Topics in Applied Econometrics for Public Policy

Master in Economics of Public Policy, BSE

Handout 5: Introduction to Quantile Regression

Laura Mayoral

IAE and BSE

Barcelona, Spring 2023

1. Introduction

- So far in this course (and probably in all your previous metrics courses): interest in conditional expectation

- This makes sense BUT

- we can be interested in other characteristics of the distribution of the outcome variable

- For example: distribution of income

- We can be interested in the drivers of income per capita; (conditional expectation)
- But this is only part of the story!
- Understanding inequality; poverty, etc. involves understanding things that happen away from the center of the distribution

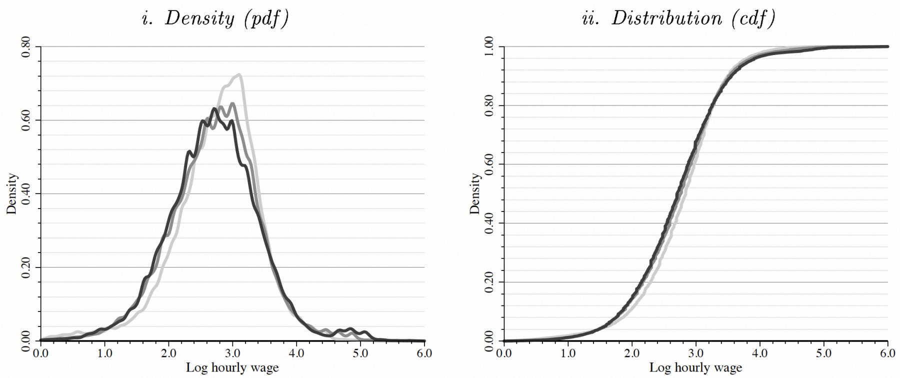
- This handout: focuses on the quantiles of the distribution of Y given X.

- τ_{th} quantile: The τ_{th} quantile of the distribution of Y is the value q_τ for which a fraction of the population has a value of Y smaller than q_τ.

- Conditional τ_{th} quantile: The τ_{th} conditional quantile of the distribution of Y given X=x is the value q_τ for which a fraction of the population for which $X = x$ has a value of Y smaller than q_τ.

- Many interesting research questions
- Example: what drives the inequality increase in the US? i.e., the poor getting poorer and the rich getting richer.

Figure I. – Distribution of U.S. Male Wages (1980-2000)



Note: Light gray: 1980; gray: 1990; dark gray: 2000. Sample restricted to working male aged 16 to 65 who worked at least 20 weeks during the reference year and at least 10 hours per week. Hourly wages are expressed in (log) US$ of year 2000. *Data source*: U.S. Census.

Table 1—Unconditional Quantiles for Wages (1980-2000)

Year	10th	25th	Percentile: 50th	75th	90th
1980	1.96	2.41	2.84	3.18	3.50
1990	1.86	2.30	2.76	3.15	3.51
2000	1.83	2.27	2.70	3.15	3.55

Note: Sample restricted to working male aged 16 to 65 who worked at least 20 weeks during the reference year and at least 10 hours per week. Hourly wages are expressed in (log) US$ of year 2000. *Data source*: U.S. Census.

(Graphs from Joan Llull's course materials)

- Quantile regression has been used in a broad range of application settings, whenever understanding things at the "tails", not at the center of the distribution is key

- In economics: wage determinants, discrimination effects, trends in income inequality and poverty; student performance at the tails, ...

- climate change: we're not only interested in average increases in temperature but also understanding what drives this increase (what places iare heating up quicker, colder or hotter ones (unconditional quantiles), and how covariates affect the increase at different points in the distribution (conditional quantiles)

- behavior, health, ...: whenever we want to understand what drives "extreme" behavior, the tails of the distribution.

- For instance: drivers of low weights in newborns.

■ A quick preview of what's coming: quantile regression, birth weight on covariates and wages as a function of education

(from Mostly harmless Econometrics)

TABLE 7.1.1

Quantile regression coefficients for schooling in the 1980, 1990, and 2000 censuses

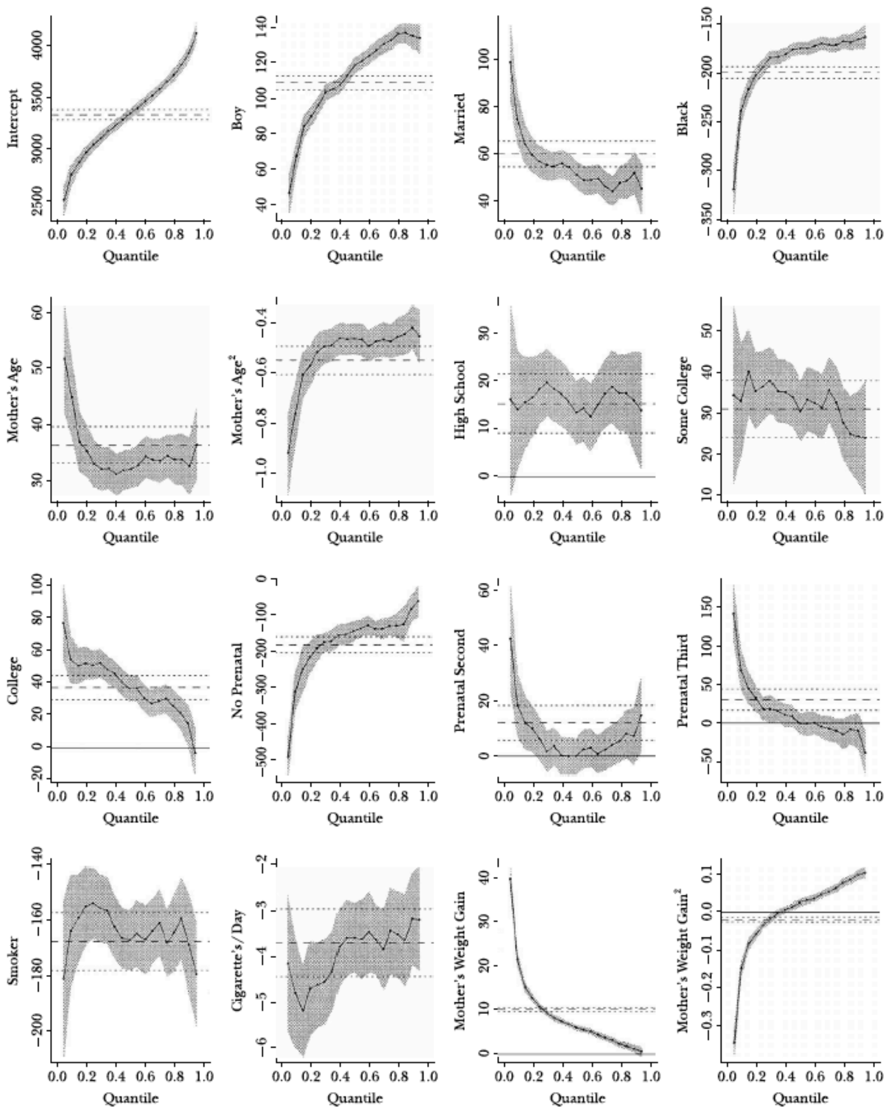
		Desc. Stats.		Quantile Regression Estimates					OLS Estimates	
ensus	Obs.	Mean	SD	0.1	0.25	0.5	0.75	0.9	Coeff.	Root MSE
980	65,023	6.4	.67	.074 (.002)	.074 (.001)	.068 (.001)	.070 (.001)	.079 (.001)	.072 (.001)	.63
990	86,785	6.5	.69	.112 (.003)	.110 (.001)	.106 (.001)	.111 (.001)	.137 (.003)	.114 (.001)	.64
2000	97,397	6.5	.75	.092 (.002)	.105 (.001)	.111 (.001)	.120 (.001)	.157 (.004)	.114 (.001)	.69

Notes: Adapted from Angrist, Chernozhukov, and Fernandez-Val (2006). The table ports quantile regression estimates of the returns to schooling in a model for log wages, ith OLS estimates shown at the right for comparison. The sample includes U.S.-born white nd black men aged 40–49. The sample size and the mean and standard deviation of log ages in each census extract are shown at the left. Standard errors are reported in paren- ieses. All models control for race and potential experience. Sampling weights were used r the 2000 census estimates

(from Koenker and Hallock, 2001)

Figure 4

Ordinary Least Squares and Quantile Regression Estimates for Birthweight Model



Roadmap

1. Preliminaries: Unconditional Quantiles
   1.1. Estimation
   1.2. Standard errors
2. Conditional Quantiles
3. Quantile Regression: Motivation
4. Quantile Regression: Generalization. Examples
5. Censored Quantile Regression

1. Preliminaries: Unconditional Quantiles

- Definition: What is a quantile $q_\tau(Y)$?

- Let $F_Y(y)$ be the cumulative distribution function (cdf) of Y.
- The τth quantile of Y, $q_\tau(Y)$, solves

$$F(q_\tau(Y)) = \tau$$

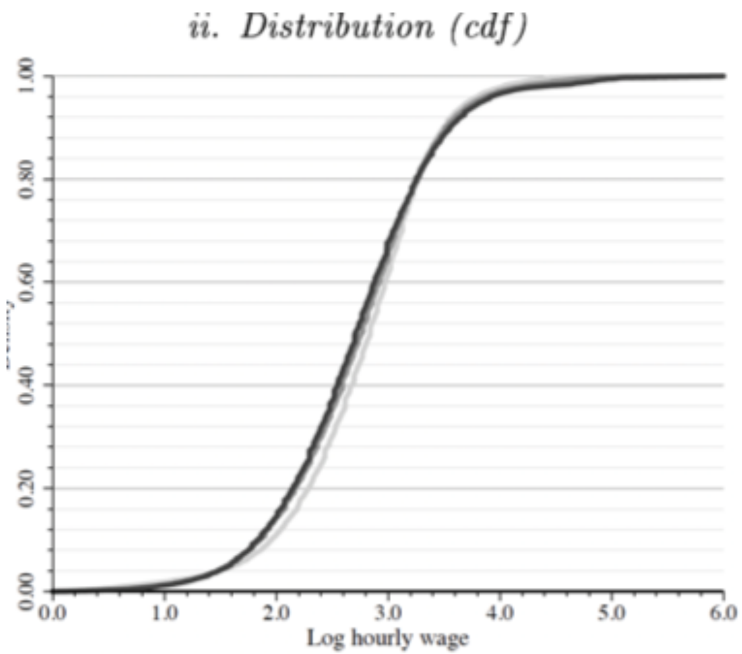
or equivalently:

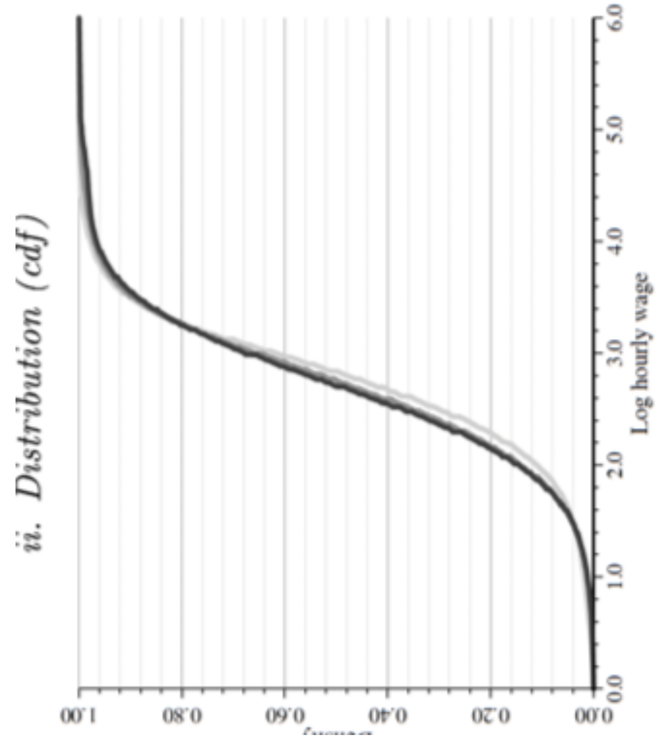
$$q_\tau(Y) = F^{-1}(\tau) = \inf\{r : F(r) \geq \tau\},$$

(this is a generalized inverse function, as F(.) is not strictly increasing for discrete variables)

- The distribution of Y is fully characterized by $\{q_\tau(Y), \tau \in (0,1\}$

- You can see the quantiles just by rotating the CDF!





Another way of writing quantiles: the check function

- Introduce the check function:

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

- It's called like this because it looks like a checkmark:

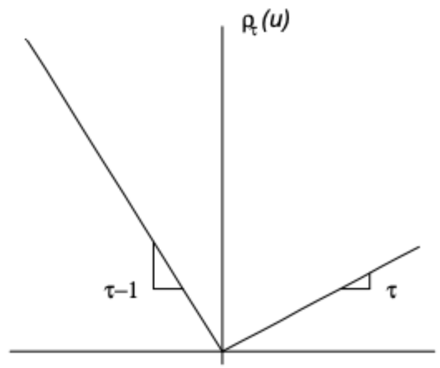


FIGURE 1. Quantile Regression ρ Function

- What is this function doing:

- The check function assigns asymmetric weights to observations larger or smaller than zero, τ and $1-\tau$.

- In one case, weights are symmetric: $\tau = .5$ (median)

- Why is this function useful:

- It can be shown that the quantile τ can be obtained by minimizing the expected value of the check function with respect to ϵ (i.e., q_τ is the value of ϵ that minimizes this function):

$$q_\tau = argmin_\epsilon E(\rho_\tau(Y-\epsilon))$$

- This fact is not immediately trivial, you can find the proof here, pag 3.)

- The proof is not required but it's copied below for your convenience:

$$
\begin{aligned}
E\left(\rho_\tau(X-\xi)\right) &= \int_{-\infty}^{+\infty} \rho_\tau(X-\xi)\ dF(x) \\
&= (\tau-1)\int_{-\infty}^{\xi}(x-\xi)\ dF(x) + \tau\int_{\xi}^{+\infty}(x-\xi)\ dF(x)
\end{aligned}
$$

Differentiating this expectation with respect to ξ,

$$
\begin{aligned}
&= \frac{d}{d\xi}\left[(\tau-1)\int_{-\infty}^{\xi}(x-\xi)\,dF(x) + \tau\int_{\xi}^{+\infty}(x-\xi)\,dF(x)\right] \\
&= \frac{d}{d\xi}\left[(\tau-1)\left(\int_{-\infty}^{\xi}x\,dF(x) - \xi\int_{-\infty}^{\xi}dF(x)\right) - \tau\left(\int_{+\infty}^{\xi}x\,dF(x) - \xi\int_{+\infty}^{\xi}dF(x)\right)\right] \\
&= (\tau-1)\left(\xi f(\xi) - \xi f(\xi) - 1.\int_{-\infty}^{\xi}dF(x)\right) - \tau\left(\xi f(\xi) - \xi f(\xi) - 1.\int_{+\infty}^{\xi}dF(x)\right) \\
&= (\tau-1)(-F(\xi)) - \tau(1-F(\xi)) \\
&= F(\xi) - \tau
\end{aligned}
$$

Some “famous” quantiles

- Median; $\tau = .5$
- First, second and third Quartiles: $\tau = \{0.25, 0.5, 0.75\}$
- Percentiles: $\tau = \{0.01, 0.02, \ldots, 0.99\}$
- Deciles: $\tau = \{0.1, 0.2, \ldots, 0.9\}$

1.1. Estimation of the unconditional Quantiles: Sample quantiles

- Consider a sample $y_1, \ldots, y_N$. We can compute sample quantiles in two ways.

1. Using the empirical cumulative distribution function:

$$\hat{F_Y}(r) = \frac{1}{N}\sum_{i=1}^{N} 1(y_i \leq r),$$

$$\hat{q_\tau}(Y) = \hat{F_Y}^{-1}(\tau) = \inf\{r : \hat{F_Y}(r) \geq \tau\},$$

where $1(\cdot)$ is the indicator function

- computationally very costly as it implies ordering all observations and picking the first observation that leaves at least a fraction τ of the sample below it.

2. Using the *check function*:

■ The sample analogue of q_τ may be found by solving,

$$\hat{q}_\tau(Y) = \underset{\epsilon}{\operatorname{argmin}} \sum_{i=1}^{N} \rho_\tau(y_i - \epsilon) =$$

$$\underset{\epsilon}{\operatorname{argmin}} \sum_{y_i \geq \epsilon}^{N} \tau |y_i - \epsilon| + \sum_{y_i \leq \epsilon}^{N} (1-\tau)|y_i - \epsilon|.$$

where $\operatorname{argmin}_\epsilon$ denotes the value of ϵ that minimizes the sum.

■ The asymmetry of the weights employed in the check function, allows us to pick up the quantiles for different values of τ

1.2. Computing standard errors

- Standard errors can be computed using 1) the asymptotic approximation or 2) bootstrap

1. Asymptotic approximation:

$$\sqrt{N}\left(\hat{q}_{\tau}(Y)-q_{\tau}(Y)\right) \xrightarrow{p} N\left(0, \frac{\tau(1-\tau)}{[f(q_{\tau}(Y))]^{2}}\right) \tag{1}$$

where $f(\cdot)$ is the probability density function of the distribution $F(\cdot)$.

2. Bootstrap: more employed in applications than A.D.

2. Conditional Quantiles

- In econometrics we're typically interested in relating different variables

- This handout is not an exception: we are interested in conditional quantiles: quantiles of the distribution of Y given $X = x$

- Conditional quantile: a measure of location that describes a particular point in the distribution of a response variable Y, given a specific value of one or more predictor variables X.

Specifically, the conditional τth quantile of Y given $X = x$, denoted $q_\tau(Y|\mathrm{X} = x)$, is the smallest value y such that the probability of Y being less than or equal to y, given X=x, is at least τ. In other words, $q_\tau(Y|\mathrm{X} = x)$ is the value of y such that:

$$P(Y \leq y|\mathrm{X} = x) \leq \tau \quad \text{and } P(Y > y|\mathrm{X} = x) \geq 1 - \tau,$$

where $0 \leq \tau \leq 1$ is the quantile level.

Or we can use a slightly different notation: Let $F_{Y|X=x}(y)$ be the conditional cumulative distribution function (cdf) of Y given $X = x$. The τth conditional quantile of Y, q_τ, solves

$$F_{Y|X=x}(q_\tau) = \tau$$

or equivalently:

$$q_\tau(Y|\mathrm{X} = x) = F^{-1}_{Y|X=x}(\tau)$$

Quantile Regression: Motivation

- As mentioned earlier, analysis of the conditional expectation only provides a partial view of the relationship among variables.

- In some applications we might be interested in understanding this relationship at different points in the conditional distribution

- Quantile Regression (QR) is a statistical tool for building such a picture.

- A few reasons why QR can be very useful.

- Provides information on the relationship between Y and X at many points of the distribution of $Y|X$

- QR is robust against outliers; also robust to departures from normality

- QR provides a potentially richer characterization of the data

- It's invariant to monotonic transformations: the quantile of Y is identical to the quantile of g(Y) if g(.) is monotone.

More on motivation: an example

- Koenker and Hallock (JEP, 2001)

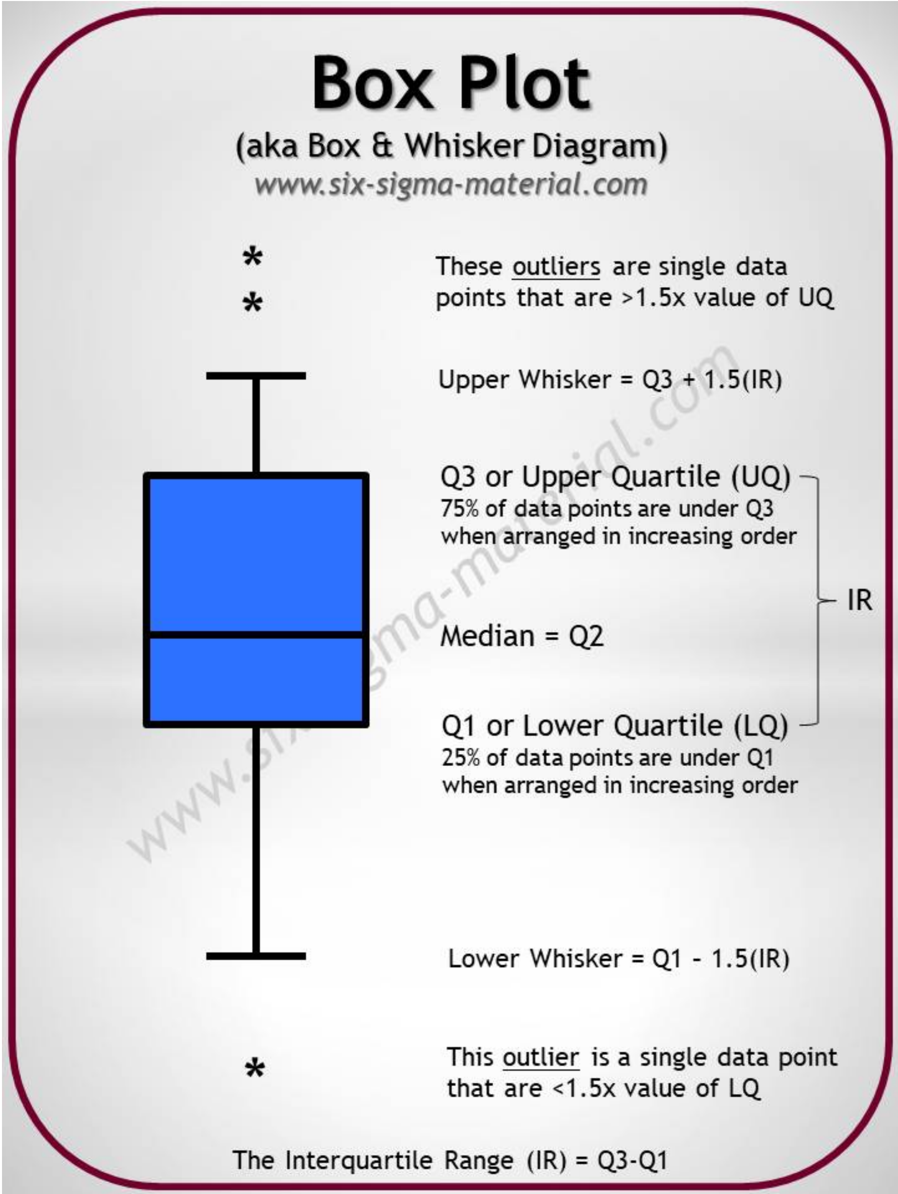
- Variables: Y log of Annual compensation for the chief executive officer (CEO); X firm's market value of equity.

- A sample of 1,660 firms was split into ten groups (deciles) according to their market capitalization (X variable).

- For each group of 166 firms, compute the three quartiles of CEO compensation, the median (middle bar in each rectangle) , the mean (arithmetic mean +, geometric mean *)

- Plot boxplots for each of the deciles

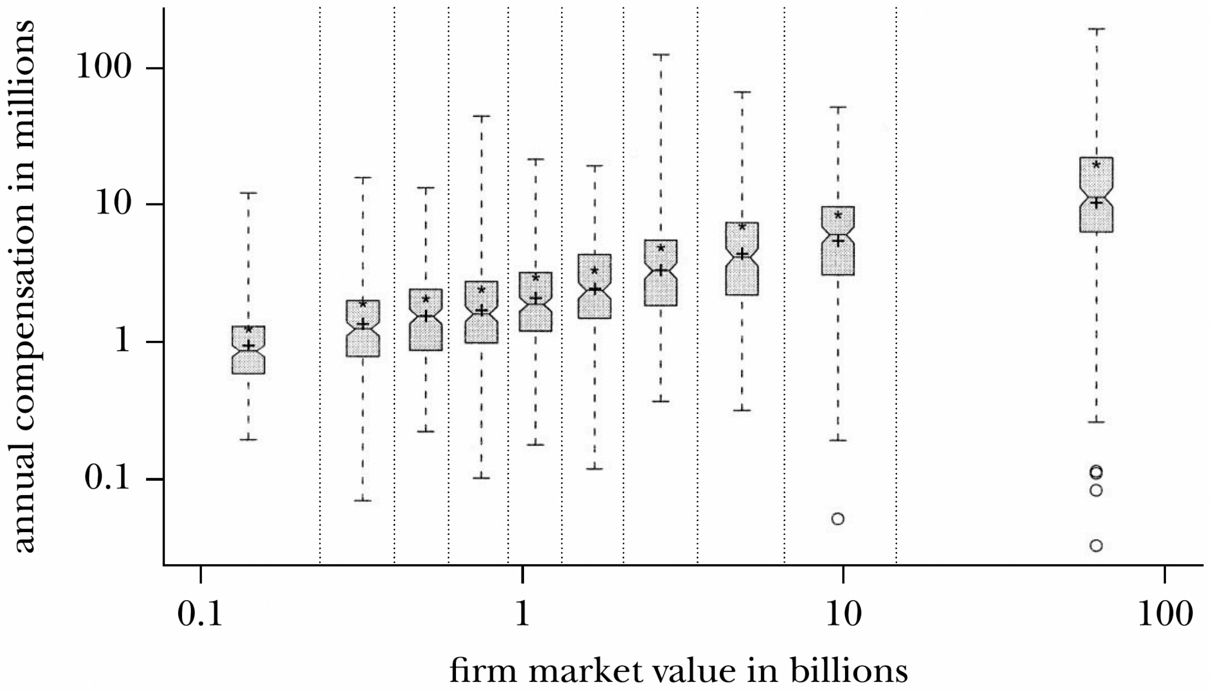
- Note: Recall the information provided by a Box plot



Koenker and Hallock (JEP, 2001)

Figure 1

Pay of Chief Executive Officers by Firm Size



Notes: The boxplots provide a summary of the distribution of CEO annual compensation for ten groupings of firms ranked by market capitalization. The light gray vertical lines demarcate the deciles of the firm size groupings. The upper and lower limits of the boxes represent the first and third quartiles of pay. The median for each group is represented by the horizontal bar in the middle of each box.

Source: Data on CEO annual compensation from EXECUCOMP in 1999.

What do you see in the graph?

- Clear tendency of average/median compensation to go up with firm size

- But other things going on in other aspects of the distribution

  - Even on the log scale, there is a tendency for dispersion, as measured by the interquartile range of log compensation, to increase with firm size.

- By characterizing the entire distribution of annual compensation for each group, the plot provides a much more complete picture than would be offered by simply plotting the group means or medians.

- Now, consider the different estimation methods you know. What do they do?

- Nonparametric estimation of the conditional mean: a flexible function estimating the means at different values of size, as firm sizes grows (i.e, a function that "joins" the $+$'s)

- Parametric OLS: assume a linear function for the $+$'s, estimate the parameters β_0, β_1 of this line

- Quantile regression: will allows us to look at any aspect of the distribution, as firm size grows. We will be able to estimation of conditional quantiles of log compensation as firm size increases.

First example of QR: Median regression

- We can derive the QR estimator in a similar way as in the conditional expectation case.

- Recall an important theorem: If the loss function is the MSE, then the best way of predicting Y using X is $E(Y|X)$, i.e.,

$$E(Y|X) = argmin_{m(x)} E(Y - m(X))^2$$

- This result depends on the loss function employed: MSE

- Consider a different loss function: Absolute-error loss function. What's the best g(X) to predict Y if the loss function is:

$$MAE = E|Y - g(X)|$$

- Answer: conditional median, $g(X) = med(Y|X)$

Parametric assumptions on $E(Y|X)$ and $med(Y|X)$

- In general both $E(Y|X)$ and $med(Y|X)$ are unspecified nonlinear functions

- If we assume that $E(Y|X)$ is linear, then $E(Y|X) = X'\beta$

- We can make a similar assumption in the case of the median, $med(Y|X) = X'\beta$

- If $med(Y|X)$ is linear $\to med(Y|X) = X'\beta$, $\to$ the optimal predictor is

$$\hat{Y} = X'\hat{\beta},$$

where $\hat{\beta}$ is the least absolute-deviations estimator that minimizes $\sum_i |y_i - x_i'\beta|$

Pros/cons of OLS vs. LAD

- Both OLS and LAD look at the evolution of central values of the distribution (averages or medians)

- OLS: non robust (very influenced by outliers). Why? by squaring the residuals, gives more weight to large residuals, that is, outliers in which predicted values are far from actual observations.

- LAD: robust to outliers. Why? LAD gives equal emphasis to all observations

- OLS: unique, stable, close-form solution.

- LAD: no closed-form solution (lack of differentiability of the objective function, no analytical method to optimize the function), unstable solution, possibly not unique solution

Quantile regression: generalization

- The (parametric) quantile regression model:

$$q_\tau(Y|X) = X'\beta_\tau$$

- Meaning: the condicional quantile τ is assumed to be a linear function of X

- Notice two features of this model:

  - Linearity: This is a parametric model (but nonparametric extensions are possible)

  - Effects change across quantiles: the vector of coefficients β_τ varies with τ

- The conditional quantile can be obtained (as in the unconditional case) by minimizing the check function:

$$q_\tau(Y|X) = argmin_{g(x)} E(\rho_\tau(Y - g(x)))$$

and if $g(x) = X'\beta_\tau$ then

$$\beta_\tau = argmin_b E(\rho_\tau(Y - X'b))$$

Quantile Regression Estimator

- Recall that

$$\beta_\tau = argmin_b E(\rho_\tau(Y - X'b)) \quad (1)$$

- Quantile regression estimator $\hat{\beta}_\tau$: sample analog of β_τ

$$\hat{\beta}_\tau = \underset{b}{\operatorname{argmin}} \sum_{i=1} \rho_\tau(|Y - X'b|)$$

- Check function is not differentiable, then optimization is not done the usual way (deriving, equating to zero, etc)

- No analytical closed-form solution.

- linear programming methods (simplex) (computationally simple).

Intepretation of QR coefficients

- Estimation of QR models is easy, understanding what's going on is a bit trickier

- Consider this example: effect of a training program on wages. We find that $\hat{\beta}_{.1} = 10\%$

- What's the meaning of this?

- Quantile coefficients tell us about effects on distributions, not on individuals

- Then, it doesn't mean that someone that was poor after the training program will be 10% richer

- It only means that the poor with training are less poor than the poor without training.

- Why is that?

- Imagine that the training program is rank preserving (i.e., the order of the individuals is not altered after the program, they all get richer but keep their relative positions)

- Then, we could give to β_τ the “individual” interpretation

- But in general, we don’t know whether an intervention is rank preserving or not

- In this case, we can only say that the poor (bottom 10%), (whoever they are) are better off

Interpreting coefficients: Marginal effects

- Recall

$$q_\tau(Y|X) = X'\beta_\tau$$

- Marginal effect (changes are infinitesimal)

$$\frac{\partial q_\tau(Y|X)}{\partial X_j} = \beta_{\tau j}$$

- ME is given by the slope coefficient (as in OLS)

- Discrete changes (larger than infinitesimal)
- A bit delicate: when we move x_j we can change the quantile!

- We need to make the assumption that by moving x_j individuals don't change the quantile!

Example 1a

- A simple case: X is a binary variable
- Question: medical expenditures with and without supplementary insurance?
- Data are from Cameron and Trivedi (2009, ch.7).
- Sample: men 65 and older who are in Medicare
  - y = ltotexp = log(total medical expenditure in 2003)
  - N = 2955 after drop 109 with zero expenditure
  - suppins = 1 if have supplementary medical insurance
  - 58% have supplementary insurance
  - may cover pharmaceutical drugs (not covered by Medicare)
  - may cover copays and coinsurance under regular Medicare

- Let's look at the mean of the conditional distributions (suppins=1/0) bysort suppins: sum totexp ltotexp

  - Sample means are substantially higher with supplementary insurance.
  - Standard deviations are higher in levels but not logs.

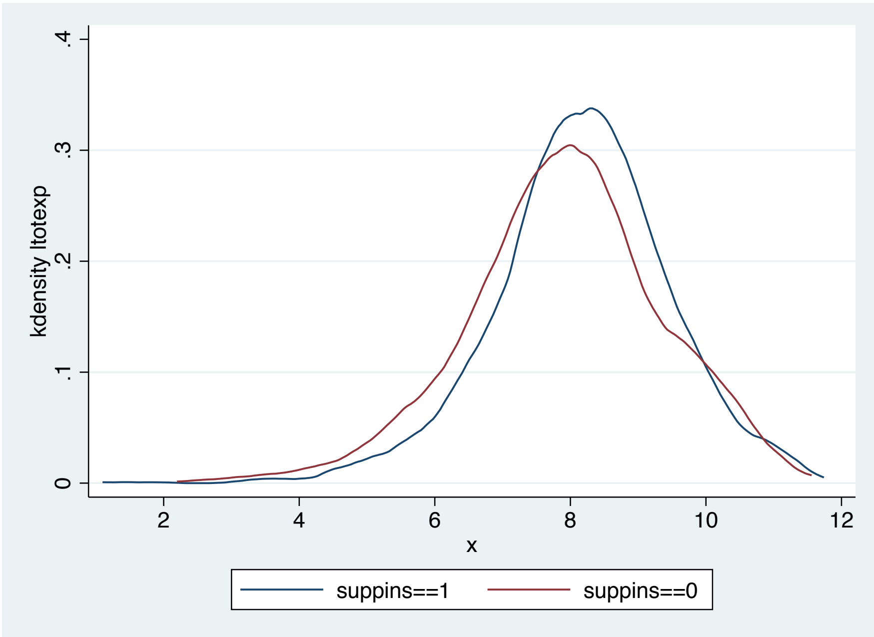
		Suppins = 1	Suppins = 0	
Means	Levels	7470	6420	+16%
	Logs	8.17	7.91	+26%
St.Devs.	Levels	12300	11200	+10%
	Logs	1.30	1.45	−15%

- But, where is the action?

- Let's plot the two conditional densities:

```
graph twoway (kdensity ltotexp if suppins == 1) (kdensity ltotexp if suppins == 0, lstyle(p2)), /// legend(
label(1 "suppins==1") label(2 "suppins==0"))
```

- More action at lower levels of expenditures



- Interpretation: Individuals in the lower quartiles of expenditure have higher medical expenditure than individuals in the lower quartiles with no additional insurance

By how much? compare different percentiles

- Obtain percentiles of the upper curve in previous slide

```
. centile ltotexp if suppins==1, centile(10 50 90)
```

Variable	Obs	Percentile	Centile	[95% Conf. Interval]
ltotexp	1748	10	6.571299	6.457681, 6.673633
		50	8.202071	8.146281, 8.258152
		90	9.771977	9.665329, 9.886245

- Obtain percentiles of the lower curve in previous slide

```
. centile ltotexp if suppins==0, centile(10 50 90)
```

Variable	Obs	Percentile	Centile	[95% Conf. Interval]
ltotexp	1748	10	6.056784	5.880274, 6.27851
		50	7.929846	7.843799, 8.019941
		90	9.796142	9.65716, 9.96981

Variable	Obs	Percentile	Difference
ltotexp	1748	10	6.571299 − 6.056784 = .514515
		50	8.202071 − 7.929846 = .272225
		90	9.771977 − 9.796142 = −.024165

- Now, let's use quantile regression.
- STATA command: QR
- Specify one quantile. Default is .5
- QR, first decile

qreg ltotexp suppins, q(.1)

```
. qreg ltotexp suppins, q(.1)
Iteration  1:  WLS sum of weighted deviations =  1406.8882

Iteration  1: sum of abs. weighted deviations =  1407.3463
Iteration  2: sum of abs. weighted deviations =  1038.0529
Iteration  3: sum of abs. weighted deviations =  758.17886

.1 Quantile regression                              Number of obs =      2,955
  Raw sum of deviations 767.2206 (about 6.3613024)
  Min sum of deviations 758.1789                    Pseudo R2     =     0.0118
```

ltotexp	Coefficient	Std. err.	t	P>|t|	[95% conf.	interval]
suppins	.5154982	.107634	4.79	0.000	.3044528	.7265435
_cons	6.056784	.0827831	73.16	0.000	5.894466	6.219103

Interpretation:

- Slope: first decile of log expenditures is .51 larger for people with additional insurance

- Constant: is 10th percentile when suppins=0.

(i.e., slope+constant gives the first decile if suppins=1)

- Estimate several quantile differences

- Stata command: sqreg
- heteroskedastic robust standard errors (bootstrap):

```
. sqreg ltotexp suppins, q(.1 .5 .9) reps(100) nodots

Simultaneous quantile regression                    Number of obs =      2,955
  bootstrap(100) SEs                                .10 Pseudo R2 =     0.0118
                                                    .50 Pseudo R2 =     0.0058
                                                    .90 Pseudo R2 =     0.0000

------------------------------------------------------------------------------
             |              Bootstrap
     ltotexp | Coefficient  std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
q10          |
     suppins |   .5154982   .1097219     4.70   0.000      .300359    .7306373
       _cons |   6.056784   .0932166    64.98   0.000     5.874008     6.23956
-------------+----------------------------------------------------------------
q50          |
     suppins |   .2715392    .054371     4.99   0.000     .1649303    .3781481
       _cons |   7.929846   .0440659   179.95   0.000     7.843443    8.016249
-------------+----------------------------------------------------------------
q90          |
     suppins |  -.0227232   .0958235    -0.24   0.813    -.2106108    .1651644
       _cons |   9.794621   .0774467   126.47   0.000     9.642766    9.946475
------------------------------------------------------------------------------
```

Example 1b

- X: discrete variable

- totchr: Number of Chronic conditions

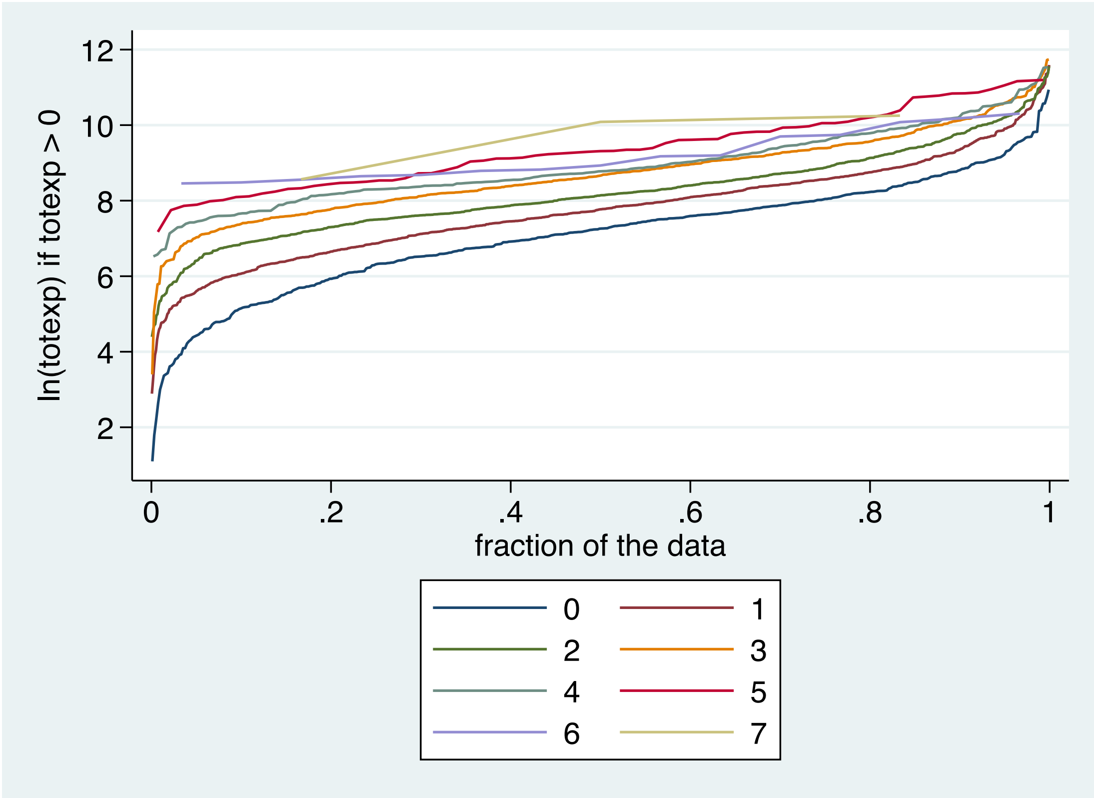
- totchr takes 7 values

```
tabulate totchr
```

# of chronic problems	Freq.	Percent	Cum.
0	466	15.77	15.77
1	865	29.27	45.04
2	809	27.38	72.42
3	506	17.12	89.54
4	222	7.51	97.06
5	69	2.34	99.39
6	15	0.51	99.90
7	3	0.10	100.00
Total	2,955	100.00	

Plot conditional quantiles (one line for each value of totchr)

qplot ltotexp, over(totchr) recast(line) scale(1.1)



Quantile regression, quantile .1

qreg ltotexp totchr, q(.1) nolog

```
qreg ltotexp totchr, q(.1) nolog

1 Quantile regression                              Number of obs =      2,955
 Raw sum of deviations 767.2206 (about 6.3613024)
 Min sum of deviations 653.1636                    Pseudo R2     =     0.1487
```

ltotexp	Coefficient	Std. err.	t	P>|t|	[95% conf.	interval]
totchr	.5674899	.032122	17.67	0.000	.5045061	.6304737
_cons	5.500936	.0714468	76.99	0.000	5.360845	5.641026

- Interpretation: The 10th conditional quantile of ltotexp increases by 0.57 with each extra chronic condition
- very large effect.

- We could create one dummy for each of the values of totchr and run the model:

quietly tabulate totchr, generate(dtotchr)

drop dtotchr1

qreg ltotexp dtotchr*, q(.1) nolog

- Omitted category: no chronic condition

```
. qreg ltotexp dtotchr*, q(.1) nolog

.1 Quantile regression                              Number of obs =      2,955
  Raw sum of deviations 767.2206 (about 6.3613024)
  Min sum of deviations 641.1541                    Pseudo R2     =     0.1643
```

ltotexp	Coefficient	Std. err.	t	P>|t|	[95% conf.	interval]
dtotchr2	.9209313	.1085837	8.48	0.000	.7080238	1.133839
dtotchr3	1.716309	.1098916	15.62	0.000	1.500837	1.931781
dtotchr4	2.236495	.1213225	18.43	0.000	1.99861	2.47438
dtotchr5	2.516852	.1540995	16.33	0.000	2.214699	2.819006
dtotchr6	2.947189	.2437451	12.09	0.000	2.469262	3.425117
dtotchr7	3.335108	.4956903	6.73	0.000	2.363174	4.307042
dtotchr8	3.418108	1.094484	3.12	0.002	1.272078	5.564138
_cons	5.147494	.0875354	58.80	0.000	4.975858	5.319131

- We can also estimate many quantiles

```
.
. sqreg ltotexp totchr, q(.1 .5 .9) reps(100) nodots

Simultaneous quantile regression                    Number of obs =      2,955
  bootstrap(100) SEs                                .10 Pseudo R2 =     0.1487
                                                    .50 Pseudo R2 =     0.0903
                                                    .90 Pseudo R2 =     0.0646
```

ltotexp	Coefficient	Bootstrap std. err.	t	P>|t|	[95% conf.	interval]
q10						
totchr	.5674899	.0258605	21.94	0.000	.5167834	.6181964
_cons	5.500936	.0725731	75.80	0.000	5.358637	5.643235
q50						
totchr	.3932115	.0195693	20.09	0.000	.3548407	.4315823
_cons	7.347944	.0527086	139.41	0.000	7.244594	7.451293
q90						
totchr	.3762154	.0286877	13.11	0.000	.3199655	.4324652
_cons	8.956738	.0816227	109.73	0.000	8.796694	9.116781

- Many variables.

```
. sqreg ltotexp suppins totchr age female white, q(.1 .5 .9) reps(100) nodots

Simultaneous quantile regression                    Number of obs =      2,955
  bootstrap(100) SEs                                .10 Pseudo R2 =     0.1640
                                                    .50 Pseudo R2 =     0.1009
                                                    .90 Pseudo R2 =     0.0687
```

ltotexp	Coefficient	Bootstrap std. err.	t	P>|t|	[95% conf.	interval]
q10						
suppins	.3957205	.0690543	5.73	0.000	.260321	.53112
totchr	.5391863	.0270711	19.92	0.000	.4861061	.5922665
age	.0192688	.0048859	3.94	0.000	.0096888	.0288489
female	-.0127282	.0806778	-0.16	0.875	-.1709188	.1454623
white	.0734392	.1826637	0.40	0.688	-.2847221	.4316006
_cons	3.867043	.4040991	9.57	0.000	3.074698	4.659388
q50						
suppins	.2769771	.0579011	4.78	0.000	.1634465	.3905077
totchr	.3942664	.0218798	18.02	0.000	.3513651	.4371676
age	.0148666	.0040762	3.65	0.000	.0068741	.022859
female	-.0880967	.060479	-1.46	0.145	-.2066821	.0304887
white	.4987457	.2304944	2.16	0.031	.0467995	.9506918
_cons	5.648891	.3507786	16.10	0.000	4.961095	6.336686
q90						
suppins	-.0142829	.0896351	-0.16	0.873	-.1900366	.1614708
totchr	.3579524	.0304578	11.75	0.000	.2982317	.4176731
age	.0059236	.0072497	0.82	0.414	-.0082914	.0201386
female	-.1576335	.0786056	-2.01	0.045	-.3117608	-.0035061
white	.3052239	.2369514	1.29	0.198	-.159383	.7698308
_cons	8.32264	.5399526	15.41	0.000	7.263918	9.381362

Interpretation

- q10 coefficient of suppins: Holding the number of chronic conditions, age, gender and race constant, if we compare people with and without supplementary health insurance, the 10th percentile of log expenditure is 0.396 higher for those with supplementary health insurance.

More on interpretation: Retransformation

- We're computing marginal effects for log expenditures, not for expenditures

- Equivalence property of QR (notice how you can't do this with expectations!):

$$q_\tau(Y|X) = exp q_\tau(log Y|X)) = exp(X'\beta_\tau)$$

- Then, if we want to compute marginal effects with respect to Y (not with respect to log Y):

$$\frac{\partial q_\tau(Y|X)}{\partial X_j} = exp(X'\beta_\tau)\beta_{\tau j}$$

```
quietly predict xb

gen expxk=exp(xb)

display "Multiplier of QR in logs coeffs to get AME in levels =" r(mean)
```

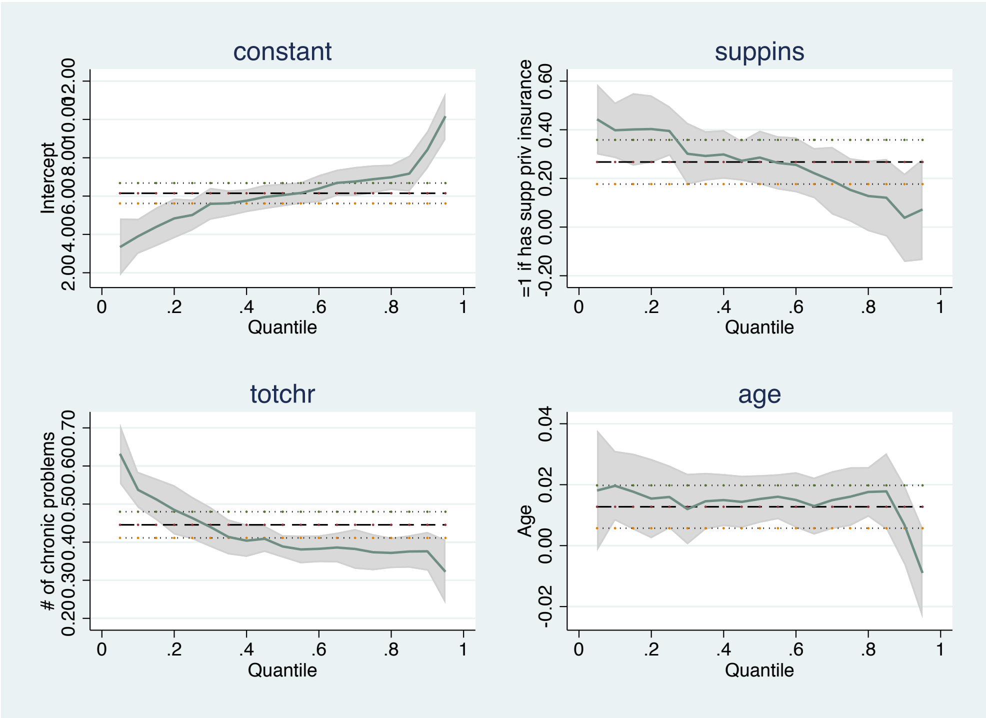
Graphical display of coefficients over quantiles

STATA command: grqreg

ssc install grqreg

bsqreg ltotexp suppins totchr age , reps(100)

grqreg, cons ci ols olsci title(constant suppins totchr age)



- Horizontal line: OLS point estimates and CI (constant across quantiles)

-