Quantile Regression

Gabriel V. Montes-Rojas

Economic motivation: Effect of training on wages

- Consider the study of the effect of subsidized training on wages

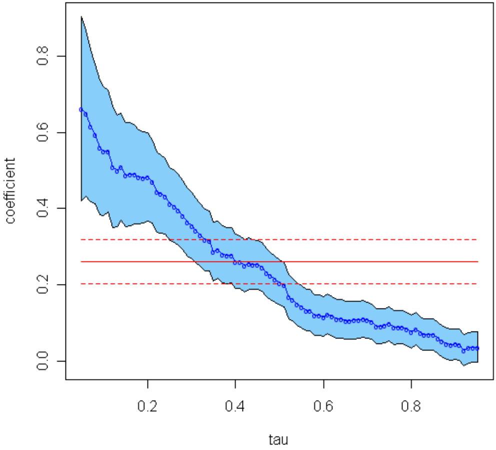
$$y = d\alpha + \boldsymbol{x}\beta + u$$

- y: wages
  d: training treatment indicator (dummy variable: 1 if received training, 0 otherwise)
  $\boldsymbol{x}$: other covariates (age, education, marital status, race, etc.)
  u: unobservables (ability, predisposition to work)
- A policy maker is interested in an estimate of whether training increases wages or not, i.e. $\alpha > 0$.
- The effect of policy variables on distributional outcomes are of fundamental interest in empirical economics. Of particular interest is the estimation of the quantile treatment effects (QTE), that is, the effect of some policy variable of interest on the different quantiles of a conditional response variable.

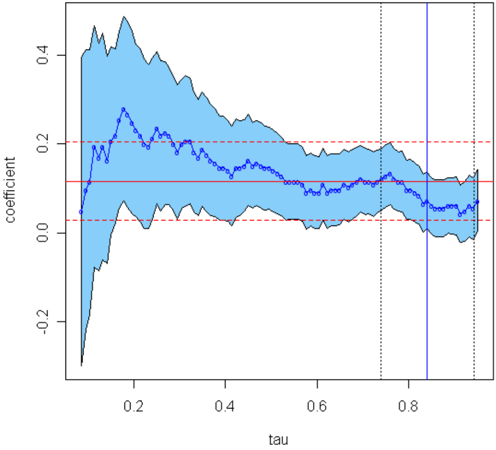
Economic motivation: Effect of training on wages

- OLS will tell you: what is the **conditional mean** effect of training?
  $\alpha = E[y|d = 1, \mathbf{x}] - E[y|d = 0, \mathbf{x}]$
- However, there may be a lot of **heterogeneity** in the responses of wages to training. Some people may benefit more than others.
- Policymakers may be more interested in seen if (conditional) disadvantaged groups are benefited more than privileged groups. That is, the training policy may be directed towards under-performers.
- What if you want to focus on a **subgroup** of the population, say those that will be ranked on the **conditional** lower 10%? $(\tau = .1)$
  $\alpha(.1) = Q_{.1}[y|d = 1, \mathbf{x}] - Q_{.1}[y|d = 0, \mathbf{x}]$
- What if you want to focus on a **subgroup** of the population, say those that will be ranked on the **conditional** upper 90%? $(\tau = .9)$
  $\alpha(.9) = Q_{.9}[y|d = 1, \mathbf{x}] - Q_{.9}[y|d = 0, \mathbf{x}]$
- What if you want to focus on a **subgroup** of the population, say those that will be ranked on the **conditional** median? $(\tau = .5)$
  $\alpha(.5) = Q_{.5}[y|d = 1, \mathbf{x}] - Q_{.5}[y|d = 0, \mathbf{x}]$

QR vs. OLS



QR vs. OLS: IV estimation



Motivation

- Consider a random variable y with $E(y) = \mu_y$, $Var(y) = \sigma^2 < \infty$, with cumulative distribution function (cdf) F_y, and a random sample $\{y_i\}_{i=1}^n$.
- The expected value is the solution to the minimization if the expected squared deviation, i.e., $E(y) = arg \min_c E(y - c)^2$.
- Then using the analogue principle

$$\hat{\mu}_y \equiv \frac{1}{N}\sum_{i=1}^{N} y_i = arg \min_c \sum_{i=1}^{N}(y_i - c)^2$$

Motivation

- The median is an **order statistic**, which tells you the number η_y where 50% of the observations are above and 50% are below. More formally η_y is any number such that $P[y \leq \eta_y] \geq 1/2$ and $P[y \geq \eta_y] \leq 1/2$.
- If the cumulative distribution function (*cdf*) of y, F_y, is strictly increasing, i.e., continuously distributed with positive density, $f_y > 0$, then $\eta_y = F_y^{-1}(1/2)$.
- The median is also a solution to the minimization of the absolute deviation, i.e. $\eta_y = arg \min_c E|y - c|$.

  Proof: (Asuming that the *cdf* is continuous.)
  $E|y-c| = E\left(1[y>c](y-c) - 1[y<c](y-c) + 1[y=c](y-c)\right)$. Taking directional derivatives, $\partial E|y-c|/\partial c = -E(1[y>c]) + E(1[y<c]) - E(1[y=c]) = -P[y>c] + P[y<c]$. In this case from the first order condition $-E(sgn(y-c)) = 0$ where $sgn(.)$ is the sign function $sgn(u) = 1 - 2 \cdot 1[u<0]$.
  Then using the analogue principle

$$\hat{\eta}_y = arg \min_c \sum_{i=1}^{N} |y_i - c|$$

Order statistics

Define $Q_\tau(y) = inf\{\xi : F_y(\xi) \geq \tau\}$ *for* $\tau \in [0,1]$ *where* $F_y(.)$ *is the cdf of* y*, as the* τ-**quantile** *or percentile of* y*.*

- ...if you want (10-90) then you have the 10-th percentile, $\tau = .1 \rightarrow Q_{.1}(y)$
- ...if you want (25-75) then you have the 25-th percentile (also called first quartile), $\tau = .25 \rightarrow Q_{.25}(y)$
- ...if you want (50-50) then you have the 50-th percentile (also called the median), $\tau = .5 \rightarrow Q_{.5}(y)$
- ...if you want (75-25) then you have the 75-th percentile (also called third quartile), $\tau = .75 \rightarrow Q_{.75}(y)$
- ...if you want (90-10) then you have the 90-th percentile, $\tau = .9 \rightarrow Q_{.9}(y)$

OLS estimator

- Consider now a linear structural equation $y = \boldsymbol{x}\boldsymbol{\beta} + u$ with $E(u|\boldsymbol{x}) = 0$, $Var(u|\boldsymbol{x}) = \sigma^2 < \infty$, with cumulative distribution function (cdf) F_U, and a random sample $\{y_i, \boldsymbol{x}_i\}_{i=1}^n$. Now $\boldsymbol{x}$ are the K covariates. A generalization to the simpler problem above is to consider the analysis of the conditional case.
- The conditional expectation is the solution to the minimization if the conditional expected squared deviation, i.e. $E(y|\boldsymbol{x}) = arg\min_{m(\boldsymbol{x})} E(y - m(\boldsymbol{x}))^2$.
- Note that in this case, if $\boldsymbol{x} = \boldsymbol{x}\boldsymbol{\beta}$, i.e., a linear function of the parameters,

$$E(y|\boldsymbol{x}) = \boldsymbol{x}\boldsymbol{\beta},\ then\ \beta = \frac{\partial E(y|\boldsymbol{x})}{\partial \boldsymbol{x}},$$

that is, the regression coefficients are the effect of a marginal change in $\boldsymbol{x}$ on the conditional expectation of y.

- Then using the analogue principle

$$\hat{\boldsymbol{\beta}} = arg\min_{\boldsymbol{b}} \sum_{i=1}^{N} (y_i - \boldsymbol{x}_i\boldsymbol{b})^2$$

This is the well-known ordinary least squares (OLS) estimator.

LAD estimator

- The conditional median is the solution to the minimization if the conditional expected absolute deviation, i.e. $Q_{.5}(y|\boldsymbol{x}) = arg\min_{\boldsymbol{b}} E|y - \boldsymbol{x}\boldsymbol{b}|$.
- We can also write the model as

$$y_i = \boldsymbol{x}_i\beta(.5) + u_i, i = 1, 2, ..., N$$

where

$$Q_{.5}(y|\boldsymbol{x}) = \boldsymbol{x}\beta(.5)$$

or equivalently

$$Q_{.5}(u|\boldsymbol{x}) = 0.$$

Then

$$\beta(.5) = \frac{\partial Q_{.5}(y|\boldsymbol{x})}{\partial \boldsymbol{x}},$$

that is, the median regression coefficients are the effect of a marginal change in $\boldsymbol{x}$ on the conditional median of y.
- Then using the analogue principle

$$\hat{\beta}(.5) = arg\min_{\boldsymbol{b}} \sum_{i=1}^{N} |y_i - \boldsymbol{x}_i\boldsymbol{b}|$$

This is the well-known **least absolute deviation** (LAD) estimator.

Quantile regression

- More generally, for any quantile $\tau \in (0,1)$ of interest, the conditional median is the solution to the minimization if the conditional expectation of a check function $\rho_\tau(.)$ where $\rho_\tau(u) = u \cdot (\tau - 1[u < 0])$, i.e. $Q_\tau(y|\boldsymbol{x}) = arg\min_{\boldsymbol{b}} \rho_\tau(y - \boldsymbol{x}\boldsymbol{b})$.
Note that $\rho_\tau(.)$ is an asymmetric function such that

$$\rho_\tau(u) = \begin{cases} \tau u & if \;\; u \geq 0 \\ (\tau - 1)u & if \;\; u < 0 \end{cases}$$

- We can also write the model as

$$y_i = \boldsymbol{x}_i \boldsymbol{\beta}(\tau) + u_i$$

where

$$Q_\tau(y|\boldsymbol{x}) = \boldsymbol{x}\boldsymbol{\beta}(\tau)$$

or equivalently

$$Q_\tau(u|\boldsymbol{x}) = 0.$$

Then

$$\boldsymbol{\beta}(\tau) = \frac{\partial Q_\tau(y|\boldsymbol{x})}{\partial \boldsymbol{x}},$$

that is, the τ-quantile regression coefficients are the effect of a marginal change in $\boldsymbol{x}$ on the conditional τ-quantile of y.

Quantile regression

Then using the analogue principle

$$\hat{\beta}(\tau) = arg \min_{\boldsymbol{b}} \sum_{i=1}^{N} \rho_\tau(y_i - \boldsymbol{x}_i \boldsymbol{b})$$

This is quantile regression (QR) estimator. Note that if $\tau = 0.5$ this is the LAD estimator. The seminal article on this topic is Koenker and Basset (1978, Econometrica).

The first order condition becomes

$$\frac{1}{N} \sum_{i=1}^{N} (\tau - 1(y_i < \boldsymbol{x}_i \boldsymbol{b}) \boldsymbol{x}_i) = \sum_{i=1}^{N} \psi(\tau, \boldsymbol{b}; y_i, \boldsymbol{x}_i) = \boldsymbol{0}_k$$

where $\boldsymbol{s}(\tau, \boldsymbol{b}; y, \boldsymbol{x}) = (\tau - 1(y < \boldsymbol{x} \boldsymbol{b}))\, \boldsymbol{x}$. Note that ρ_τ is not differentiable but has directional derivative.

Location-scale model

Consider a location-scale model

$$y = \mathbf{x}^\top \gamma + (\mathbf{x}^\top \boldsymbol{\alpha})\epsilon \;\; with \;\; \epsilon \sim F_\epsilon \;\; , \;\; \epsilon \perp\!\!\!\perp \mathbf{x}.$$

For this model,

$$\frac{\partial Q_\tau(y|\mathbf{x})}{\partial \mathbf{x}} = \boldsymbol{\beta}(\tau) = \gamma + \boldsymbol{\alpha} Q_\tau(\epsilon),$$

where $Q_\tau(\epsilon)$ corresponds to the τ-quantile of ϵ, while the conditional mean

$$\frac{\partial E(y|\mathbf{x})}{\partial \mathbf{x}} = \gamma \;\; (i.e.\ constant)$$

Note that in order to have heterogeneity in the QR estimators we require **heteroskedasticity**. In general in QR, observations are assumed independent across *i* but not necessarily identically distributed (allowing heteroskedasticity).

A likelihood approach

Bera, A., Galvao, A. Montes-Rojas, G. & Park, S-Y. (2016). "Asymmetric Laplace regression: Maximum likelihood, maximum entropy and quantile regression," Journal of Econometric Methods, 5(1), 79-101.

- The conditional mean/OLS approach is based on the normal probability density:

$$f(y;\mu,\sigma)=\frac{1}{\sqrt{2\pi\sigma}}\exp\left(-\frac{(y-\mu)^2}{\sigma^2}\right),$$

assuming $y\sim N(\mu,\sigma^2)$.

- The QR model is based on the asymmetric Laplace probability density (ALPD):

$$f(y;\mu,\tau,\sigma)=\frac{\tau(1-\tau)}{\sigma}\exp\left(-\frac{\rho_\tau(y-\mu)}{\sigma}\right)$$

for given (τ,σ). The well-known symmetric Laplace (double exponential) distribution is a special case when τ=1/2.

Asymptotics

- Quantile regression models deviates from the M-estimators theory because the objective function $\rho_\tau(.)$ is not twice differentiable.
- Rewrite the objective function in the M-estimator environment

$$\rho_\tau(y - \boldsymbol{x}\boldsymbol{\beta}) = q(\boldsymbol{w}_i, \boldsymbol{\beta}) = \tau 1[y_i - \boldsymbol{x}_i\boldsymbol{\beta} \geq 0](y_i - \boldsymbol{x}_i\boldsymbol{\beta}) - (1-\tau)1[y_i - \boldsymbol{x}_i\boldsymbol{\beta} < 0](y_i - \boldsymbol{x}_i\boldsymbol{\beta}).$$

- And the score function as

$$\boldsymbol{s}_i(\boldsymbol{\beta}) = -\boldsymbol{x}_i'\{\tau 1[y_i - \boldsymbol{x}_i\boldsymbol{\beta} \geq 0] - (1-\tau)1[y_i - \boldsymbol{x}_i\boldsymbol{\beta} < 0]\}.$$

- Note that if u_i has a continuous distribution at zero, $E[\boldsymbol{s}_i(\boldsymbol{\beta}_0)|\boldsymbol{x}_i] = \boldsymbol{0}$ because $E(1[y_i - \boldsymbol{x}_i\boldsymbol{\beta}_0 \geq 0]|\boldsymbol{x}_i) = P(1[y_i - \boldsymbol{x}_i\boldsymbol{\beta}_0 \geq 0]|\boldsymbol{x}_i) = (1-\tau)$ and . $E(1[y_i - \boldsymbol{x}_i\boldsymbol{\beta}_0 < 0]|\boldsymbol{x}_i) = P(1[y_i - \boldsymbol{x}_i\boldsymbol{\beta}_0 < 0]|\boldsymbol{x}_i) = \tau$.
- Moreover the solutions are not exact zeros but satisfy

$$N^{-1/2}\sum_{i=1}^{N}\boldsymbol{s}_i(\hat{\boldsymbol{\beta}}) = o_p(1)$$

Asymptotics

- Assuming that the conditional cdf of u $F_u(.|\boldsymbol{x})$ is continuously differentiable at 0 with density $f_u(.|\boldsymbol{x}) > 0$,

$$
\begin{aligned}
E[\boldsymbol{s}_i(\boldsymbol{\beta}_0)|\boldsymbol{x}_i] &= -\boldsymbol{x}_i'\{\tau P[y_i - \boldsymbol{x}_i\boldsymbol{\beta} \geq 0|\boldsymbol{x}_i] - (1-\tau)P[y_i - \boldsymbol{x}_i\boldsymbol{\beta} < 0|\boldsymbol{x}_i]\} \\
&= -\boldsymbol{x}_i'\{\tau P[u_i \geq \boldsymbol{x}_i(\boldsymbol{\beta} - \boldsymbol{\beta}_0)|\boldsymbol{x}_i] - (1-\tau)P[u_i < \boldsymbol{x}_i(\boldsymbol{\beta} - \boldsymbol{\beta}_0)|\boldsymbol{x}_i]\} \\
&= -\boldsymbol{x}_i'\{\tau(1 - F_u[\boldsymbol{x}_i(\boldsymbol{\beta} - \boldsymbol{\beta}_0)|\boldsymbol{x}_i]) - (1-\tau)F_u[\boldsymbol{x}_i(\boldsymbol{\beta} - \boldsymbol{\beta}_0)|\boldsymbol{x}_i]\} \\
&= -\boldsymbol{x}_i'[\tau - F_u(\boldsymbol{x}_i(\boldsymbol{\beta} - \boldsymbol{\beta}_0)|\boldsymbol{x}_i)]
\end{aligned}
$$

Asymptotics

- QR is an M-estimator and it is asymptotically normal:

$$\sqrt{N}(\hat{\beta} - \beta_0) \xrightarrow{d} N(\mathbf{0}, \mathbf{A}_0^{-1}\mathbf{B}_0\mathbf{A}_0^{-1}).$$

- $$\nabla_\beta E[\mathbf{s}(\beta_0)|\mathbf{x}] = f_u(\mathbf{x}\beta_0|\mathbf{x})\mathbf{x}'\mathbf{x}$$

  so that

$$\mathbf{A}_0 = E[f_u(\mathbf{x}\beta_0|\mathbf{x})\mathbf{x}'\mathbf{x}].$$

- Also

$$\mathbf{B}_0 \equiv E[\mathbf{s}(\beta_0)\mathbf{s}(\beta_0)'] = \tau(1-\tau)E[\mathbf{x}'\mathbf{x}].$$

- If we assume u independent of $\mathbf{x}$, then the asymptotic variance simplifies to $\frac{\tau(1-\tau)}{(f_u(F_u^{-1}(\tau)))^2}[E(\mathbf{x}'\mathbf{x})]^{-1}$

"Make sure that unertain parts of the evidence never have overriding influence on the final conclusions." (Huber and Ronchetti, p.21)

QR as robust estimator

- Our interest is to measure the location of the random variable x, either by the mean, $E(x)$ or the median, $Median(x)$.
- Consider a given sample $\{x_i\}_{i=1}^{N} = \{x_1, x_2, ..., x_N\}$. Denote it by F_N.
- Now consider an *outlier* observation x_{N+1}. Denote the new sample by F_{N+1}.
- What is the influence of this particular observation on the estimators $\bar{x}$ (average of x for the mean) and $\hat{\eta}_x$ (sample median for the median of x)?

QR as robust estimator

- For the sample average, $\bar{x}$ define $\bar{x}_{F_N}$ and $\bar{x}_{F_{N+1}}$ as the sample average for the two samples. Note that

$$\bar{x}_{F_{N+1}} - \bar{x}_{F_N} = \frac{1}{N+1}(\sum_{i=1}^{N} x_i + x_{N+1}) - \frac{1}{N}(\sum_{i=1}^{N} x_i) = \frac{1}{N+1}(x_{N+1} - \bar{x}_{F_N})$$

  Thus the effect of the outlier is given by the difference between the new observation and the original sample average.

- For the sample median, order the samples as $\{x_{(1)}, x_{(2)}, ..., x_{(N)}\}$ and $\{x_{(1)}, x_{(2)}, ..., x_{(N+1)}\}$ such that $x_{(i)} \leq x_{(j)}$ for $i < j$. Then

$$\hat{\eta}_{F_{N+1}} - \hat{\eta}_{F_N} = x_{(\frac{N+1}{2})} - x_{(\frac{N}{2})}$$

QR as robust estimator

- Consider the unidimensional model above with *cdf* F. Consider a perturbation/outlier/unusual observation in the sample of mass ϵ at the value y, and the corresponding *cdf*: $F_\epsilon = \epsilon\delta_y + (1-\epsilon)F$.
- The influence function for an statistic $\hat{\beta}(F)$ is

$$IF_{\hat{\beta}}(y, F) = \lim_{\epsilon \to 0} \frac{\hat{\beta}(F) - \hat{\beta}(F_\epsilon)}{\epsilon}$$

- For the mean,

$$\hat{\mu}(F_\epsilon) = \int y dF_\epsilon = \epsilon y + (1-\epsilon)\hat{\mu}(F)$$

$$IF_{\hat{\mu}}(y, F) = y - \hat{\mu}(F)$$

QR as robust estimator

- For the quantile estimator,

$$\hat{\eta}_\tau(F_\epsilon) = F_\epsilon^{-1}(\tau)$$

$$IF_{\hat{\eta}_\tau}(y, F) = \frac{sgn(y - F^{-1}(\tau))}{f(F^{-1}(\tau))}$$

- There is a dramatic difference between the two influence functions. For the mean, the influence depends on y, implying that an "outlier" can take the estimator far away from the mean. However, for the quantile function the influence essentially depends on $1/f(F^{-1}(\tau))$ which is called the *sparsity* at a particular quantile.

QR as robust estimator

- For the OLS estimator,

$$IF_{\hat{\beta}}((y, \mathbf{x}), F) = (\mathbf{x}'\mathbf{x})^{-1}\mathbf{x}(y - \mathbf{x}\hat{\beta})$$

- For the quantile regression estimator,

$$IF_{\hat{\beta}(\tau)}((y, \mathbf{x}), F) = Q^{-1}\ \mathbf{x}\ sgn(y - \mathbf{x}\hat{\beta}(\tau))$$

where $Q = \int(\mathbf{x}'\mathbf{x})f(\mathbf{x}\hat{\beta}(\tau))dG(x)$ and $G(.)$ is the cdf of $\mathbf{x}$.

The problem of endogeneity

- Consider a model now with an endogenous variable where

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + ... + \beta_K x_{Ki} + u_i, i = 1, 2, ..., N,$$

  where x_1 is endogenous (i.e. $cov(x_1, u) \neq 0$) and $(x_2, ..., x_K)$ exogenous variables (i.e. $cov(x_j, u) = 0, j = 2, 3, ..., K$).
- Moreover consider an instrumental variable z, with the usual assumptions $cov(z, u) = 0, cov(z, x_1) = 0$.

IVQREG estimator

Chernozhukov and Hansen (2004,2006) instrumental variables estimator for quantile regression is implemented as

$$\hat{\beta}_1(\tau) = argmin_{\beta_1} \|\hat{\gamma}(\beta_1, \tau)\|_A,$$

where $\hat{\gamma}(\beta_1, \tau)$ *is obtained from*

$$argmin_{\beta_0, \beta_2, \ldots, \beta_K, \gamma} \sum_{i=1}^{N} \rho_\tau \left(y_i - (\beta_0 + \beta_2 x_{2i} + \ldots + \beta_K x_{Ki} + \gamma z_i)\right),$$

with $\rho_\tau(.)$ *the* τ*-quantile regression check function,* $\|x\|_A = \sqrt{x' A x}$ *and* A *is a positive definite matrix.*

Note: This estimator may seem unintuitive. However, Galvao and Montes-Rojas (2016) show that if $\rho(u) = u^2$*, i.e. we have an OLS estimator, then* $\hat{\beta} = \hat{\beta}_{2SLS}$*, i.e. the 2SLS IV estimator.*

IVQREG estimator, asymptotic distribution

$$\sqrt{n}(\hat{\boldsymbol{\beta}}(\tau) - \boldsymbol{\beta}(\tau)) \xrightarrow{d} N(\mathbf{0}, J(\tau)^{-1} S(\tau) J(\tau)^{-1})'$$

where $\boldsymbol{\beta} = (\beta_1, \beta_0, \beta_2, ..., \beta_K)$, $J(\tau) = E[f_{\epsilon(\tau)}(0|\boldsymbol{x}, z)(\boldsymbol{x}, z)(\boldsymbol{x}, z)']$ *with* $\epsilon(\tau) = y_i - (\boldsymbol{\beta x} + \gamma z_i)$, $f_{\epsilon(\tau)}(.)$ *its density function, and* $S(\tau) = (\min(\tau, \tau') - \tau\tau')E[(\boldsymbol{x}, z)(\boldsymbol{x}, z)']$.

QR in STATA

- OLS: `reg y x1 x2`
- Median regression: `qreg y x1 x2, q(50)`
- Quantile regression, $\tau = .1$: `qreg y x1 x2, q(10)`
- Quantile regression, $\tau = .9$: `qreg y x1 x2, q(90)`
- To make graph with the quantile process $(\tau, \beta(\tau)), \tau \in (0, 1)$: `gen beta1s=.`

```
gen beta1ols=.
reg y x1 x2
replace beta1ols=_b[x1]
gen tau=.
forvalues tau = 1(1)100 {
qreg y x1 x2, q(‘tau’)
replace beta1s=_b[x1] in ‘tau’
replace tau=‘tau’ in ‘tau’
}
line beta1s beta1ols tau
```

- See also the external command `grqreg`.

QR in STATA

- One important aspect of the quantile regression estimators is that it requires an estimation of the density function.
- Bootstrap methods are useful: `bsqreg y x1 x2`
- Simultaneous quantile regression. Many quantile can be estimated at the same time: `sqreg y x1 x2, q(10 25 50 75 90)`

QR in STATA

- In order to implement the ivqreg estimator follow the command in `https://www.msu.edu/~kwakdo/ivqreg.pdf`
- You need to install it in your PC first.
- Ex: `ivqreg y x2 x3 (x1=z), q(10)`

QR in STATA

- For QTE (quantile treatment effects):
  `http://www.stata-journal.com/sjpdf.html?articlenum=st0203`
- You need to install it first...
- `ivqte` Y `(`X`)` `(`D`)`
- `ivqte` Y `(`X`)` `(`$D = Z$`)`
- where
  - Y: independent variable.
  - X: control variable.
  - D: endogenous/treatment varible.
  - Z: instrumental variable.

References

This slides are based on

- *Chapter 12 of Wooldridge.*
- *Chernozhukov, V. and Hansen, C. (2004) The Impact of 401(k) Participation on the Wealth Distribution: An Instrumental Quantile Regression Analysis, Review of Economics and Statistics 86(3), 735-751.* *http://www.mit.edu/~vchern/papers/ch_401k.pdf*
- *Chernozhukov, V. and Hansen, C. (2006) An IV Model of Quantile Treatment Effects, Econometrica 73(1), pp. 245-261.* *http://www.mit.edu/~vchern/papers/ch_iqr_ema.pdf*
- *Firpo, S. (2007) Efficient Semiparametric Estimation of Quantile Treatment Effects, Econometrica 75(1), 259-276.* *http://epge.fgv.br/files/1125.pdf*
- *Frolich, M., and Melly, B. (2010) Estimation of Quantile Treatment Effects with Stata, Stata Journal, 10, 423-457.* *https://www.stata-journal.com/sjpdf.html?articlenum=st0203*
- *Koenker, R. (2005), Quantile Regression. Cambridge: Cambridge University Press.*
- *Koenker, R. and Hallock, K. (2001) "Quantile Regression," Journal of Economic Perspectives, 15(4), 143–156.*