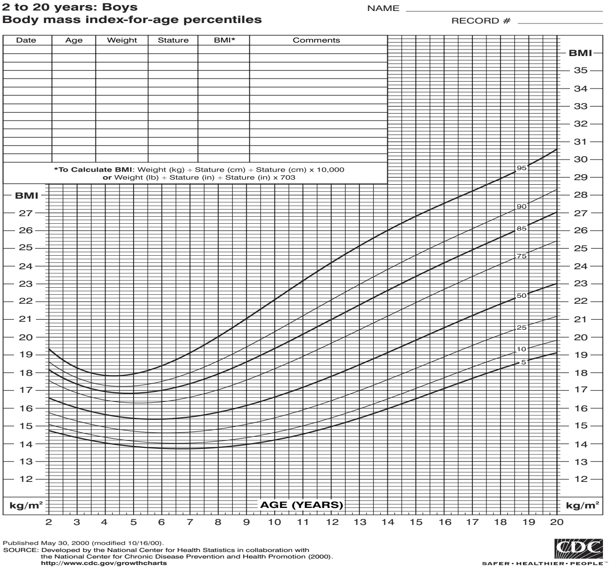
- Let $F_{Y|X}(y|x)$ denote the conditional CDF of Y given X. The uth conditional quantile of Y given X is

$$\mathbb{Q}_{Y|X}(u|x) = \inf\{y : F_{Y|X}(y|x) \geq u\}, \quad u \in (0,1)$$

  i.e., the minimum value of Y such that the amount u of the conditional probability distribution lies below this value

- $x \mapsto \mathbb{Q}_{Y|X}(u|x)$ describes how the uth conditional quantile depends on X, e.g., how the percentiles of height (Y) vary with age (X)

Growth Charts: Quetelet (1870)



Relation with Conditional Distribution

Assume Y is continuous and $y \mapsto F_{Y|X}(y|x)$ is strictly increasing

1 The conditional quantiles solve the moment equations

$$\Pr[Y \leq \mathbb{Q}_{Y|X}(u|x) \mid X = x] = u$$

2 Quantile Function is the Inverse of the Distribution Function

$$F_{Y|X}(y|x) = \Pr[Y \leq y \mid X = x]$$

the inverse is

$$F_{Y|X}^{-1}(u|x) = \{y : \Pr[Y \leq y \mid X = x] = u\}.$$

Thus,

$$\mathbb{Q}_{Y|X}(u|x) = F_{Y|X}^{-1}(u|X).$$

Non-Additive Error Representation

- Any CEF model has an additive error term representation

$$Y = \mathbb{E}(Y|X) + U, \quad \mathbb{E}(U|X) = 0$$

- Skorohod representation for conditional quantiles:

$$Y = \mathbb{Q}_{Y|X}(U|X), \quad U \mid X \sim U(0,1)$$

  Proof: Let

$$U = F_{Y|X}(Y|X) \mid X \sim U(0,1) \text{ check!}$$

  and apply $\mathbb{Q}_{Y|X}(\cdot|X) \equiv F^{-1}_{Y|X}(\cdot|X)$ to both sides

- The error term U represents the rank in the conditional distribution, absorbing ability, skills, or "proneness" (Doksum,74)

- Skorohod representation is non-additive in the error term U

Prediction Properties

- $\mathbb{E}(Y|X)$ is the solution to the prediction problem

$$\mathbb{E}(Y|X) = \arg\min_{b(\cdot)} \mathbb{E}[(Y - b(X))^2]$$

- Let $\rho_u(\varepsilon) = [u - 1(\varepsilon < 0)]\varepsilon$ be the absolute loss or "check" function

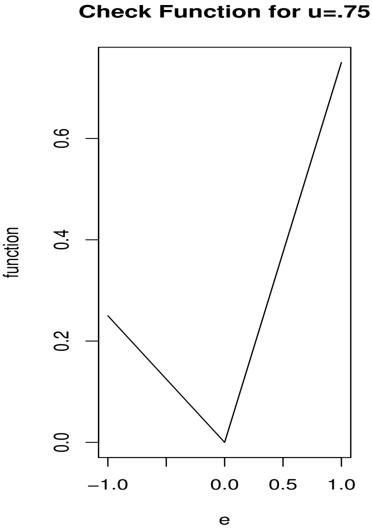
- $\mathbb{Q}_{Y|X}(u|X)$ is a solution to the prediction problem

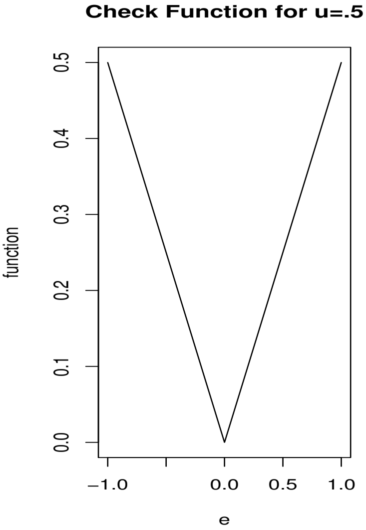
$$\mathbb{Q}_{Y|X}(u|X) \in \arg\min_{b(\cdot)} \mathbb{E}[\rho_u(Y - b(X))]$$

  Proof: for simplicity assume that $X = 1$, then
  $\mathbb{E}[\rho_u(Y - b)] = u \int_b^{\infty} (y - b) dF_Y(y) - (1 - u) \int_{-\infty}^{b} (y - b) dF_Y(y)$.
  Then, taking derivatives wrt b

$$F_Y(b) = u \Rightarrow b = F_Y^{-1}(u)$$

Check Function: $\rho_u(\varepsilon) = u\varepsilon^+ + (1-u)\varepsilon^-$





Linear Quantile Regression Model

- Approximate $\mathbb{Q}_{Y|X}$ using linear forms for convenience:

$$\mathbb{Q}_{Y|X}(u|x) \approx P(x)'\beta(u),$$

  where $P(X)$ includes transformations of original regressor X (polynomials, interactions, ...). We shall use X for $P(X)$.

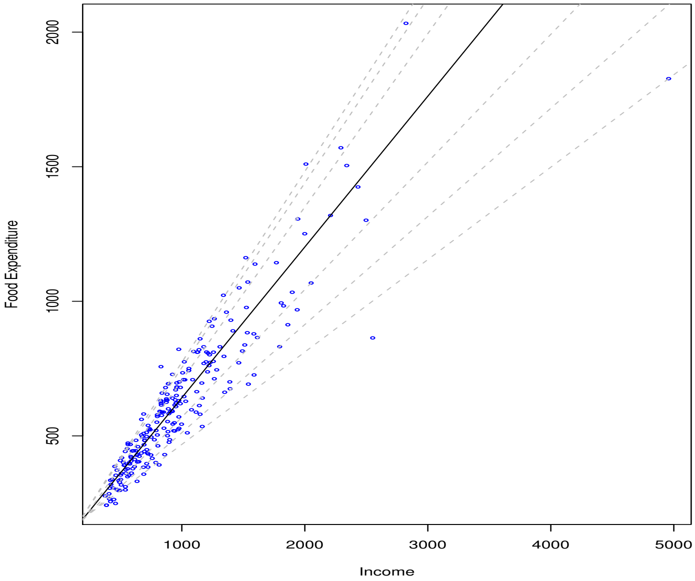
- Coefficient $\beta(u)$ can depend on u. Examples:
  - Conditional Median $X'\beta(1/2), \quad u = 1/2$
  - Conditional Quartile $X'\beta(1/4), \ u = 1/4$
  - Conditional Minimum $X'\beta(0^+), \ u = 0^+$

- X has heterogeneous impact across quantiles: the impact on tails can be very different than the impact in the middle

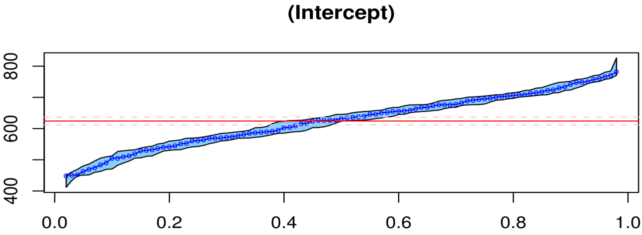
- Random-coefficient representation

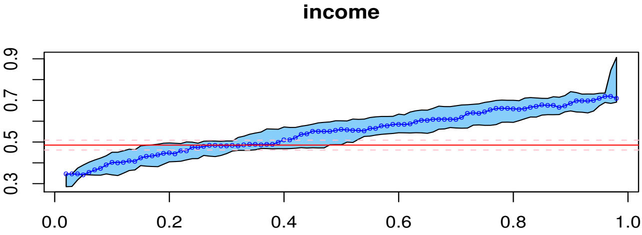
$$Y = X'\beta(U), \quad U \mid X \sim U(0,1)$$

Engel Curves: $x \mapsto x'\beta(u)$



Engel Curves: $u \mapsto \beta(u)$





Location Shift Model

- Classical location shift model

$$Y = X'\beta + \epsilon, \quad \epsilon \perp\!\!\!\perp X$$

- Special case of QR model

$$Y = \mathbb{Q}_{Y|X}(U|X) = X'\beta + \mathbb{Q}_\epsilon(U), \quad U \mid X \sim U(0,1)$$

so that $\mathbb{Q}_{Y|X}(u|x) = x'\beta(u)$ for all $u \in (0,1)$ with

$$\beta(u) = (\beta_1 + \mathbb{Q}_\epsilon(u), \beta_2, ..., \beta_p)'.$$

- Impact of X is summarized by a single number β

Location-Scale Shift Model

- Classical location-scale model

$$Y = X'\beta + X'\gamma \cdot \epsilon, \quad \epsilon \perp\!\!\!\perp X, \quad X'\gamma > 0$$

- Special case of QR model

$$Y = \mathbb{Q}_{Y|X}(U|X) = X'\beta + X'\gamma \cdot \mathbb{Q}_{\epsilon}(U).$$

so that $\mathbb{Q}_{Y|X}(u|x) = x'\beta(u)$ for all $u \in (0,1)$ with

$$\beta(u) = \beta + \gamma \cdot \mathbb{Q}_{\epsilon}(u),$$

- All the slope functions $u \mapsto \beta(u)$
  - are monotone in u, and
  - are affine transformations of each other.
- The Engel example is of this kind.

Treatment Effect Model

- Effect of a binary treatment D on an outcome of interest Y
- Conditional quantile curves $d \mapsto \mathbb{Q}_{Y|D}(u|d)$ are linear:

$$\mathbb{Q}_{Y|D}(u|d) = \alpha(u) + \delta(u)d,$$

$$\alpha(u) = \mathbb{Q}_{Y|D}(u|0), \quad \delta(u) = \mathbb{Q}_{Y|D}(u|1) - \mathbb{Q}_{Y|D}(u|0)$$

- $\delta(u)$ has causal interpretation as quantile treatment effect
- Potential outcomes indexed by treatment status Y_0 and Y_1 and treatment effect $Y_1 - Y_0$
- Observed outcome is $Y = (1 - D)Y_0 + DY_1$
- Under random assignment $(Y_1, Y_0) \perp\!\!\!\perp D$,

$$\delta(u) = \mathrm{QTE}(u) = \mathbb{Q}_{Y_1}(u) - \mathbb{Q}_{Y_0}(u) \overset{?}{=} \mathbb{Q}_{Y_1 - Y_0}(u)$$

- $\mathbb{E}[Y_1] - \mathbb{E}[Y_0] = \mathbb{E}[Y_1 - Y_0]$? why?

Conditional vs Unconditional Effects

- Consider the linear QR model

$$Y = \alpha(U) + \beta(U)X, \quad U \mid X \sim U(0,1)$$

- Conditional mean effect of increasing X by one unit:

$$\mathbb{E}[Y \mid X = x+1] - \mathbb{E}[Y \mid X = x] = \mathbb{E}[\beta(U)](x+1) - \mathbb{E}[\beta(U)]x = \mathbb{E}[\beta(U)]$$

- Unconditional mean effect:

$$\mathbb{E}[\alpha(U) + \beta(U)(X+1)] - \mathbb{E}[\alpha(U) + \beta(U)X] = \mathbb{E}[\beta(U)]$$

- Conditional u-quantile effect:

$$\mathbb{Q}_{Y|X}(u \mid x+1) - \mathbb{Q}_{Y|X}(u \mid x) = \alpha(u) + \beta(u)(x+1) - \alpha(u) - \beta(u)x = \beta(u)$$

- Unconditional u-quantile effect:

$$\mathbb{Q}_{\alpha(U)+\beta(U)(X+1)}(u) - \mathbb{Q}_{\alpha(U)+\beta(U)X}(u) \neq \beta(u)$$

Misspecification of QR Model

- Let $\beta = \arg\min_b \mathbb{E}[(Y - X'b)^2]$. White (80) shows that $X'\beta$ is also the BLP of $\mathbb{E}(Y|X)$ since

$$\beta = \arg\min_b \mathbb{E}[(\mathbb{E}(Y|X) - X'b)^2]$$

- Let $\beta(u) = \arg\min_b \mathbb{E}[\rho_u(Y - X'b)]$. Angrist, Chernozhukov and F.-V. (06) shows that $X'\beta(u)$ is a BLP of $\mathbb{Q}_{Y|X}(u|X)$ since

$$\beta(u) = \arg\min_b \mathbb{E}[w_u(X,b)(\mathbb{Q}_{Y|X}(u|X) - X'b)^2]$$

where

$$w_u(X,b) = \int_0^1 (1-u) f_{Y|X}(uX'b + (1-u)\mathbb{Q}_{Y|X}(u|X)|X)du$$

Applications of Quantile Regression Models

1. Growth charts – dependence of height and weight quantiles on age (Quetelet, 1870)
2. Stochastic dominance/welfare (Doksum, *Annals*, 1974, Heckman and Smith, *REStud*, 1997)
3. Residual wage inequality (Hogg, *JASA*, 1975, Buchinsky, *Econometrica*, 1994)
4. Impact of maternal behavior and prenatal care on quantiles of infant birth-weights
    - "normal" quantiles (Abrevaya, 2001, Koenker, 2005)
    - extremal quantiles (250-1000 grams) (Chernozhukov and F.-V., 2011)

Applications of Quantile Regression Models

5. Heterogeneous Engel curves for food (Deaton, 1997, Koenker, 2005)

6. Value-at-risk (extreme risk) forecasting (Chernozhukov and Umantsev, 2000, Engle and Maganelli, 2000)

7. Production Frontiers and Probabilistic Frontiers (Timmer, 1971)

   $$\mathbb{Q}_{Y|X}(1|x) \qquad \mathbb{Q}_{Y|X}(1-\epsilon|x)$$

   Y production, X factors

8. Reservation Rules in Search Model (Flinn and Heckman, 1981)

   $$\mathbb{Q}_{W|X}(\epsilon|X) = \text{Approximate Reservation Wage Function}$$

   W accepted wage, X covariates

2. Estimation: Sample Quantiles

$(Y_1, ..., Y_n)$ random sample of Y. There are 2 ways to obtain sample quantiles

1. Sorting: $Y_{(1)} \leq ... \leq Y_{(n)}$ ordered sample. Then,

$$\hat{\mathbb{Q}}_Y(u) = \min\{Y_{(j)} : j/n \geq u\} = Y_{(\lceil nu \rceil)}$$

2. Optimization: recall $\rho_u(\varepsilon) = [u - 1(\varepsilon < 0)]\varepsilon$

$$\hat{\mathbb{Q}}_Y(u) \in \arg\min_q \frac{1}{n} \sum_{i=1}^{n} \rho_u(Y_i - q) = Y_{(\lceil nu \rceil)}$$

Optimization approach extends naturally to the regression case and is convenient to derive statistical properties

Koenker and Bassett (78) Quantile Regression Estimator

- QR originated in the work of Laplace (1818)
- Let $(Y_1, X_1), ..., (Y_n, X_n)$ be a random sample of (Y, X)
- Regression u-Quantile

  $$\hat{\beta}(u) = \arg\min_{\beta} \mathbb{E}_n[\rho_u(Y_i - X_i'\beta)]$$

  where $\mathbb{E}_n$ is the empirical average $\mathbb{E}_n[Z_i] = n^{-1}\sum_{i=1}^n Z_i$
- The great motivation for this objective function is computational, as the solution can be easily computed due to convexity of the objective function
- The first order conditions for this problem are the “right” moments

Properties of Quantile Regression Program

- Sample analog of the population program

$$\beta(u) = \arg\min_{\beta} \mathbb{E}[\rho_u(Y - X'\beta)]$$

- Since $\rho_u(Y - X'\beta) = [u - 1(Y < X'\beta)](Y - X'\beta)$,

$$\nabla_{\beta}\rho_u(Y - X'\beta) = -(u - 1\{Y < X'\beta\})X$$

- Thus, the population parameter solves the moment conditions

$$\mathbb{E}\left[\{u - \Pr(Y < X'\beta(u)|X)\}X\right] = 0.$$

- Correct moment restrictions for conditional quantiles
- The solution at $\beta = \beta(u)$ is unique if the Hessian of the limit objective function

$$J(u) = \nabla_{\beta\beta'}\mathbb{E}[\rho_u(Y - X'\beta)]\Big|_{\beta=\beta(u)} = \mathbb{E}[f_{Y|X}(X'\beta(u)|X)XX'],$$

is full rank (and thus is positive definite due to convexity)

Asymptotic Theory

- Theorem (Koenker, 05): Let $u \in (0,1)$ be a fixed index. Under appropriate regularity conditions, the estimator $\hat{\beta}(u)$ is CAN:

$$\sqrt{n}(\hat{\beta}(u) - \beta(u)) \to_d N(0, J(u)^{-1}\Omega(u,u)J(u)^{-1}),$$

  where $\Omega(u,u) = \mathbb{E}[(u - 1\{Y \leq X'\beta(u)\})^2 XX']$ which simplifies to $\Omega_0(u,u) = u(1-u)\mathbb{E}[XX']$ if $\mathbb{Q}_{Y|X}(u|X)$ is linear

- For standard errors, use $\hat{\Omega}(u,u) = \mathbb{E}_n[(u - 1\{Y_i \leq X_i'\hat{\beta}(u)\})^2 X_i X_i']$ and Powell's Hessian estimator $\hat{J}(u) = \mathbb{E}_n[1\{|Y_i - X_i'\hat{\beta}(u)| \leq h_n\} X_i X_i']/h_n$ where $h_n \to 0$ and $\sqrt{n}h_n \to \infty$. Then,

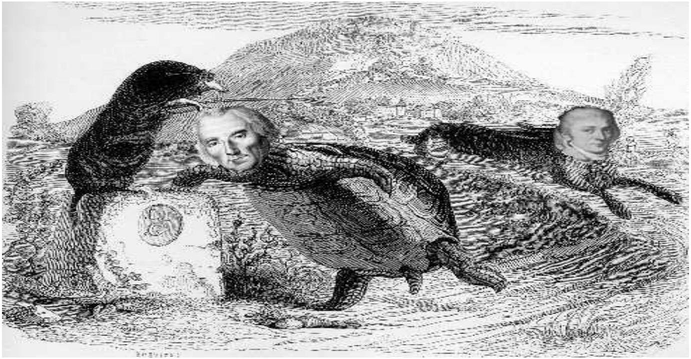
$$\hat{\Omega}(u,u) \to_P \Omega(u,u), \quad \text{and } \hat{J}(u) \to_P J(u)$$

- Nonparametric bootstrap is also consistent to obtain standard errors (Hahn, 95, 97)

Computation

- Unlike OLS, QR problem does not have closed form solution; standard nonlinear optimization methods do not work because objective function is non-differentiable
- QR problem has a dual linear programming representation
- Computation is very fast using interior point methods & preprocessing (Portnoy and Koenker, Statistical Science)
- There are efficient algorithms to estimate the entire QR process
$$\{\hat{\beta}(u) : u \in \mathcal{U} \subset (0,1)\}$$
- Koenker's R-package $\mathrm{quantreg}$ implements all these methods
- Running time is $O_p(n^{1+\delta}\dim(\beta)^3)$, faster than OLS.

Laplacian Tortoise has caught up with Gaussian Hare



3. Uniform Inference

- Let $X = (1, D)$ and $\beta(u) = (\alpha(u), \delta(u))$, where $D \in \{0, 1\}$ is a treatment indicator
- Many hypotheses of interest in policy analysis involve the entire distribution
  1. No effect: $\delta(u) = 0, \quad u \in \mathcal{U} \subset (0, 1)$
  2. Constant effect: $\delta(u) = c, \quad u \in \mathcal{U} \subset (0, 1)$
  3. Stochastic dominance: $\delta(u) \geq 0, \quad u \in \mathcal{U} \subset (0, 1)$
- Angrist, Chernozhukov and F.-V. (06) derive a functional central limit theorem for QR:

$$J(u)\sqrt{n}(\hat{\beta}(u) - \beta(u)) \rightarrow_d Z(u), \quad u \in \mathcal{U} \subset (0, 1)$$

  where Z is a gaussian process with covariance function

$$\Omega(u, v) = \mathbb{E}[(u - 1\{Y \leq X'\beta(u)\})(v - 1\{Y \leq X'\beta(v)\})XX']$$

- The limit process is useful to perform uniform inference on $\beta(\cdot)$ using Kolmogorov-Smirnov type methods

Kolmogorov-Smirnov Tests

- Consider the general functional hypothesis:

$$H_0 : R(u)\beta(u) = r(u), \quad u \in \mathcal{U} \subset (0, 1)$$

- Kolmogorov-Smirnov test statistic is

$$S_n = \sup_{u \in \mathcal{U}} \|\sqrt{n}[R(u)\hat{\beta}(u) - r(u)]\|_{\hat{V}(u)}$$

where $V(u) = R(u)J(u)^{-1}\Omega(u, u)J(u)^{-1}R(u)'$

- Chernozhukov and F-V. (05) and Angrist, Chernozhukov and F.-V. (06) show that recentered bootstrap consistently estimates the distribution of S_n
- $R(u)$ and $r(u)$ are replaced by estimates if they are unknown

Confidence Bands for Coefficients $u \mapsto \beta_k(u)$

- An asymptotic $(1-\alpha)$-confidence band for $u \mapsto \beta_k(u)$ is formed by two functions $u \mapsto L_k(u)$ and $u \mapsto U_k(u)$ such that

$$\Pr(L_k(u) \leq \beta_k(u) \leq U_k(u), u \in \mathcal{U}) \to 1-\alpha$$

- Invert Kolmogorov-Smirnov test for $\beta_k(u)$

$$\Pr\left(\sup_{u\in\mathcal{U}} \frac{|\hat{\beta}_k(u) - \beta_k(u)|}{\sqrt{[J(u)^{-1}\Omega(u,u)J(u)^{-1}/n]_{(k,k)}}} \leq c_{1-\alpha}\right) \to 1-\alpha,$$

  where $c_{1-\alpha}$ is the $(1-\alpha)$-quantile of

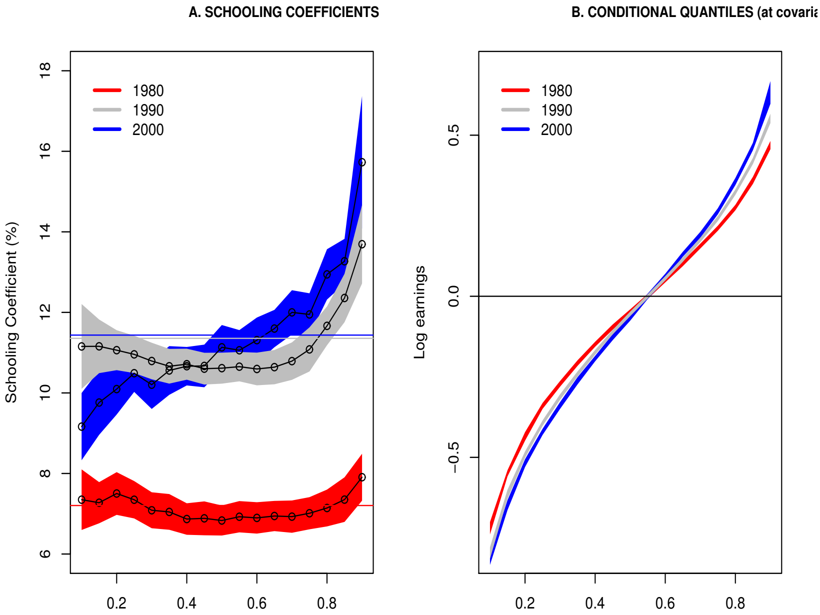
$$\sup_{u\in\mathcal{U}} |[\Omega(u,u)^{-1/2}Z(u)]_k|$$

- Replace $J(u)^{-1}\Omega(u,u)J(u)^{-1}$ and $c_{1-\alpha}$ by uniformly consistent estimators

Example: Evolution of Returns to Schooling 1980–2000

- From Angrist, Chernozhukov and F.-V. (06)
- Y is log-wages for prime age white men, and X includes schooling and quadratic function in experience
- 95% confident bands for returns to schooling in 1980, 1990, and 2000, constructed by subsampling
- Significant increase in returns to schooling between 1980 and 1990, returns become more step in 2000
- Reject hypothesis of constant returns in 1990 and 2000; cannot reject in 1980 if we exclude upper tail
- 95% confidence bands for recentered wage distribution: increase in wage dispersion

Returns to Schooling in the U.S.: 1980–2000



4. Equivariance and Censored Quantile Regression

- Quantile of an increasing function is the function of the quantile
- Theorem (Equivariance to Increasing Transformations): For any weakly increasing transformation $T[\cdot \mid X]$, which is possibly X-dependent,

$$T[Y \mid X] \equiv T[\mathbb{Q}_{Y|X}(U \mid X) \mid X] \equiv \mathbb{Q}_{T[Y|X]|X}(U \mid X)$$

- Proof: $T[\mathbb{Q}_{Y|X}(\cdot \mid X) \mid X]$ has domain $[0, 1]$ and is weakly increasing because it is the composition of two weakly increasing functions. Thus it is the proper quantile function.
- In general, expectations do not satisfy this invariance:

$$\mathbb{E}(T[Y \mid X] \mid X) \neq T[\mathbb{E}(Y \mid X) \mid X].$$

  Affine or location-scale transformations commute with conditional expectations, but more general transformations usually do not.

Examples

1. $Y = \ln W$ and $\mathbb{Q}_{Y|X}(u \mid x) = x'\beta(u)$, then

$$\mathbb{Q}_{W|X}(u \mid x) = \exp(x'\beta(u)).$$

   The same can not be done *generally* for mean regression.

2. $Y^b = 1\{Y > 0\}$ and $\mathbb{Q}_{Y|X}(u \mid x) = x'\beta(u)$, then

$$\mathbb{Q}_{Y^b|X}(u \mid x) = 1\{x'\beta(u) > 0\}$$

   This is the binary quantile regression or *maximum score model*. Estimation can be done using the maximum score estimator (Manski, 1975, Horowitz, 1997, Kordas, 2005)

$$\hat{\beta}(u) \in \arg\min_{\beta} \mathbb{E}_n\{\rho_u(Y_i^b - 1\{X_i'\beta > 0\})\}$$

   This estimator is very hard to compute.

Censored Quantile Regression

- $Y^c = \max[0, Y]$ and $\mathbb{Q}_{Y|X}(u \mid x) = x'\beta(u)$, then

$$\mathbb{Q}_{Y^c|X}(u \mid x) = \max[0, x'\beta(u)]$$

  This is the *censored quantile regression model*.

- Estimation can be done using nonlinear quantile regression

$$\hat{\beta}(u) \in \arg\min_{\beta} \mathbb{E}_n\{\rho_u(Y_i^c - \max[0, X_i'\beta])\}$$

- This is the censored quantile estimator (Powell, 84)
- Computation is challenging because CQR problem is not convex in β. Chernozhukov and Hong (02) and Bilias, Florios and Skouras (19) give convenient algorithms
- Implemented in R command $\mathrm{quantreg}$ (Koenker, 18) and Stata command cqiv (Chernozhukov, F.-V., Han and Kowalski, 18)

5. Models with Endogeneity

- Standard IV algebra in the error representation $Y = X'\beta(U)$, $U \mid Z \sim U(0,1)$, does not work due to nonlinearity of quantile operator

$$\mathbb{Q}_{Y|Z}(u \mid Z) = \mathbb{Q}_{X'\beta(U)|Z}(u \mid Z) \neq \mathbb{Q}_{X|Z}(u \mid Z)'\beta(u),$$

  whereas $\mathbb{E}[Y \mid Z] = \mathbb{E}[X \mid Z]'\beta$ in $Y = X'\beta + U$, $\mathbb{E}[U \mid Z] = 0$

- Conditional quantiles satisfy the moment conditions

$$\mathbb{E}[u - 1\{Y \leq \mathbb{Q}_{Y|X}(u \mid X)\} \mid X] = 0$$

- $u - 1\{Y \leq \mathbb{Q}_{Y|X}(u \mid X)\}$ plays the role of residual
- An IV version of these moment conditions is

$$\mathbb{E}[u - 1\{Y \leq Q(u, X)\} \mid Z] = 0$$

  where Z is some "instrument"

- Does $Q(u, x)$ correspond to any structural function?

An Endogenous Model (Chernozhukov and Hansen, 05)

- Consider the general nonseparable model

$$Y = g(X, \varepsilon)$$

- Assumption IVQR: (1) $\dim \varepsilon = 1$; (2) $\varepsilon \mapsto g(\cdot, \varepsilon)$ is strictly increasing; (3) there exists Z such that $Z \perp\!\!\!\perp \varepsilon$
- Normalize $\varepsilon \sim U(0,1)$
- Examples:
  1. Location-scale shift model:

$$Y = X'\beta + (X'\gamma) \cdot \varepsilon, \quad X'\gamma > 0, \quad \varepsilon \perp\!\!\!\perp Z$$

  2. Linear quantile regression model:

$$Y = X'\beta(\varepsilon), \quad X'\dot{\beta}(\varepsilon) > 0, \quad \varepsilon \perp\!\!\!\perp Z$$

- Lemma QSF: Under IVQR $g(x, u)$ is the u-quantile of $g(x, \varepsilon)$ and

$$\mathrm{QTE}_{x,x'}(u) = g(x', u) - g(x, u)$$

Demand-Supply Example

- Consider the log-linear demand-supply system

$$\begin{aligned}
&\ln Q_p = \beta_0(\varepsilon) + \beta_1(\varepsilon) \cdot \ln(p),\\
&\ln S_p = f(p, Z, \eta),\\
&P \in \{p : Q_p = S_p\},\\
&\varepsilon \text{ independent of } Z \text{ and normalize } \varepsilon \mid Z \sim U(0,1).
\end{aligned}$$

- ε is a demand disturbance, "level of demand"; η is a supply disturbance, "level of supply", Z are supply shifters;
- $p \mapsto \beta_0(u) + \beta_1(u)\ln(p)$ is the u-quantile demand curve
- Demand elasticity $\beta_1(u)$ varies with u
- Equilibrium quantity Y and price P will satisfy

$$\begin{aligned}
\ln Y &= \beta_0(\varepsilon) + \beta_1(\varepsilon)\ln P,\\
P &= \delta(Z, \underbrace{\varepsilon, \eta, \text{"}\textit{sunspots}\text{"}}_{V}),
\end{aligned}$$

Moment Conditions

- The main statistical implication of the model is

$$\mathbb{E}[u - 1\{Y \leq g(X, u)\} \mid Z] = 0$$

i.e., $Q(u, x) = g(x, u)$

- This follows from independence and

$$\{Y \leq g(X, u)\} \Leftrightarrow \{g(X, \varepsilon) \leq g(X, u)\} \Leftrightarrow \{\varepsilon \leq u\}.$$

- Under completeness conditions the solution to conditional moment equation is unique (Newey and Powell, 03, and Chernozhukov and Hansen, 13)
- Two possibilities for estimation
  1. Quasi-Bayesian GMM (Chernozhukov and Hong, 03)
  2. Inverse Quantile Regression (Chernozhukov and Hansen, 06)

Inverse Quantile Regression

- Assume linear quantile model:

$$Y = g(X, \varepsilon) = \alpha(\varepsilon) + X\beta(\varepsilon), \quad \varepsilon \mid Z \sim U(0,1)$$

- Principle: Find $X\beta$ such that the QR of $Y - X\beta$ on Z returns $\mathbf{0}$ as the estimate of the coefficients on Z since

$$\mathbb{E}[Z(u - 1\{Y - X\beta(u) < \alpha(u) + Z'\gamma(u)\})] = 0 \text{ for } \gamma(u) = 0$$

- Two-step procedure
  1. Given $\beta \in \mathcal{B}$, run QR of $Y - X'\beta$ on Z:

  $$[\hat{\alpha}(\beta, u), \hat{\gamma}(\beta, u)] \equiv \arg\min_{\alpha,\gamma} \mathbb{E}_n\Big[\rho_u(Y - X\beta - \alpha - Z'\gamma)\Big],$$

  2. Pick β such that the *Wald* statistic for testing the exclusion of Z is as small as possible:

  $$\hat{\beta}(u) \in \arg\inf_{\beta\in\mathcal{B}} \; W_n[\beta] = n\Big(\hat{\gamma}(\beta, u)'\widehat{\Omega}_{\gamma}^{-1}[\beta]\hat{\gamma}(\beta, u)\Big)$$

  Estimate $\alpha(u)$ by $\hat{\alpha}(\hat{\beta}(u), u)$

- Possible to incorporate covariates replacing α by $C'\alpha$

Properties of IQR

- Implementation details:
  1. Given a grid $\{\beta_j : j = 1, ..., J\}$, run J QRs of $Y - X'\beta_j$ on Z
  2. Pick $\hat{\beta}(u)$ as the value of β_j that minimizes $W_n[\beta_j]$ – the testing statistic for excluding the instruments.

- IQR is consistent and asymptotically normal

- IQR works very well if X is one- or two-dimensional; when X is high-dimensional, the quasi-Bayesian GMM approach is computationally more attractive

- A confidence region for $\beta(u)$ can be formed by collecting all α_j' st.

$$W_n[\beta_j] \leq c_p,$$

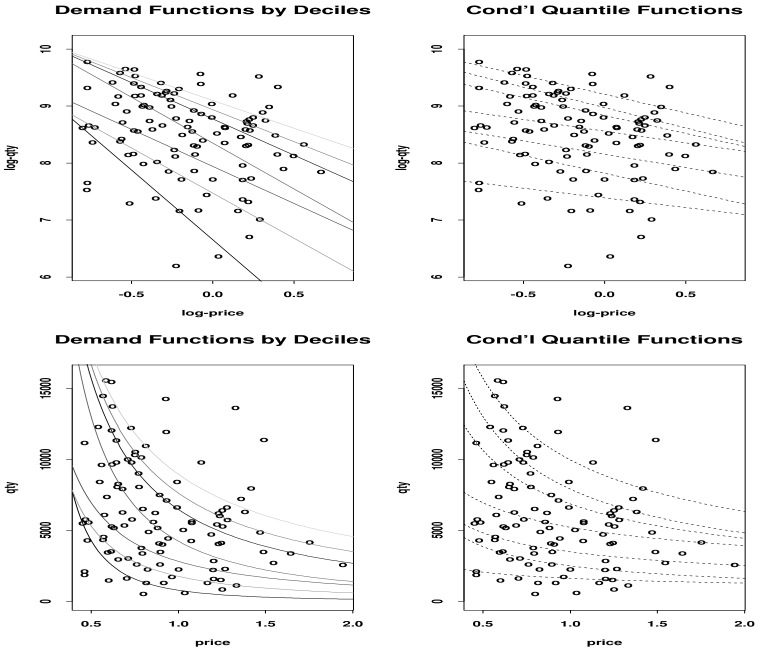
  where c_p is the critical value for the Wald statistic.

- This confidence region is robust to partial identification and weak identification.

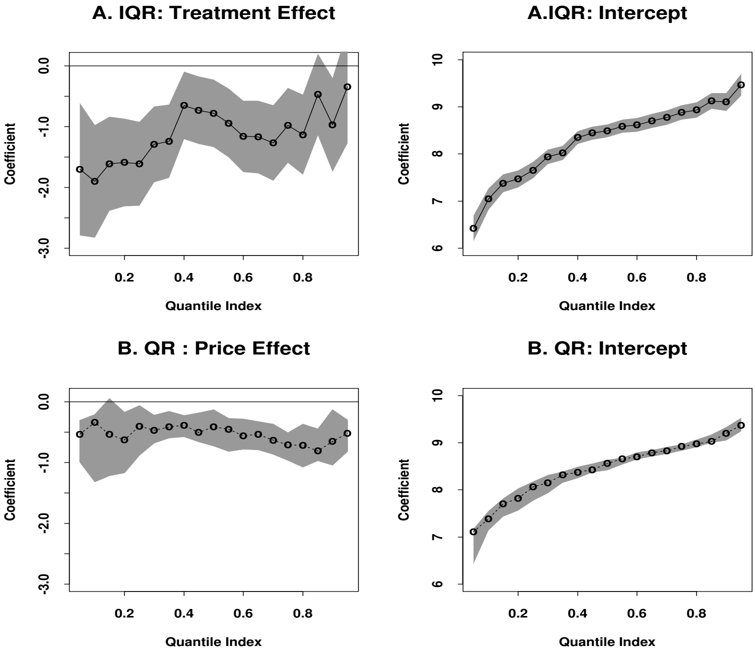
Demand Application: Chernozhukov and Hansen (08)

- Data from Fulton Fish Market over the period from December 2, 1991 to May 8, 1992 (Graddy, 95)
- Q and P are daily total quantity purchased and average price of whiting fish
- Z are supply shifters such as weather and conditions of the sea.
- Compare standard QR estimates of the price elasticity with IQR estimates that allow for endogeneity
- Cobb-Douglas log-linear demand specification

Demand Function Estimates



Elasticity Estimates



6. Treatment Effects

- What restrictions does Assumption IVQR impose in a treatment effect model?
- Recall the treatment effect model with IV

$$Y = Y_0 + (Y_1 - Y_0)D, \quad (Y_0, Y_1) \perp\!\!\!\perp Z$$

- By the Skorohod representation $Y_d = q(\varepsilon_d, d)$ with $\varepsilon_d \sim U(0,1)$ and $d \in \{0,1\}$, so that for $\varepsilon = (\varepsilon_0, \varepsilon_1)$

$$Y = g(D, \varepsilon) = q(\varepsilon_0, 0)(1 - D) + q(\varepsilon_1, 1)D$$

- Rank invariance ($\varepsilon_0 = \varepsilon_1$) is sufficient for IVQR
- Individual rank does not change across treatment statuses

Rank Similarity

- CH (05) give a weaker rank similarity condition

$$\varepsilon_0 \overset{d}{=} \varepsilon_1 \mid Z, V, \quad D = \delta(Z, V),$$

which delivers the statistic implication of IVQR in the treatment effect model

- Allows random deviations or slippages from rank invariance
- Alternative representation for observed outcomes

$$Y = q(\varepsilon_D, D)$$

- By LIE and rank similarity

$$\begin{aligned}\Pr(\varepsilon_D \leq u \mid Z) = \mathbb{E}[\Pr(\varepsilon_{\delta(Z,V)} \leq u \mid Z, V) \mid Z] \\ = \Pr(\varepsilon_d \leq u \mid Z) = u\end{aligned}$$

Local Quantile Treatment Effects

- Without rank invariance we can identify quantile treatment effects for a particular group under alternative assumptions
- Focus on binary instrument case

$$Z \in \{0, 1\}, \text{ with } 0 < \Pr(Z = 1) < 1$$

- Assumption LQTE:
  (1) Independence: $D = \Pi(Z, V)$ and $(Y_d, V) \perp\!\!\!\perp Z$, $d \in \{0, 1\}$;
  (2) Relevance: $\Pr(\Pi(1, V) > \Pi(0, V)) > 0$;
  (3) Monotonicity: $\Pi(1, V) \geq \Pi(0, V)$
- Let the local distribution structural function be

$$G_c(y, d) = \Pr[Y_d \leq y \mid \Pi(1, V) > \Pi(0, V)]$$

- The local quantile treatment effect is

$$LQTE(u) = G_c^{-1}(u, 1) - G_c^{-1}(u, 0)$$

- Quantile treatment effect for *compliers*, individuals whose behavior changes with the instrument

LQSF and IV

- Theorem LQSF (Abadie, 02): Under assumption LQTE

$$G_c(y,d) = \frac{Cov(Z, 1\{Y \leq y\}1\{D = d\})}{Cov(Z, 1\{D = d\})}$$

- Proof: Let $d = 1$. Define $T_0 = 0$ and $T_1 = 1(Y_1 \leq y)$. Then

$$T := 1(Y \leq y)1(D = 1) = T_0 + (T_1 - T_0)D$$

- By independence

$$(T_0, T_1, V) = [0, 1(Y_1 \leq y), V] \perp\!\!\!\perp Z$$

- By the LATE argument

$$\frac{Cov(Z,T)}{Cov(Z,D)} = \mathbb{E}[T_1 - T_0 \mid \Pi(1,V) > \Pi(0,V)] = G_c(y,1)$$

- Similar argument for $d = 0$ with $T = 1(Y \leq y)1(D = 0)$

LQTE and QR

- Lemma (Abadie, Angrist, and Imbens, 02): Under LQTE

$$(Y_0, Y_1) \perp\!\!\!\perp D \mid \Pi(1, V) > \Pi(0, V)$$

  i.e., random assignment holds for compliers

- QR in the complier group, but we do not observe compliance
- Abadie (03) kappa function κ_u finds compliers

$$\kappa_u = 1 - \frac{D(1-Z)}{1-\mathbb{E}[Z]} - \frac{(1-D)Z}{\mathbb{E}[Z]}$$

- AAI (02) show that $LQTE(u) = \beta(u)$, where

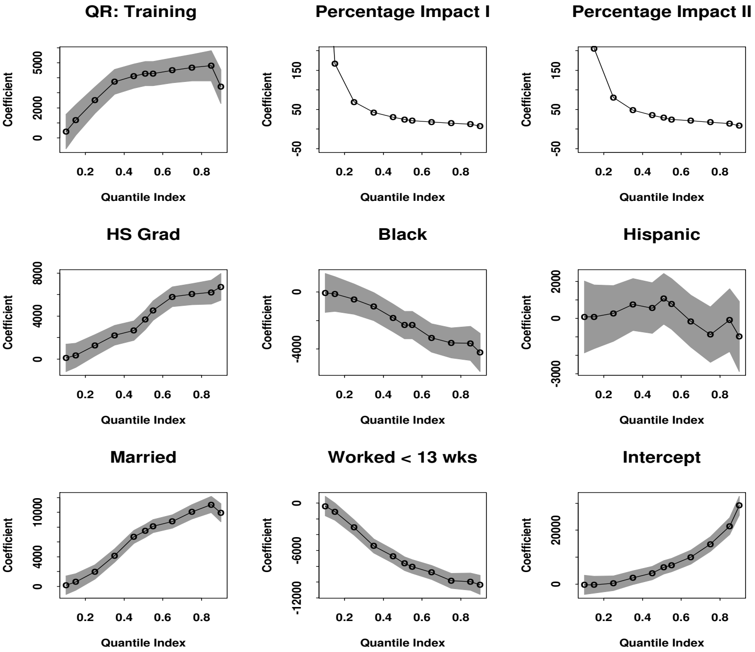
$$[\alpha(u), \beta(u)] = \arg\min_{\alpha,\beta} \mathbb{E}[\kappa_u \rho_u (Y - \alpha - \beta D)]$$

- Need to estimate κ_u; see AAI for implementation details

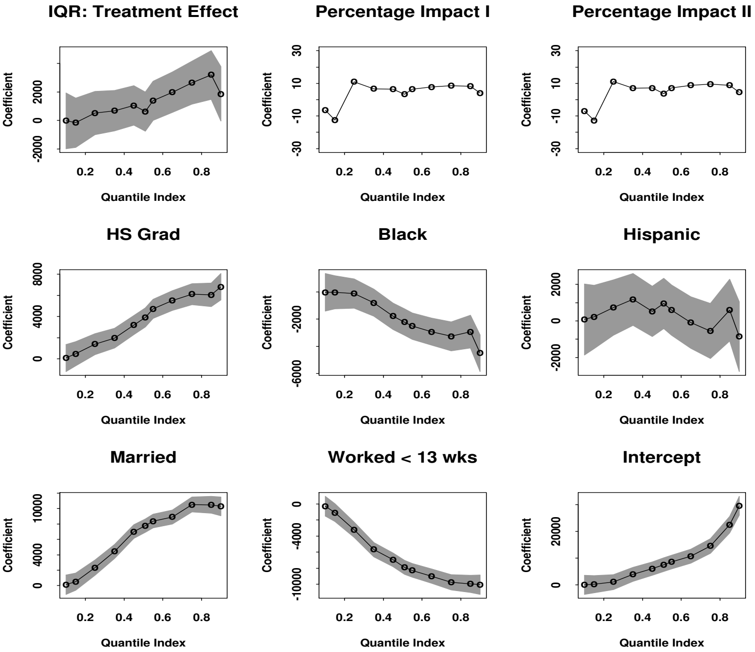
Training Program Application

- JTPA: Job Training Partnership Act
- Discussed in detail in Heckman and Smith (97) and AAI (02)
- Samples based on a randomized experiment with imperfect compliance
- Focus on sample of men with 5,102 observations
- Variables:
  - $Y =$ 30 month Earnings;
  - $D =$ Participation;
  - $Z =$ Random offer of Participation;
  - $C =$ Good Set of 15 Controls (age, education, race, marital status, previous work experience)

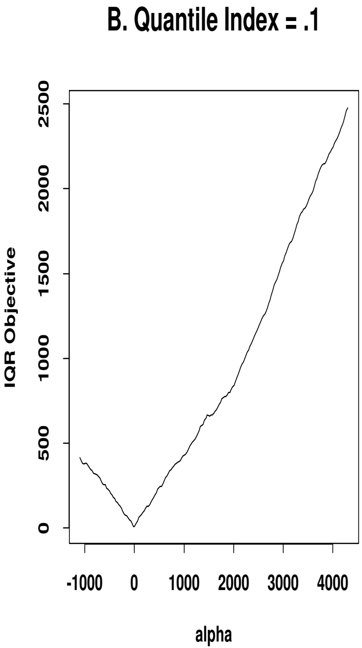
Effect of JTPA: QR Estimates

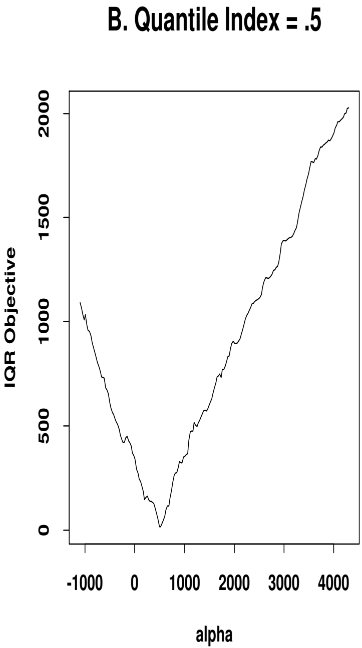


Effect of JTPA: IQR Estimates



IQR Objective Function





Effect of JTPA: OLS and QR Estimates

TABLE II

QUANTILE REGRESSION AND OLS ESTIMATES

Dependent Variable: 30-month Earnings

		Quantile				
	OLS	0.15	0.25	0.50	0.75	0.85
A. Men						
Training	3,754	1,187	2,510	4,420	4,678	4,806
	(536)	(205)	(356)	(651)	(937)	(1,055)
% Impact of Training	21.2	135.6	75.2	34.5	17.2	13.4
High school or GED	4,015	339	1,280	3,665	6,045	6,224
	(571)	(186)	(305)	(618)	(1,029)	(1,170)
Black	−2,354	−134	−500	−2,084	−3,576	−3,609
	(626)	(194)	(324)	(684)	(1,087)	(1,331)
Hispanic	251	91	278	925	−877	−85
	(883)	(315)	(512)	(1,066)	(1,769)	(2,047)
Married	6,546	587	1,964	7,113	10,073	11,062
	(629)	(222)	(427)	(839)	(1,046)	(1,093)
Worked less than 13	−6,582	−1,090	−3,097	−7,610	−9,834	−9,951
weeks in past year	(566)	(190)	(339)	(665)	(1,000)	(1,099)
Constant	9,811	−216	365	6,110	14,874	21,527
	(1,541)	(468)	(765)	(1,403)	(2,134)	(3,896)

Effect of JTPA: LQTE Estimates

TABLE III

QUANTILE TREATMENT EFFECTS AND 2SLS ESTIMATES

Dependent Variable: 30-month Earnings

		Quantile				
	2SLS	0.15	0.25	0.50	0.75	0.85
A. Men						
Training	1,593	121	702	1,544	3,131	3,378
	(895)	(475)	(670)	(1,073)	(1,376)	(1,811)
% Impact of Training	8.55	5.19	12.0	9.64	10.7	9.02
High school or GED	4,075	714	1,752	4,024	5,392	5,954
	(573)	(429)	(644)	(940)	(1,441)	(1,783)
Black	−2,349	−171	−377	−2,656	−4,182	−3,523
	(625)	(439)	(626)	(1,136)	(1,587)	(1,867)
Hispanic	335	328	1,476	1,499	379	1,023
	(888)	(757)	(1,128)	(1,390)	(2,294)	(2,427)
Married	6,647	1,564	3,190	7,683	9,509	10,185
	(627)	(596)	(865)	(1,202)	(1,430)	(1,525)
Worked less than 13	−6,575	−1,932	−4,195	−7,009	−9,289	−9,078
weeks in past year	(567)	(442)	(664)	(1,040)	(1,420)	(1,596)
Constant	10,641	−134	1,049	7,689	14,901	22,412
	(1,569)	(1,116)	(1,655)	(2,361)	(3,292)	(7,655)

Effect of JTPA: IQR Estimates

Table 1
Results from empirical examples

	Example 1: demand for fish			Example 2: returns to training	
	QR (1)	IVQR (2)	IVQR (3)	QR (4)	IVQR (5)
$\hat{\alpha}(0.15)$	−0.53	−1.5	−1.5	1188	−200
Wald CI	(−1.24,0.16)	(−3.69,−0.69)	(−2.51,−0.49)	(553,1822)	(−1435,1035)
Dual CI		[−5.0,0.5)	(−3.2,0.1)		(−1300,1500)
$\hat{\alpha}(0.25)$	−0.40	−1.0	−1.4	2510	500
Wald CI	(−0.87,0.07)	(−2.51,0.51)	(−2.52,−0.28)	(1742,3278)	(−887,1887)
Dual CI		(−4.4,0.0)	(−3.1,0.1)		(−1000,2000)
$\hat{\alpha}(0.50)$	−0.41	−0.7	−0.9	4420	300
Wald CI	(−0.81,−0.01)	(−1.67,0.27)	(−1.82,0.02)	(3220,5621)	(−1589,2189)
Dual CI		(−3.0,0.6)	(−3.0,0.6)		(−1400,2700)
$\hat{\alpha}(0.75)$	−0.70	−1.2	−1.3	4678	2700
Wald CI	(−1.18,−0.22)	(−2.02,−0.38)	(−2.07,−0.53)	(2901,6455)	(−260,5660)
Dual CI		(−2.0,−0.1)	(−2.1,0.1)		(−400,5600)
$\hat{\alpha}(0.85)$	−0.81	−1.3	−1.1	4806	3200
Wald CI	(−1.24,−0.38)	(−2.10,−0.50)	(−1.82,−0.38)	(2751,6861)	(32, 6368)
Dual CI		(−2.0,5.0]	(−2.6,5.0]		(500,5800)

Notes: Columns (1)–(3) report results from estimation of the demand for fish, and columns (4) and (5) report results from estimation of the returns to training from the JTPA experiment. Columns (1) and (4) report ordinary QR results, and columns (2), (3), and (5) report instrumental QR results. In column (2), one instrument, *Stormy*, is used, and in column (3), two instruments, *Stormy* and *Mixed* are used. Rows labeled $\alpha(\tau)$ for $\tau \in \{0.15, 0.25, 0.50, 0.75, 0.85\}$ report point estimates, and the numbers in parentheses are confidence intervals.

Discussion

- IVQR and LQTE produce similar results in this application
- Wuthrich (18) shows that under the LQTE conditions IVQR estimates LQTE but at a shifted quantile index

$$g(1,u) - g(0,u) = LQTE(G_c(g(1,u),1)) = LQTE(G_c(g(0,u),0))$$

- There are 3 populations under LQTE:
  1. Compliers: $\Pi(0,V) = 0$, $\Pi(1,V) = 1$,
  2. Always takers: $\Pi(0,V) = \Pi(1,V) = 1$,
  3. Never takers: $\Pi(0,V) = \Pi(1,V) = 0$
- Distribution of Y_0 identified for compliers and never takers, and distribution of Y_1 identified for compliers and always takers
- Without rank invariance, IVQR uses the compliers to extrapolate the distributions of potential outcomes that are not identified