Microeconometric Modeling

William Greene

Stern School of Business

New York University

4. Quantile Regresion

Quantile Regression

- $Q(y|\mathbf{x},\alpha) = \beta'\mathbf{x}$, α = quantile
- Estimated by linear programming
- $Q(y|\mathbf{x},.50) = \beta'\mathbf{x}$, .50 ➔ median regression
- Median regression estimated by LAD (estimates same parameters as mean regression if symmetric conditional distribution)
- Why use quantile (median) regression?
  - Semiparametric
  - Robust to some extensions (heteroscedasticity?)
  - Complete characterization of conditional distribution

Estimated Variance for Quantile Regression

- Asymptotic Theory
- Bootstrap – an ideal application

Asymptotic Theory Based Estimator of Variance of Q - REG

Model: $y_i = \boldsymbol{\beta}'\mathbf{x}_i + u_i,\;\; Q(y_i \mid \mathbf{x}_i, \alpha) = \boldsymbol{\beta}'\mathbf{x}_i,\;\; Q[u_i \mid \mathbf{x}_i, \alpha] = 0$

Residuals: $\hat{u}_i = y_i - \hat{\boldsymbol{\beta}}'\mathbf{x}_i$

Asymptotic Variance: $\frac{1}{N}\left(\mathbf{A}^{-1}\;\mathbf{C}\;\mathbf{A}^{-1}\right)$

$\mathbf{A} = \mathrm{E}[\mathrm{f}_u(0)\mathbf{x}\mathbf{x}']$ Estimated by $\frac{1}{N}\sum_{i=1}^{N}\frac{1}{\mathrm{B}}\frac{1}{2}1\left[|\hat{u}_i| < \mathrm{B}\right]\mathbf{x}_i\mathbf{x}_i'$

Bandwidth B can be Silverman's Rule of Thumb:

$$\frac{1.06}{N^{.2}}Min\left(s_u, \frac{Q(\hat{u}_i \mid .75) - Q(\hat{u}_i \mid .25)}{1.349}\right)$$

$\mathbf{C} = \alpha(1-\alpha)E[\mathbf{x}\mathbf{x}']$ Estimated by $\frac{\alpha(1-\alpha)}{N}\mathbf{X}'\mathbf{X}$

For α=.5 and normally distributed u, this all simplifies to $\frac{\pi}{2}s_u^2\left(\mathbf{X}'\mathbf{X}\right)^{-1}$.

But, this is an ideal application for bootstrapping.

```
-----------------------------------------------------------------------------
Quantile Regression Model. Quantile =       .250000
Linear Programming estimation method
LHS=HHNINC     Mean                  =        .44583
               Standard deviation    =        .21650
               Number of observs.    =          3377
               Minimum               =        .04000
               t= .25000 quantile    =        .30000
               Maximum               =       3.00000
Model size     Parameters            =             5
               Degrees of freedom    =          3372
Residuals      Sum of squares        =     193.75951
               Standard error of e   =        .20226
Fit            R-squared             =        .12721
               PseudoR2=1-F(0)/F(b)  =        .11046
Not using OLS or no constant. Rsquared may be <= 0
Functions F= Sum r(t)[y(i)-x(i)b] =     164.31749
          F0=Sum r(t)[y(i)-Qy(t)] =     184.72281
             r(t)[u]=t*u-u*[u<0].t=       .250000
--------+--------------------------------------------------------------------
        |                Standard            Prob.      95% Confidence
  HHNINC|  Coefficient     Error       z    |z|>Z*         Interval
--------+--------------------------------------------------------------------
Constant|   -.07580***    .01839    -4.12  .0000     -.11185    -.03975
     AGE|   -.00036**     .00016    -2.25  .0244     -.00068    -.00005
    EDUC|    .02393***    .00137    17.51  .0000      .02125     .02661
 MARRIED|    .11459***    .00547    20.96  .0000      .10388     .12531
    HSAT|    .00773***    .00122     6.31  .0000      .00533     .01013
--------+--------------------------------------------------------------------
Constant|   -.01504       .03479     -.43  .6656     -.08323     .05315
     AGE|   -.00035       .00039     -.90  .3669     -.00112     .00041
    EDUC|    .02707***    .00167    16.19  .0000      .02379     .03035
 MARRIED|    .11361***    .01115    10.19  .0000      .09175     .13547
    HSAT|    .00777***    .00195     3.99  .0001      .00396     .01158
--------+--------------------------------------------------------------------
Constant|    .03738       .03246     1.15  .2495     -.02624     .10099
     AGE|    .00020       .00039      .51  .6100     -.00057     .00097
    EDUC|    .03240***    .00237    13.68  .0000      .02776     .03704
 MARRIED|    .08042***    .01112     7.23  .0000      .05863     .10222
    HSAT|    .00693***    .00231     3.00  .0027      .00240     .01145
--------+--------------------------------------------------------------------
Note: ***, **, * ==>  Significance at 1%, 5%, 10% level.
-----------------------------------------------------------------------------
```

$\alpha = .25$

$\alpha = .50$

$\alpha = .75$

OLS vs. Least Absolute Deviations

```
----------------------------------------------------------------------
Least absolute deviations estimator...............
Residuals     Sum of squares        =     1537.58603
              Standard error of e   =        6.82594
Fit           R-squared             =         .98284
              Adjusted R-squared    =         .98180
Sum of absolute deviations          =    189.3973484
--------+-------------------------------------------------------------
Variable| Coefficient    Standard Error  b/St.Er. P[|Z|>z]   Mean of X
--------+-------------------------------------------------------------
        |Covariance matrix based on   50 replications.
Constant|   -84.0258***     16.08614       -5.223   .0000
       Y|     .03784***       .00271       13.952   .0000      9232.86
      PG|   -17.0990***      4.37160       -3.911   .0001      2.31661
--------+-------------------------------------------------------------
Ordinary      least squares regression ............
Residuals     Sum of squares        =     1472.79834
              Standard error of e   =        6.68059   Standard errors are based on
Fit           R-squared             =         .98356   50 bootstrap replications
              Adjusted R-squared    =         .98256
--------+-------------------------------------------------------------
Variable| Coefficient    Standard Error  t-ratio  P[|T|>t]   Mean of X
--------+-------------------------------------------------------------
Constant|   -79.7535***      8.67255       -9.196   .0000
       Y|     .03692***       .00132       28.022   .0000      9232.86
      PG|   -15.1224***      1.88034       -8.042   .0000      2.31661
--------+-------------------------------------------------------------
```

HEALTH ECONOMICS
Health Econ. **19**: 1063–1074 (2010)
Published online in Wiley InterScience (www.interscience.wiley.com).10.1002/hec.1647

QUANTILE REGRESSION ANALYSIS OF THE RATIONAL ADDICTION MODEL: INVESTIGATING HETEROGENEITY IN FORWARD-LOOKING BEHAVIOR

AUDREY LAPORTE[a,*], ALFIA KARIMOVA[b] and BRIAN FERGUSON[c]

[a]*Department of Health Policy, Management and Evaluation, University of Toronto, Toronto, Ont., Canada*
[b]*Department of Economics, University of Toronto, Toronto, Ont., Canada*
[c]*Department of Economics, University of Guelph, Guelph, Ont., Canada*

SUMMARY

The time path of consumption from a rational addiction (RA) model contains information about an individual's tendency to be forward looking. In this paper, we use quantile regression (QR) techniques to investigate whether the tendency to be forward looking varies systematically with the level of consumption of cigarettes. Using panel data,

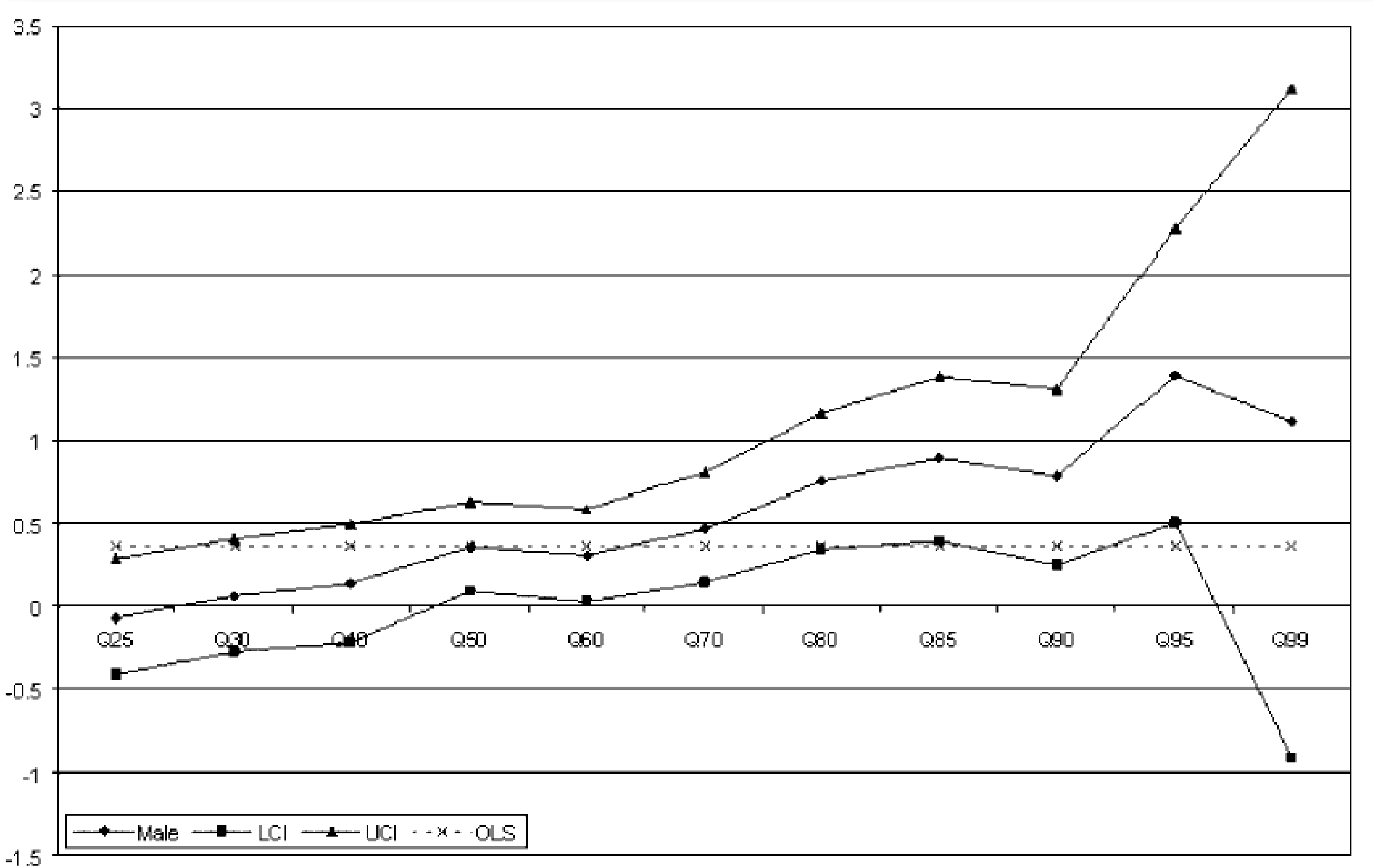

Figure 4. Male