Lecture 8: An Introduction to Quantile Methods

Cite as: Victor Chernozhukov, course materials for 14.385 Nonlinear Econometric Analysis, Fall 2007. MIT OpenCourseWare (http://ocw.mit.edu), Massachusetts Institute of Technology. Downloaded on [DD Month YYYY].

Outline

1. Exogenous Quantile Models
2. Quantile Regression
3. Applications

Cite as: Victor Chernozhukov, course materials for 14.385 Nonlinear Econometric Analysis, Fall 2007. MIT OpenCourseWare (http://ocw.mit.edu), Massachusetts Institute of Technology. Downloaded on [DD Month YYYY].

Exogenous Conditional Quantile Models

1. Conditional Quantiles

$$Q_Y(u|X), \quad u \in (0,1),$$

denotes u-quantile of Y conditional on X.

Approximate using linear forms for convenience:

$$Q_Y(u|X) \approx P(X)'\beta(u),$$

where $P(X)$ is a collection of transforms of original regressor X. For simplicity, we can use notation X to mean $P(X)$.

Coefficient $\beta(u)$ can depend on u.

2. Moment Equations that Define Quantiles take a very natural and simple form:

$$P[Y \leq Q_Y(u|X)|X] = u,$$

when Y has a continuous distribution.

Examples:

- Conditional Median $X'\beta(1/2), \quad u = 1/2$
- Conditional Quartile $X'\beta(1/4), \quad u = 1/4$
- Conditional Minimum $X'\beta(0^+), \quad u = 0^+$

- X have heterogeneous impact across quantiles

Cite as: Victor Chernozhukov, course materials for 14.385 Nonlinear Econometric Analysis, Fall 2007. MIT OpenCourseWare (http://ocw.mit.edu), Massachusetts Institute of Technology. Downloaded on [DD Month YYYY].

Quetelet (1870) pioneered the use of growth charts – conditional quantile estimation.

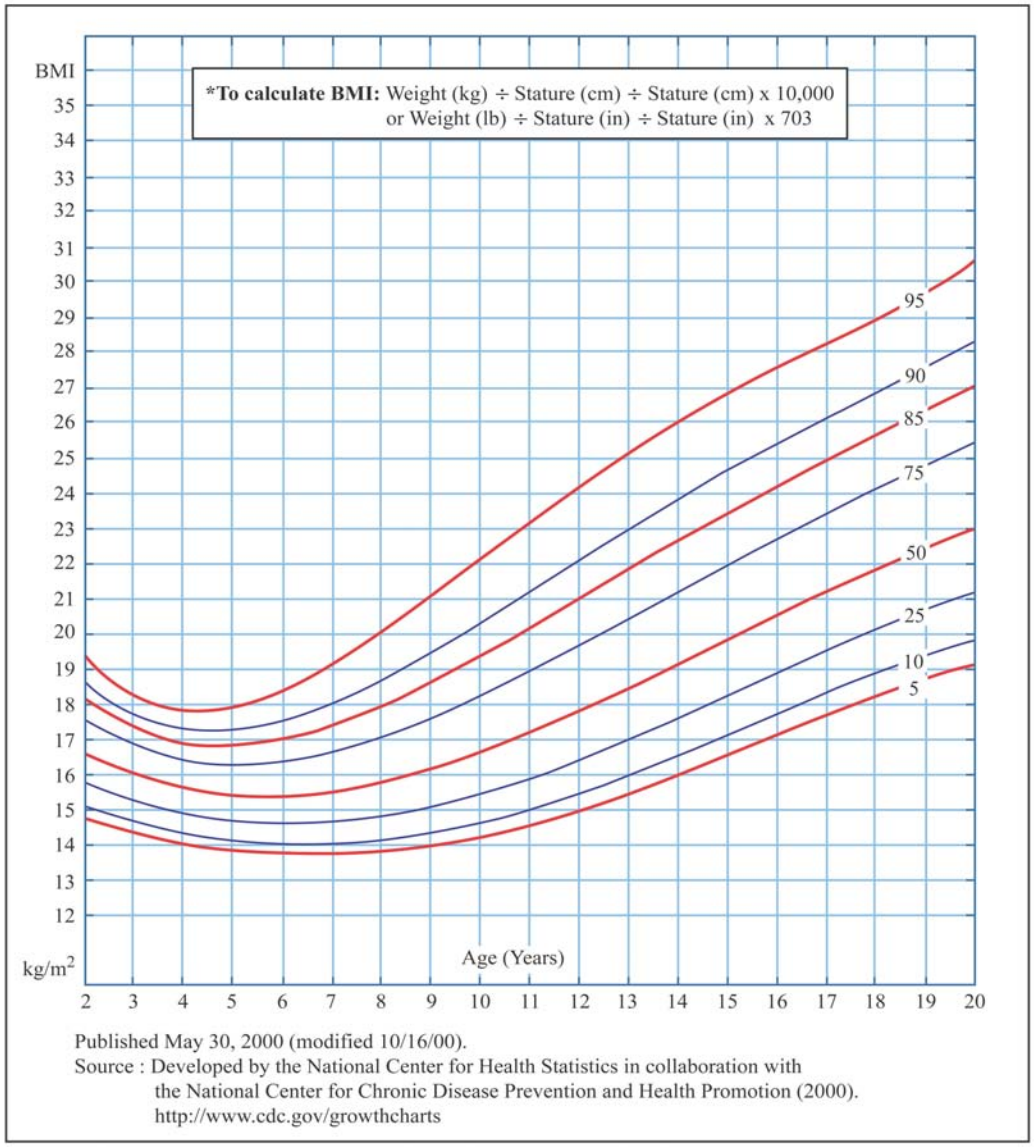


Image by MIT OpenCourseWare.

Cite as: Victor Chernozhukov, course materials for 14.385 Nonlinear Econometric Analysis, Fall 2007. MIT OpenCourseWare (http://ocw.mit.edu), Massachusetts Institute of Technology. Downloaded on [DD Month YYYY].

3. Quantile Function is the Inverse of the Distribution Function

$$F_Y(y|X) = Pr[Y \leq y|X]$$

The inverse is

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

Thus,

$$Q_Y(\cdot|X) = F_Y^{-1}(\cdot|X).$$

4. Non-Additive Error/Random Coefficient Representation

Theorem. Skorohod Representation:

$$Y = Q_Y(U|X), \quad U|X \sim U(0,1)$$

- U is the disturbance or rank variable (absorbing ability, skills, "proneness", (Doksum,74)).

- **Proof.** Construction proceeds by taking

  $$U = F_Y(Y|X),$$

  and we have that $U \sim U(0,1)|X$. Please check this! Now apply the inverse transform $Q_Y(\cdot|X) \equiv F_Y^{-1}(\cdot|X)$ to both sides of the previous display. □

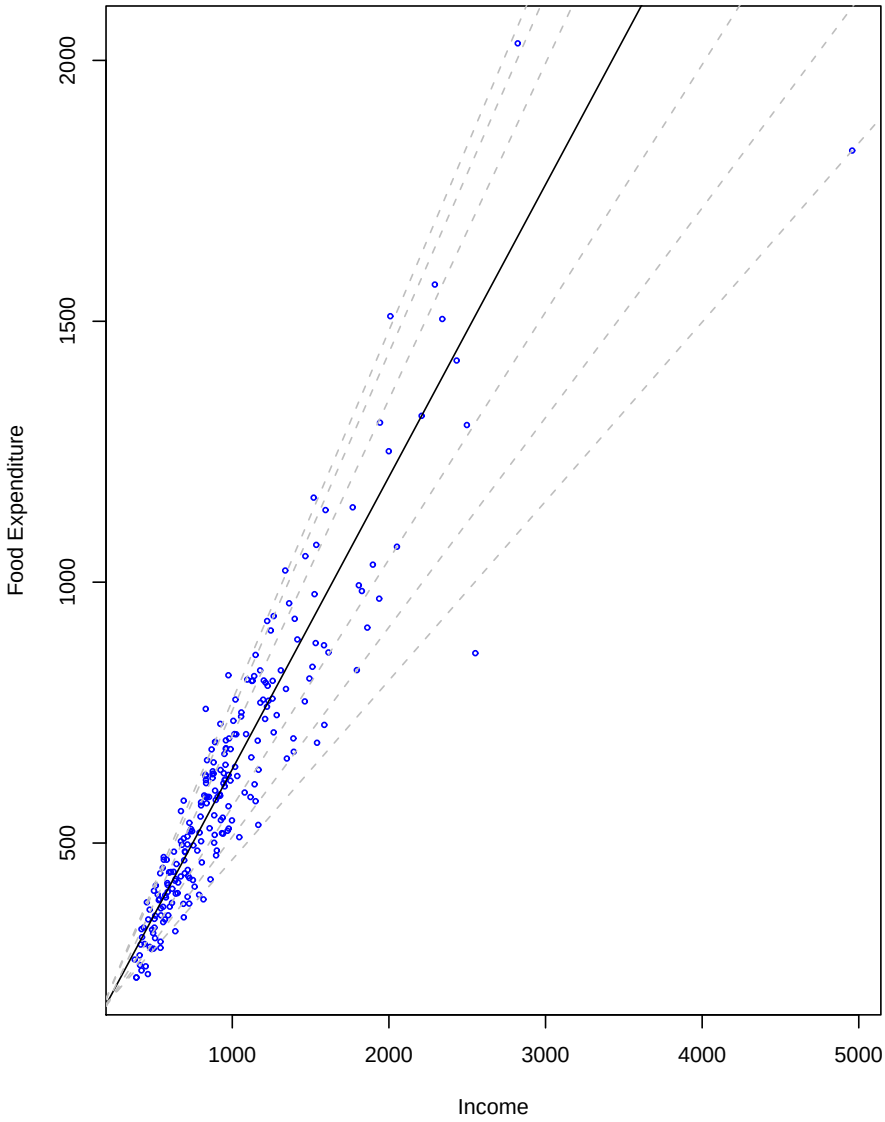
Example:

$$Y = X'\beta(U).$$

Cite as: Victor Chernozhukov, course materials for 14.385 Nonlinear Econometric Analysis, Fall 2007. MIT OpenCourseWare (http://ocw.mit.edu), Massachusetts Institute of Technology. Downloaded on [DD Month YYYY].

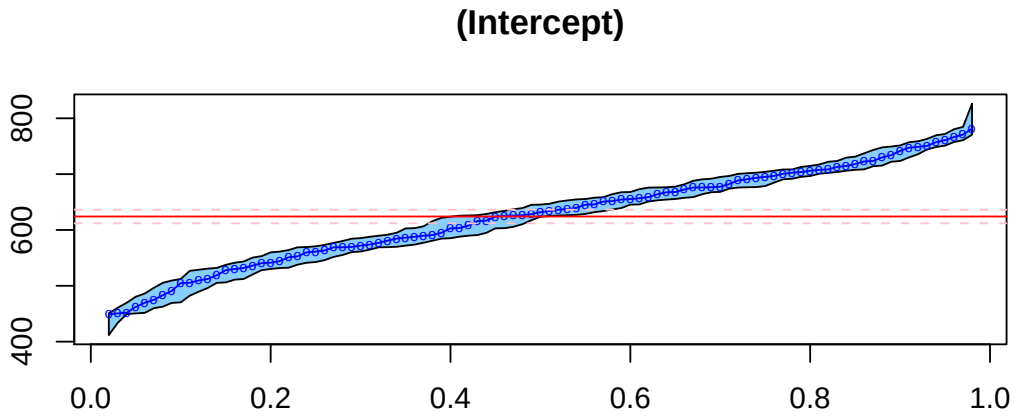
- X affects the *entire shape* of Y distribution. The impact on tail quantile can be very different than the impact on the central quantiles.

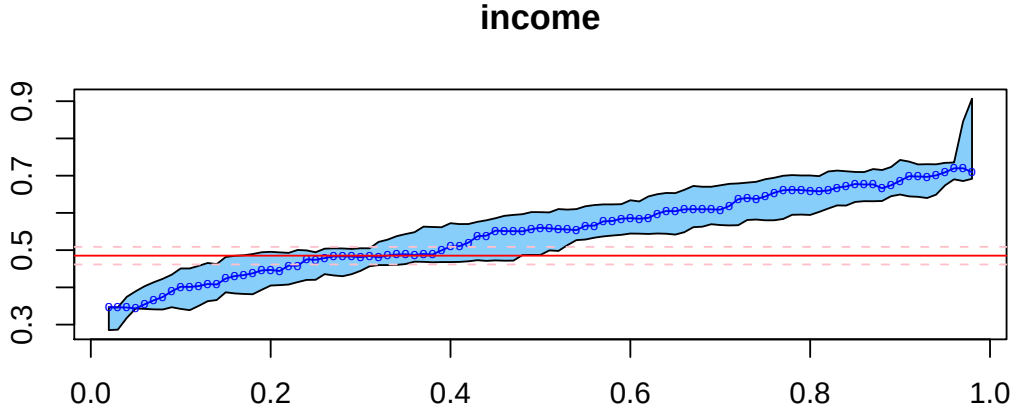
  Example: Engel Curves by Quantile:



Cite as: Victor Chernozhukov, course materials for 14.385 Nonlinear Econometric Analysis, Fall 2007. MIT OpenCourseWare (http://ocw.mit.edu), Massachusetts Institute of Technology. Downloaded on [DD Month YYYY].

- The impact is summarized by the functional coefficient: $u \mapsto \beta(u)$. Example: Coefficient estimates $\widehat{\beta}_0(u)$ and $\widehat{\beta}_1(u)$ as a function of index u. Bands indicate 90% confidence bands. Estimates obtained by quantile regression.





Cite as: Victor Chernozhukov, course materials for 14.385 Nonlinear Econometric Analysis, Fall 2007. MIT OpenCourseWare (http://ocw.mit.edu), Massachusetts Institute of Technology. Downloaded on [DD Month YYYY].

- **Location Model.** Classical location model $Y = X'\beta + \epsilon$ is a very special case, where

$$Y = Q_Y(U|X) = X'\beta + Q_\epsilon(U),$$

in which the impact is summarized by a single number β. That is, in this model, for all $u \in (0,1)$,

$$\beta(u) = (\beta_1 + Q_\epsilon(u), \beta_2, ..., \beta_p)'.$$

Slopes are constant across probability indices u.

- **Location-Scale Model.** Classical location-scale model

$$Y = X'\beta + X'\gamma \cdot \epsilon,$$

where X is a very special case, where

$$Y = Q_Y(U|X) = X'\beta + X'\gamma \cdot Q_\epsilon(U).$$

In this model, for all $u \in (0,1)$

$$\beta(u) = \beta + \gamma \cdot Q_\epsilon(u).$$

That is, all the slope functions $u \mapsto \beta(u)$

  - are monotone in u and

  - are affine transformations of each other.

The Engel curves example is of this kind.

Cite as: Victor Chernozhukov, course materials for 14.385 Nonlinear Econometric Analysis, Fall 2007. MIT OpenCourseWare (http://ocw.mit.edu), Massachusetts Institute of Technology. Downloaded on [DD Month YYYY].

Applications of Quantile Regression Models

1. Growth charts – dependence of height and weight quantiles on age– **Quetelet** c. 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 birthweights:

- "normal" quantiles (Abrevaya, 2001, Koenker, 2005)
- extremal quantiles (250-1000 grams) (Chernozhukov, 2004)

Cite as: Victor Chernozhukov, course materials for 14.385 Nonlinear Econometric Analysis, Fall 2007. MIT OpenCourseWare (http://ocw.mit.edu), Massachusetts Institute of Technology. Downloaded on [DD Month YYYY].

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

6. Value-at-risk (extreme risk) forecasting (Chernozhukov and Umantesev, 2000, Engle and Manganelli, 2000).

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

$$Q_Y(1|x) \qquad Q_Y(1-\epsilon|x)$$

Y production, X factors.

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

$Q_W(\epsilon|X) =$ Approximate Reservation Wage Function

W accepted wage, X covariates.

Cite as: Victor Chernozhukov, course materials for 14.385 Nonlinear Econometric Analysis, Fall 2007. MIT OpenCourseWare (http://ocw.mit.edu), Massachusetts Institute of Technology. Downloaded on [DD Month YYYY].

Estimation of Exogenous Models

a. Quantile Regression (Laplace, 1818, Koenker and Bassett, 1978)

Regression u-Quantile

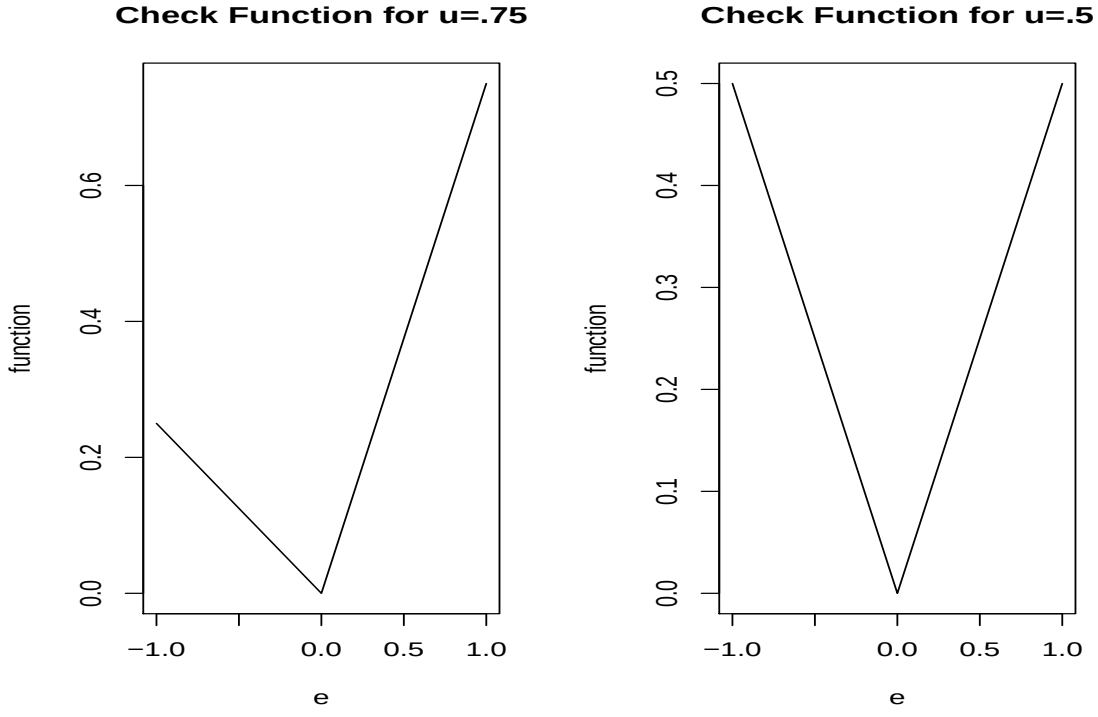
$$\widehat{\beta}(u) = \arg\min_{\beta} E_n[m(Z_i, \beta)]$$

where

$$m(Z_i, \beta) = \rho_u(Y_i - X_i'\beta)$$

and ρ_u is the check function:

$$\rho_u(\epsilon) = u\epsilon^+ + (1-u)\epsilon^- = (u - 1(\epsilon < 0))\epsilon.$$



Cite as: Victor Chernozhukov, course materials for 14.385 Nonlinear Econometric Analysis, Fall 2007. MIT OpenCourseWare (http://ocw.mit.edu), Massachusetts Institute of Technology. Downloaded on [DD Month YYYY].

The population parameter is

$$\beta(u) = \arg\min_{\beta} Em(Z_i, \beta).$$

Note that

$$\nabla_\beta m(Z_i, \beta) = (u - 1\{Y \leq X'\beta(u)\})X_i,$$

Thus, the population parameter solves the moment conditions

$$E[u - 1(Y \leq X'\beta)]X = 0 \text{ or } E[u - P(Y \leq X'\beta|X)]X = 0.$$

These are the correct moment restrictions that arise from original conditional moment restrictions (page 1). The solution at $\beta = \beta(u)$ is unique if and the Hessian of the limit objective function

$$J = \nabla^2_\beta E[m(Z_i, \beta)] = E[f_Y(X'\beta|X)XX'],$$

evaluated at $\beta = \beta(u)$, is full rank (and thus is positive definite due to convexity of the objective function).

Discussion. The great motivation for this objective function is computational, as the solution can be easily computed due to convexity of the objective function. Moreover, the first order conditions for this problem are the "right" moments. Furthermore, another motivation is that in the non-regression case the optimization program finds the sample u-quantile and is thus equivalent to a sorting algorithm.

Note that this is an M-estimator. The key terms in the analysis are thus the gradient and the Hessian.

Cite as: Victor Chernozhukov, course materials for 14.385 Nonlinear Econometric Analysis, Fall 2007. MIT OpenCourseWare (http://ocw.mit.edu), Massachusetts Institute of Technology. Downloaded on [DD Month YYYY].

Theorem. (Koenker, 2005) Let $u \in (0,1)$ be a fixed index. Under appropriate regularity conditions, the estimator $\widehat{\beta}(u)$ is CAN:

$$\sqrt{n}(\widehat{\beta}(u) - \beta(u)) = J^{-1}\sqrt{n}E_n[\nabla_\beta m(Z_i, \beta(u))] + o_p(1)$$
$$\xrightarrow{d} N(0, J^{-1}\Omega J^{-1}),$$

where

$$\Omega = Var[\nabla_\beta m(Z_i, \beta(u))]$$

which simplifies to $\Omega_0 = u(1-u)E[XX']$ in the correctly specified case.

Proof. The basic idea is the same as in the smooth m-estimation case, except that "stochastic" differentiability approach mentioned in Lecture 5 has to be used due to non-smoothness of the finite sample objective function. The details of the proof are assigned as a theoretical exercise.

Remarks:

1) For estimation of J and Ω see the book – Koenker (2005). The best software package quantreg by Koenker is implemented in R.

2) Extremal Case.* The idea is that extremal quantiles (extreme order statistics) do not behave in a normal fashion. This corresponds to the case where $u \approx 0$ or $u \approx 1$. For example, .05 and .95 quantiles in samples of

Cite as: Victor Chernozhukov, course materials for 14.385 Nonlinear Econometric Analysis, Fall 2007. MIT OpenCourseWare (http://ocw.mit.edu), Massachusetts Institute of Technology. Downloaded on [DD Month YYYY].

up to 500 observations do not seem to have distributions close to normal.

Formal asymptotic analysis that mimics such situations is done under hypothetical limits taken as $u \to 0$ as $n \to \infty$, but $un \to c > 0$. Then

$A(n)(\widehat{\beta}(u) - \beta(u)) \approx_d$ Marked Poisson Process Functionals

These results were obtained by Knight, 2002 ($c = 0$) and Chernozhukov, Annals of Statistics, 2005 ($c > 0$). A practical inference theory is available from Chernozhukov and Du (2007).

Cite as: Victor Chernozhukov, course materials for 14.385 Nonlinear Econometric Analysis, Fall 2007. MIT OpenCourseWare (http://ocw.mit.edu), Massachusetts Institute of Technology. Downloaded on [DD Month YYYY].

Computation using interior point methods & preprocessing is very fast (Portnoy and Koenker, Statistical Science):

Koenker's R-package quantreg implements the method.

Running time is $O_p(n^{1+\delta} \dim(\beta)^3)$, faster than OLS.

Laplacian Tortoise has caught up with Gaussian Hare.

Image removed due to copyright restrictions.

Cite as: Victor Chernozhukov, course materials for 14.385 Nonlinear Econometric Analysis, Fall 2007. MIT OpenCourseWare (http://ocw.mit.edu), Massachusetts Institute of Technology. Downloaded on [DD Month YYYY].

Birthweight Application. Here Y is birthweight in grams of live infants born to black mothers, and X are various social and economic factors.

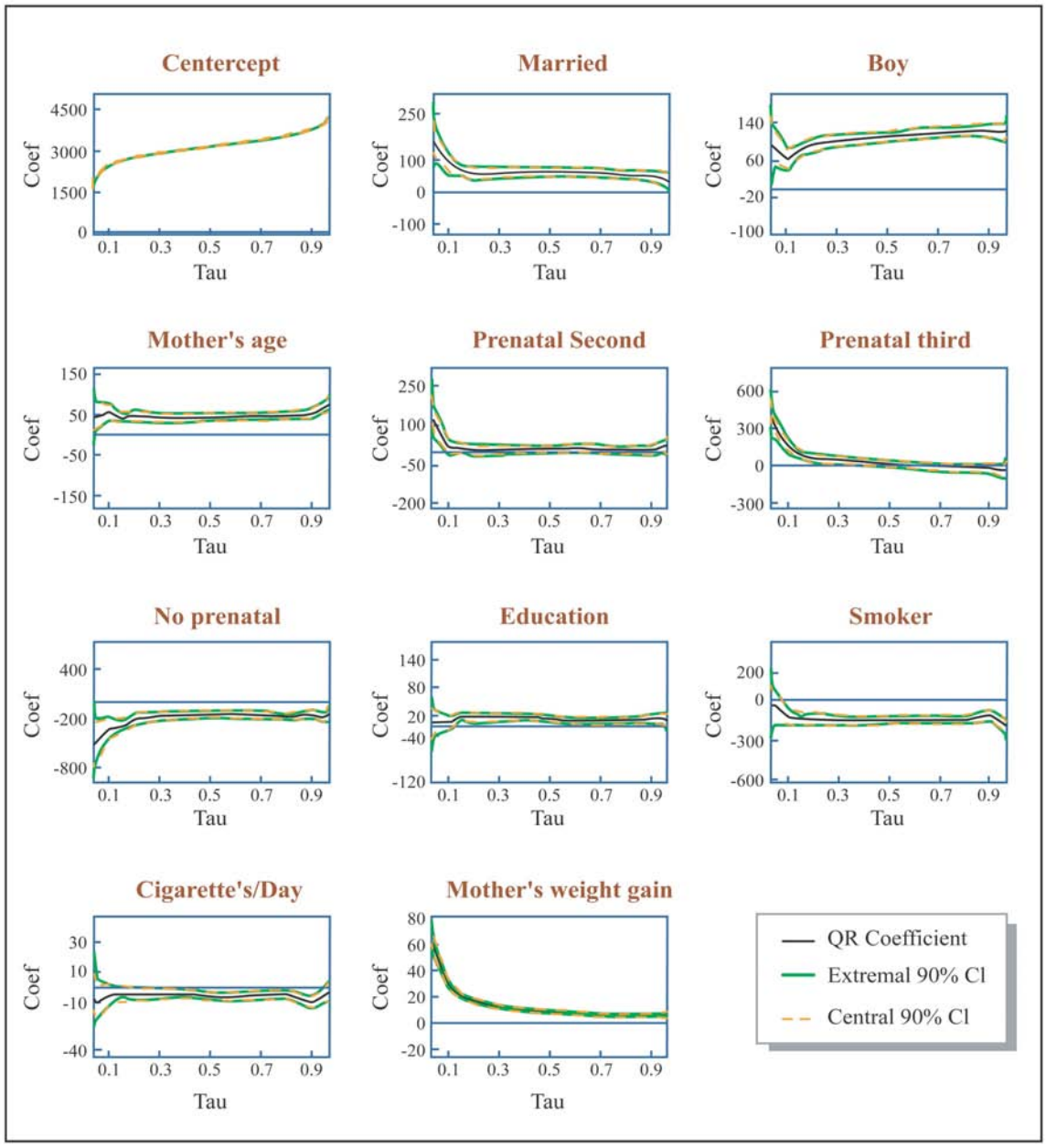


Image by MIT OpenCourseWare.

Cite as: Victor Chernozhukov, course materials for 14.385 Nonlinear Econometric Analysis, Fall 2007. MIT OpenCourseWare (http://ocw.mit.edu), Massachusetts Institute of Technology. Downloaded on [DD Month YYYY].

Changes in Wage Structure in the U.S. in 1980-2000. Here Y records log-wages for prime age white men, and X includes schooling and quadratic function in experience. Ref: Angrist, Chernozhukov and Frenandez (2006, Econometrica)

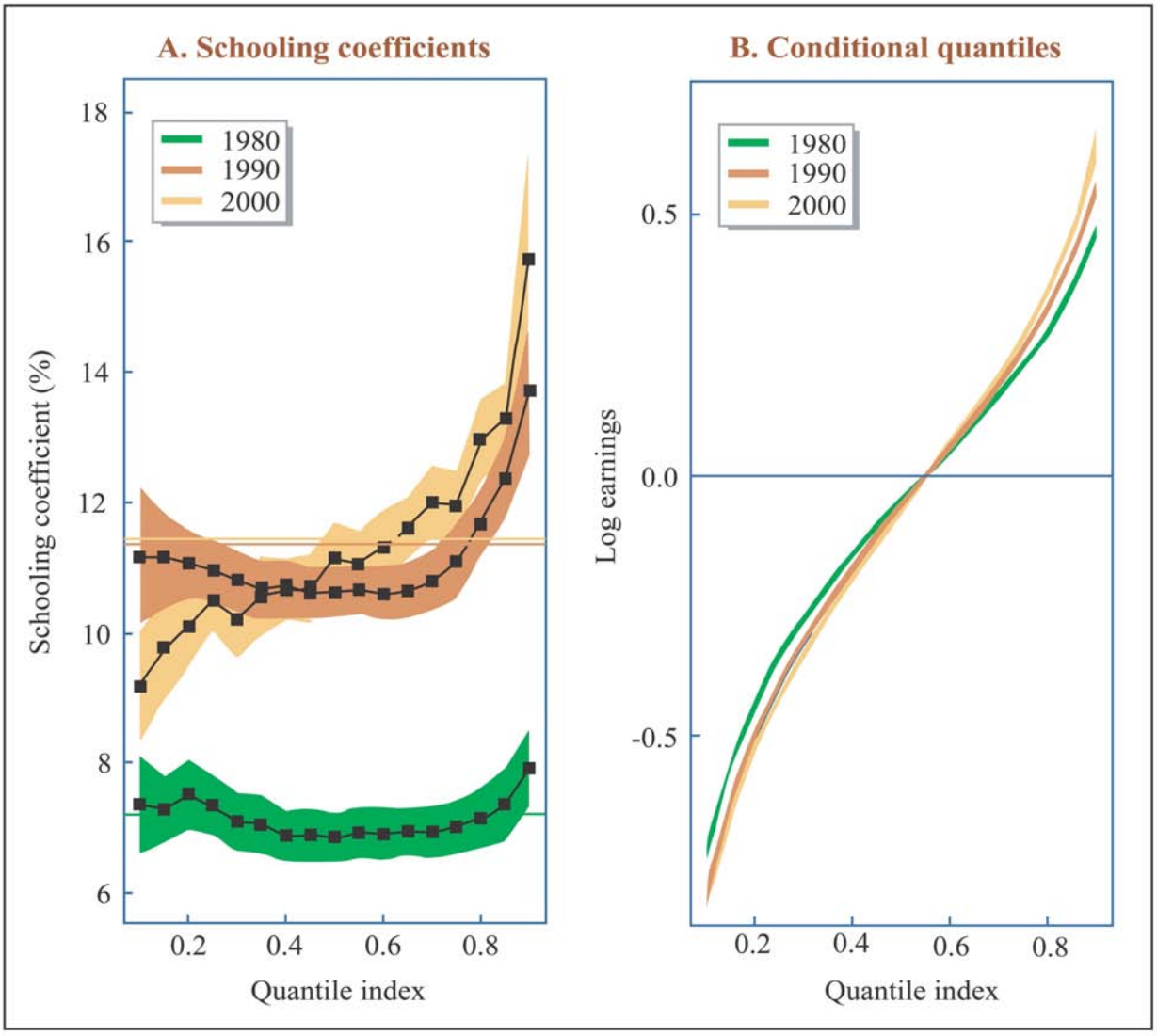


Image by MIT OpenCourseWare.

Cite as: Victor Chernozhukov, course materials for 14.385 Nonlinear Econometric Analysis, Fall 2007. MIT OpenCourseWare (http://ocw.mit.edu), Massachusetts Institute of Technology. Downloaded on [DD Month YYYY].