Quantile Regression

Kaden Grace Lucas Gurgel John Paul Lynn

PS 3 Presentation, March 30, 2023

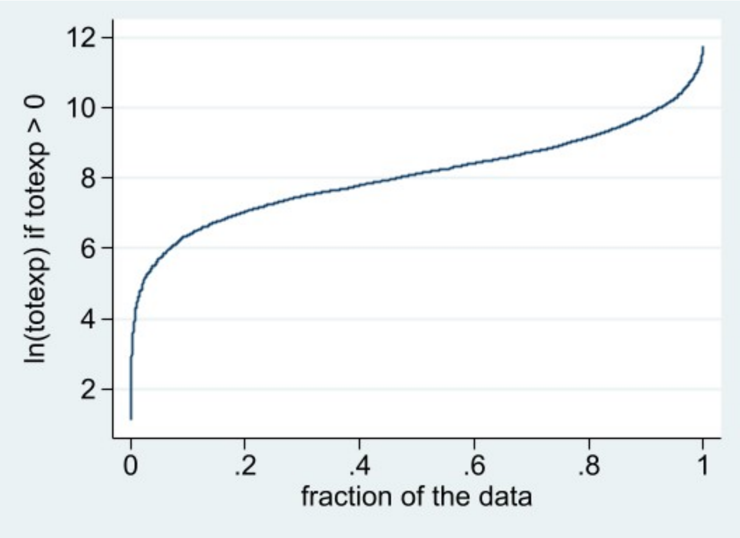
THE UNIVERSITY OF
TENNESSEE
KNOXVILLE

Q1 - Summary of Data

Table: Summary of Log Total Expenditures

	N	Mean	SD	Min	Max
ltotexp	2955	8.06	1.37	1.10	11.74
suppins	2955	.59	.49	0	1
totchr	2955	1.81	1.29	0	7
age	2955	74.245	6.38	65	90
female	2955	.58	.49	0	1
white	2955	.97	.16	0	1

Q2 - Quantile Plot



Q3 - Median Regression ($\tau = 0.5$)

Table: Basic Quantile Regression for tau = .5

Sup Priv Insurance	0.277***	(0.054)
Chronic Condit.	0.394***	(0.020)
Age	0.015***	(0.004)
Female	-0.088*	(0.053)
White	0.499***	(0.163)
Constant	5.649***	(0.341)

Q4 -

- OLS estimates the conditional mean. Jensen's inequality states that $\mathbf{E}(ln(y))$ does not equal $ln(\mathbf{E}(y))$, because means are not invariant to monotonic transformations.
- On the other hand, the quantile regression method is invariant to monotonic transformations (equivariance property), so that the median $ln(y) = ln(median(y))$

Q4 ...

- By how much does one more chronic condition increase the conditional median of expenditures (in levels)?

$$\begin{aligned}\frac{\partial Q_{0.5}(y|x)}{\partial chronic} &= exp(x'\beta)\beta_{chronic} \\ &= (3746.7178)(0.39) \\ &= \$1476.21\end{aligned}$$

Q5 -

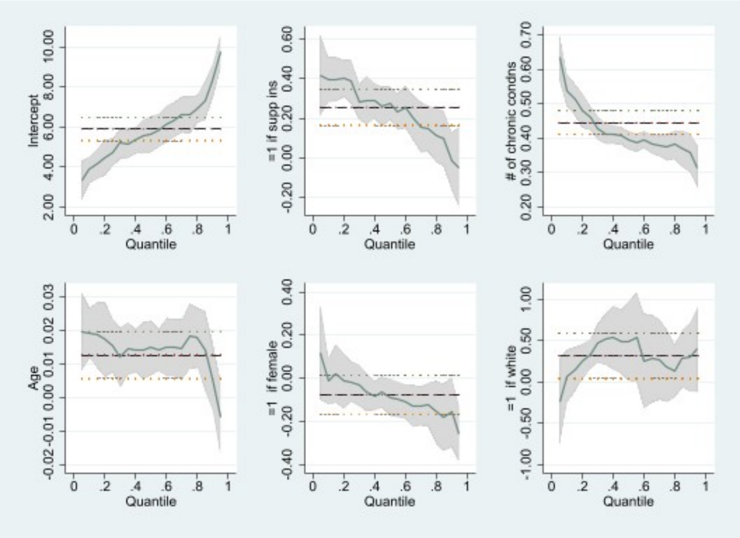
	OLS	QR_25	QR_50	QR_75	BSQR_50
Priv Ins	0.26***	0.39***	0.28***	0.15**	0.28***
	(0.05)	(0.06)	(0.05)	(0.06)	(0.06)
Chronic	0.45***	0.46***	0.39***	0.37***	0.39***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Age	0.01***	0.02***	0.01***	0.02***	0.01***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Female	-0.08*	-0.02	-0.09*	-0.12**	-0.09*
	(0.05)	(0.06)	(0.05)	(0.06)	(0.05)
White	0.32**	0.34*	0.50***	0.19	0.50**
	(0.14)	(0.18)	(0.16)	(0.19)	(0.20)
Constant	5.90***	4.75***	5.65***	6.60***	5.65***
	(0.30)	(0.37)	(0.34)	(0.39)	(0.39)

Q5 ...

- By how much does one more chronic condition increase expenditures (in levels) for $\tau = 0.25$ and $\tau = 0.75$?

$$\begin{aligned}
\frac{\partial Q_{0.25}(y|x)}{\partial chronic} &= exp(x'\beta)\beta_{chronic} \\
&= (3746.7178)(0.46) \\
&= \$1723.24 \\
\frac{\partial Q_{0.75}(y|x)}{\partial chronic} &= exp(x'\beta)\beta_{chronic} \\
&= (3746.7178)(0.37) \\
&= \$1386.29
\end{aligned}$$

Q6 -



Intercept
2.00
4.00
6.00
8.00
10.00
Quantile
=1 if supp ins
-0.20
0.00
0.20
0.40
0.60
# of chronic condns
0.20
0.30
0.40
0.50
0.60
0.70
Age
-0.02
-0.01
0.00
0.01
0.02
0.03
=1 if female
-0.40
-0.20
0.00
0.20
0.40
=1 if white
-1.00
-0.50
0.00
0.50
1.00
0
.2
.4
.6
.8
1

Q7 -

- Santos, Silva, and Tenreyro (2006) argue that important implications of Jensen's inequality have been neglected in econometrics; mainly that under heteroskedasticity, log-linear models estimated by OLS will lead to biased estimates.
- In practice, assuming positive y-values leads to heteroskedasticity; the variance of y becomes small as y approaches 0.
- If there are a large mass of zeros (as is typical in trade data) then log-linearization becomes impossible since $ln(0)$ is undefined and will lead to a discontinuous function.
- Using the mean could also give problems with Jensen's inequality.
- Suggest a Poisson Pseudo Maximum Likelihood (PPML) estimator, which addresses both problems: a large mass of zeros and heteroskedastic errors.

Q7...

- Figueiredo, Lima, and Schaur (2015) appropriately address the previous issues brought to light by Santos, Silva, and Tenreyro (2006).
- Their European Union data set does not have a large mass of zeros, making log-linearization methods feasible.
- In the presence of heteroskedasticity, they use the quantile regression method to account for potentially heterogenous results across different regions of the distribution.
- Unlike the mean function, the quantile regression method is invariant to monotone transformations; the model is not subject to Jensen's inequality.

Q7...

- Their model takes the form:

$$f_{ij} = exp(x_{ij}\beta)\eta_{ij}$$
$$\eta_{ij} = exp[(x_{ij}\gamma)\epsilon_{ij}]$$
$$\epsilon_{ij} \; i.i.d.F_{\epsilon}(0, 1),$$

where $F_{\epsilon}()$ is the unknown distribution function of ϵ_{ij} and

$$F_{\epsilon}^{-1}(\tau) = Q_{\tau}(\epsilon)$$

Q7...

The conditional quantile of f_{ij} is defined:

$$Q_\tau(f_{ij}|x_{ij}) = exp[x_{ij}\beta(\tau)]$$
$$Q_\tau(ln(f_ij)|x_{ij}) = x_{ij}\beta(\tau)$$
$$\beta(\tau) = \beta + \gamma * Q_\tau(\epsilon)$$