Quantile Regression: A simple experiment

MS8956 Presentation

LI Jinze

April 16, 2024

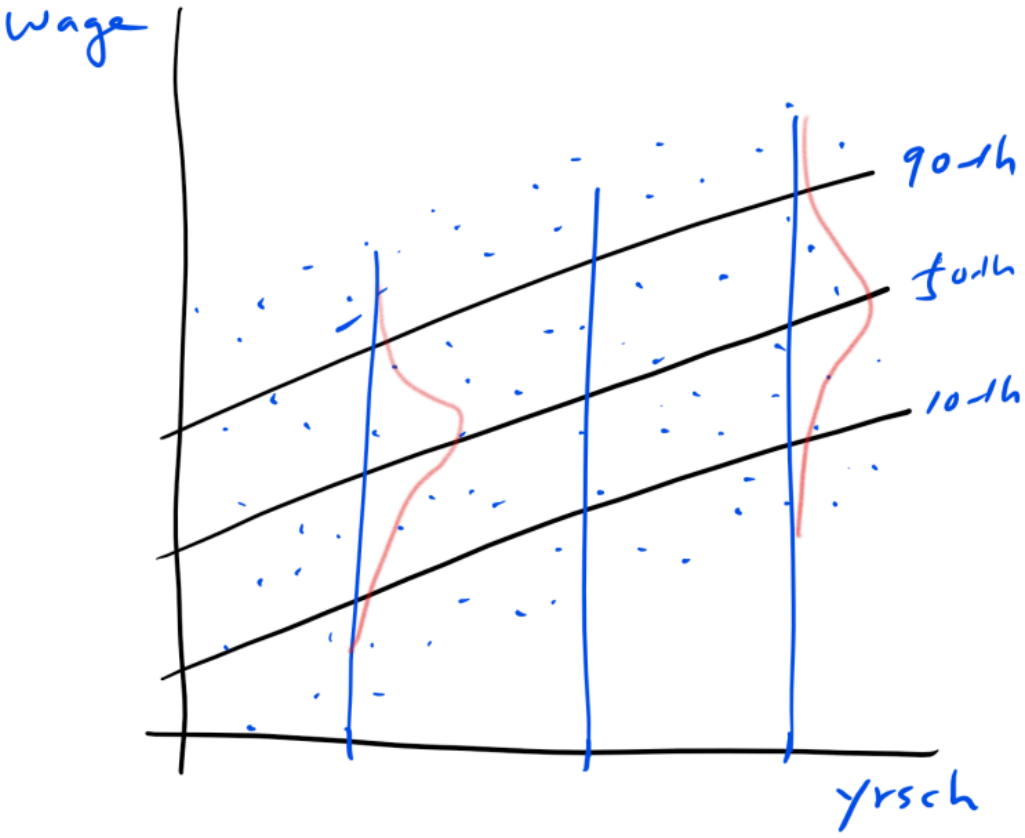
Quantile Regression

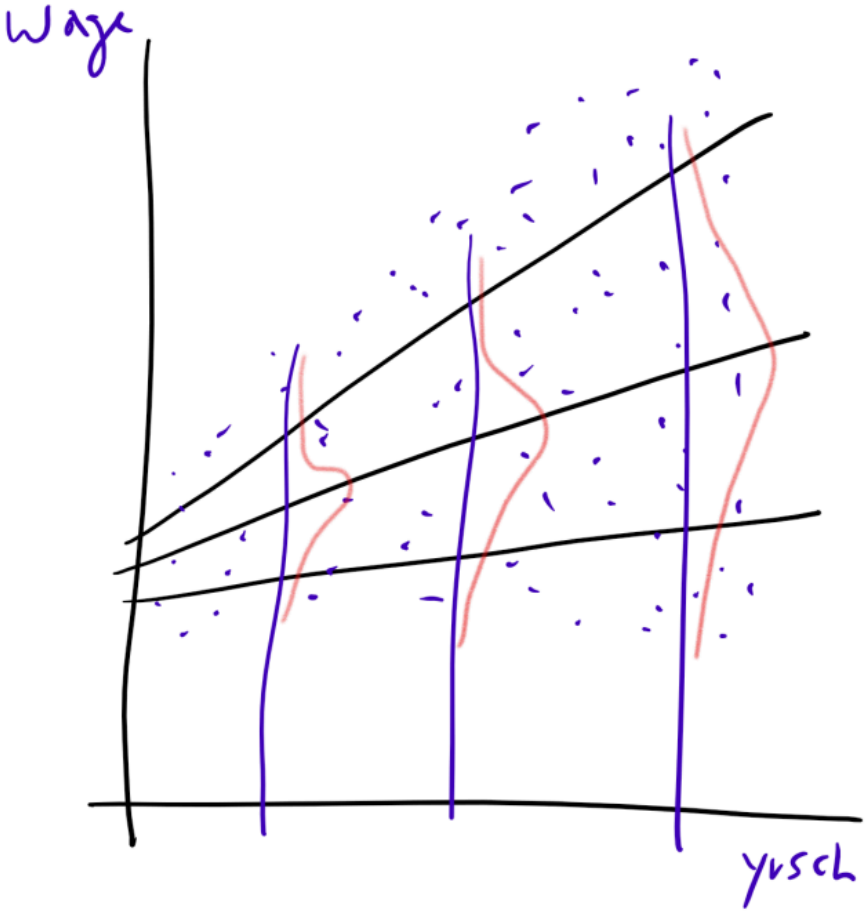
Quantile regression is a type of regression analysis.

OLS	Quantile Regression
-|-	<–>
estimate the mean	a specific quantile
less robust	more robust to outliers
not	invariant to monotonic transform

Risk Management: model and predict extreme events.
Income Inequality: study income distribution and inequality.

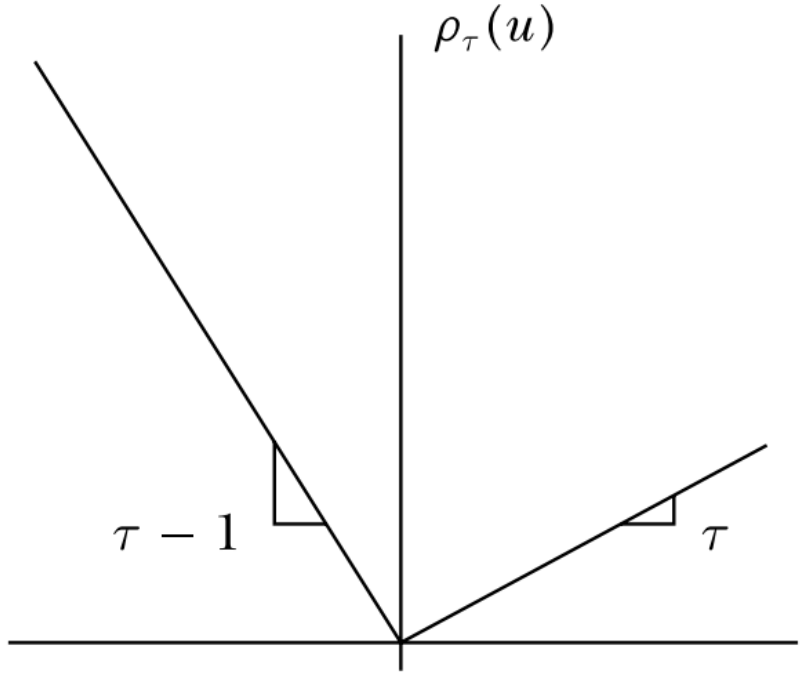
Quantile Regression





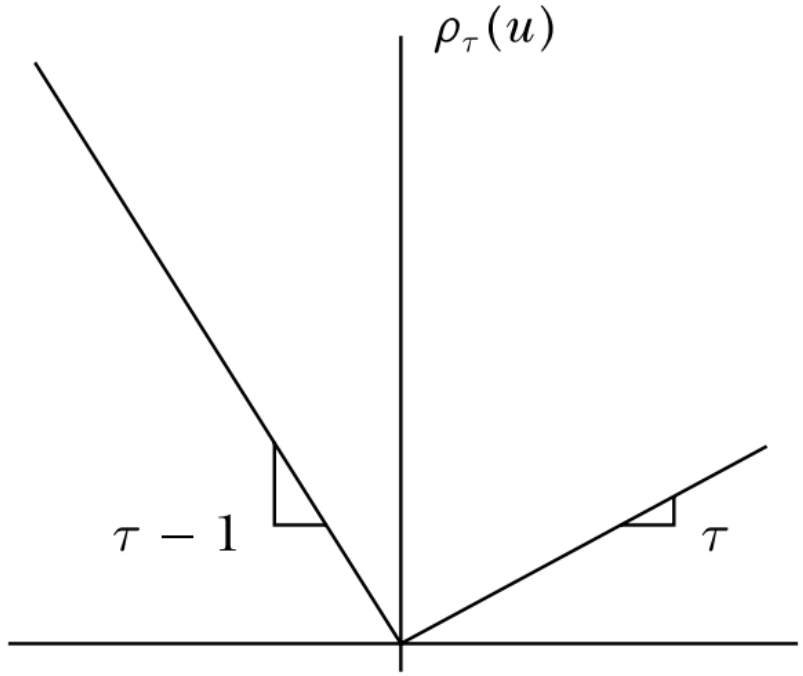
Loss Function

- $\rho_\tau(u) = (\tau - 1)\min(0, u) + \tau \max(0, u)$



Quantile Regression

$$
\begin{aligned}
E[\rho_\tau(X-\hat{x})] &= (\tau-1)\int_{-\infty}^{\hat{x}}(x-\hat{x})\,dF(x)+\tau\int_{\hat{x}}^{\infty}(x-\hat{x})\,dF(x),\\
0 &= (1-\tau)\int_{-\infty}^{\hat{x}}dF(x)-\tau\int_{\hat{x}}^{\infty}dF(x)\\
\hat{x} &= \mathsf{F}^{-1}(\tau)
\end{aligned}
$$



Conditional QR

- conditional linear QR

$$(\hat{\alpha}, \hat{\beta}) = \text{argmin}_{\alpha,\beta} \sum_{i=1}^{n} \rho_\tau(y_i - \alpha - \beta' x_i) \tag{1}$$

- conditional non-linear QR

$$(\hat{\alpha}, \hat{\beta}) = \text{argmin}_{\alpha,\beta} \sum_{i=1}^{n} \rho_\tau(y_i - g(\alpha + \beta' x_i)) \tag{2}$$

- Conditional linear QR can be solved by Linear Programming

Data View

Automobile Data:
https://archive.ics.uci.edu/dataset/10/automobile

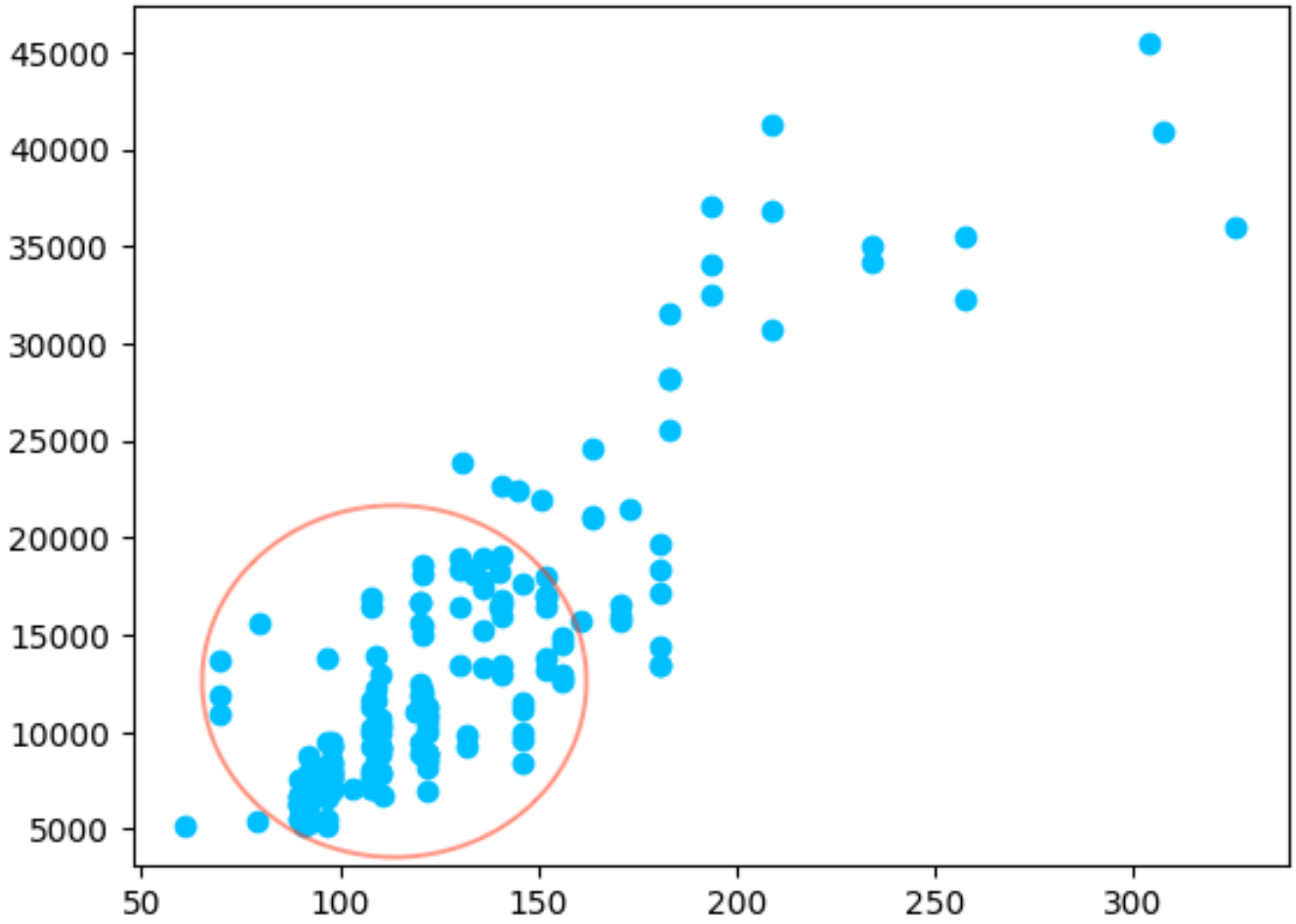
```
In [5]: data.head()
```

Out[5]:

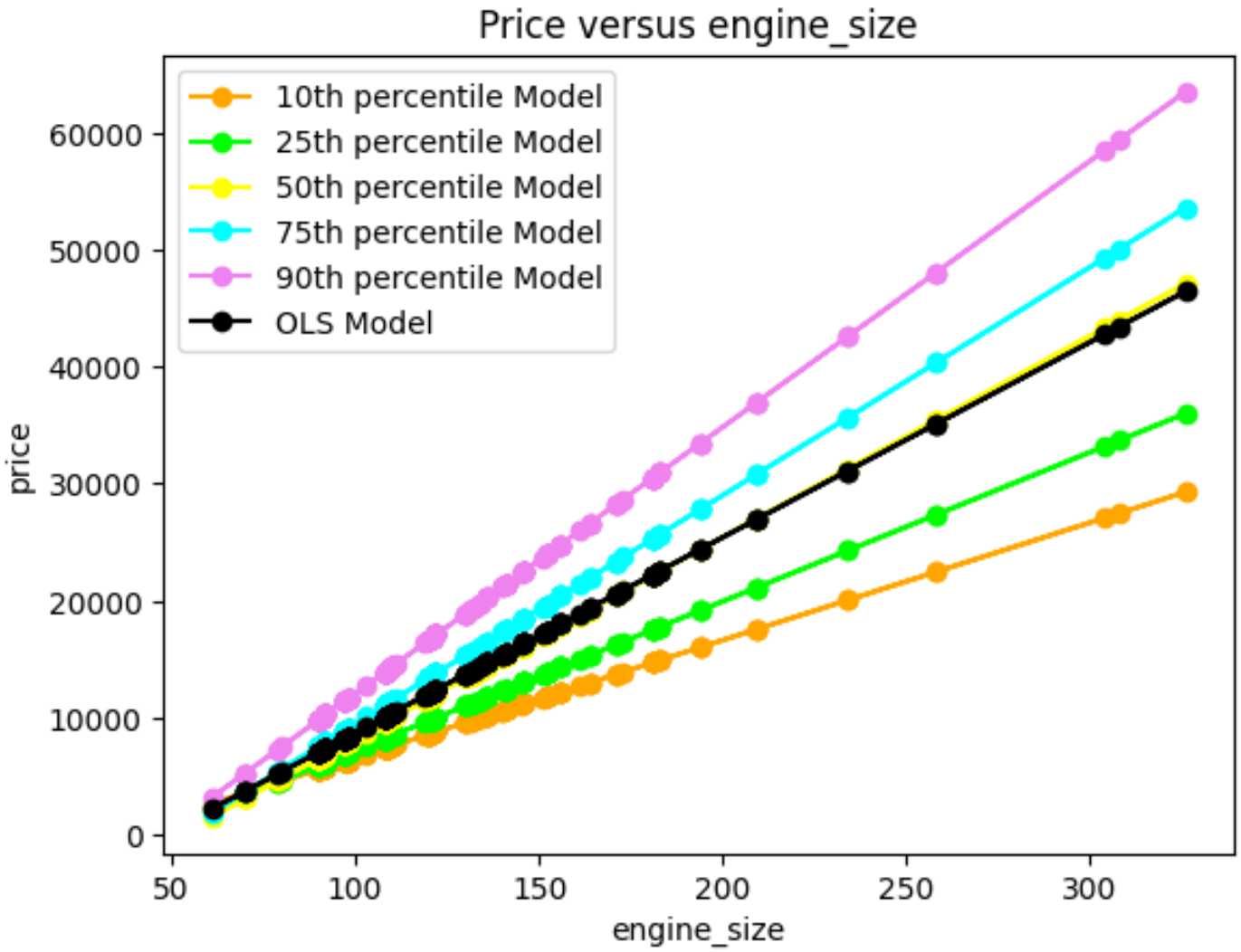
	make	aspiration	body_style	curb_weight	num_of_cylinders	engine_size	price
0	alfa-romero	std	convertible	2548	4	130	13495
1	alfa-romero	std	convertible	2548	4	130	16500
2	alfa-romero	std	hatchback	2823	6	152	16500
3	audi	std	sedan	2337	4	109	13950
4	audi	std	sedan	2824	5	136	17450

Data Visualization

y: Price x: Engine Size



Result 1

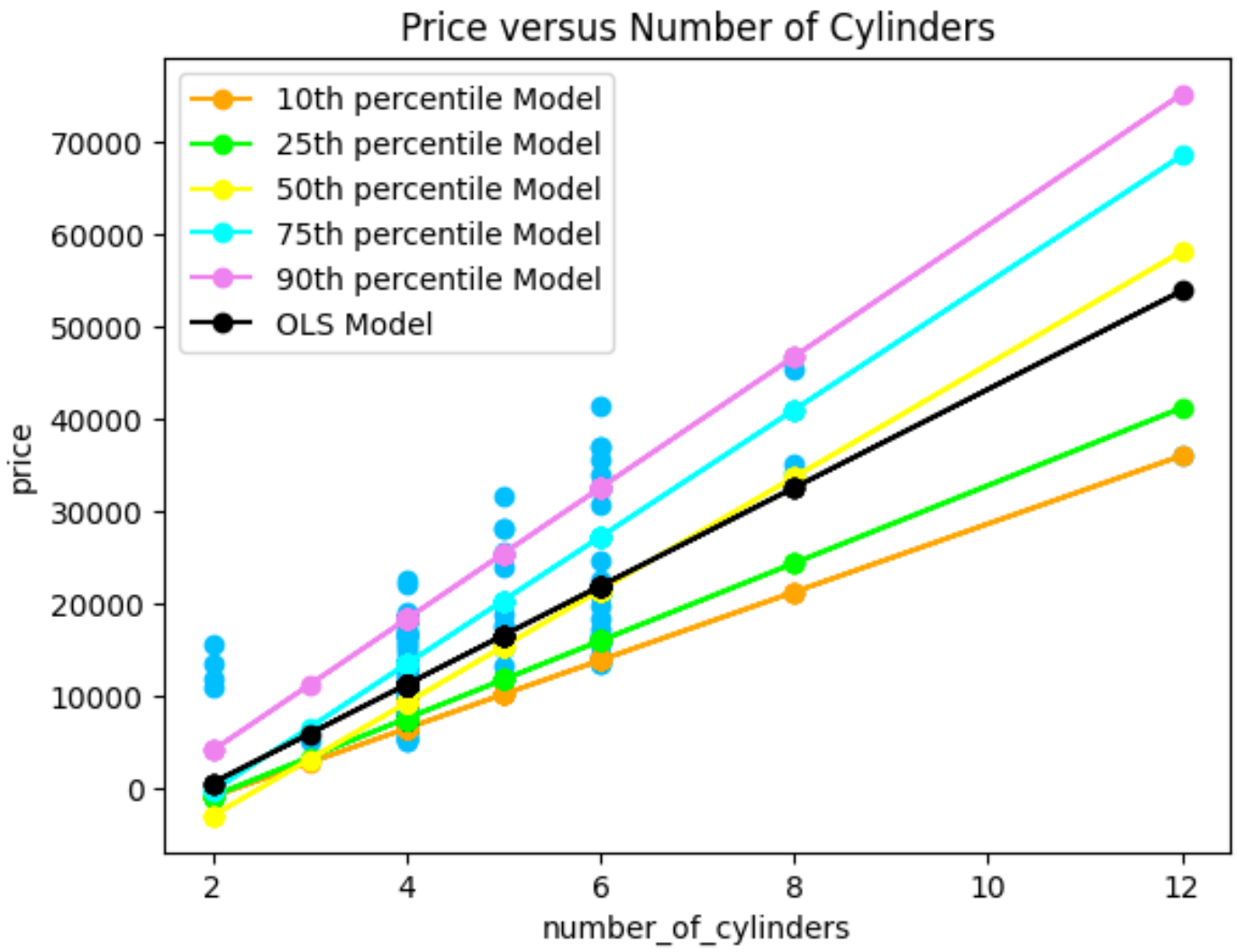

Price versus engine_size
10th percentile Model
25th percentile Model
50th percentile Model
75th percentile Model
90th percentile Model
OLS Model
price
engine_size
0
10000
20000
30000
40000
50000
60000
50
100
150
200
250
300

Result 2: 10th quantile

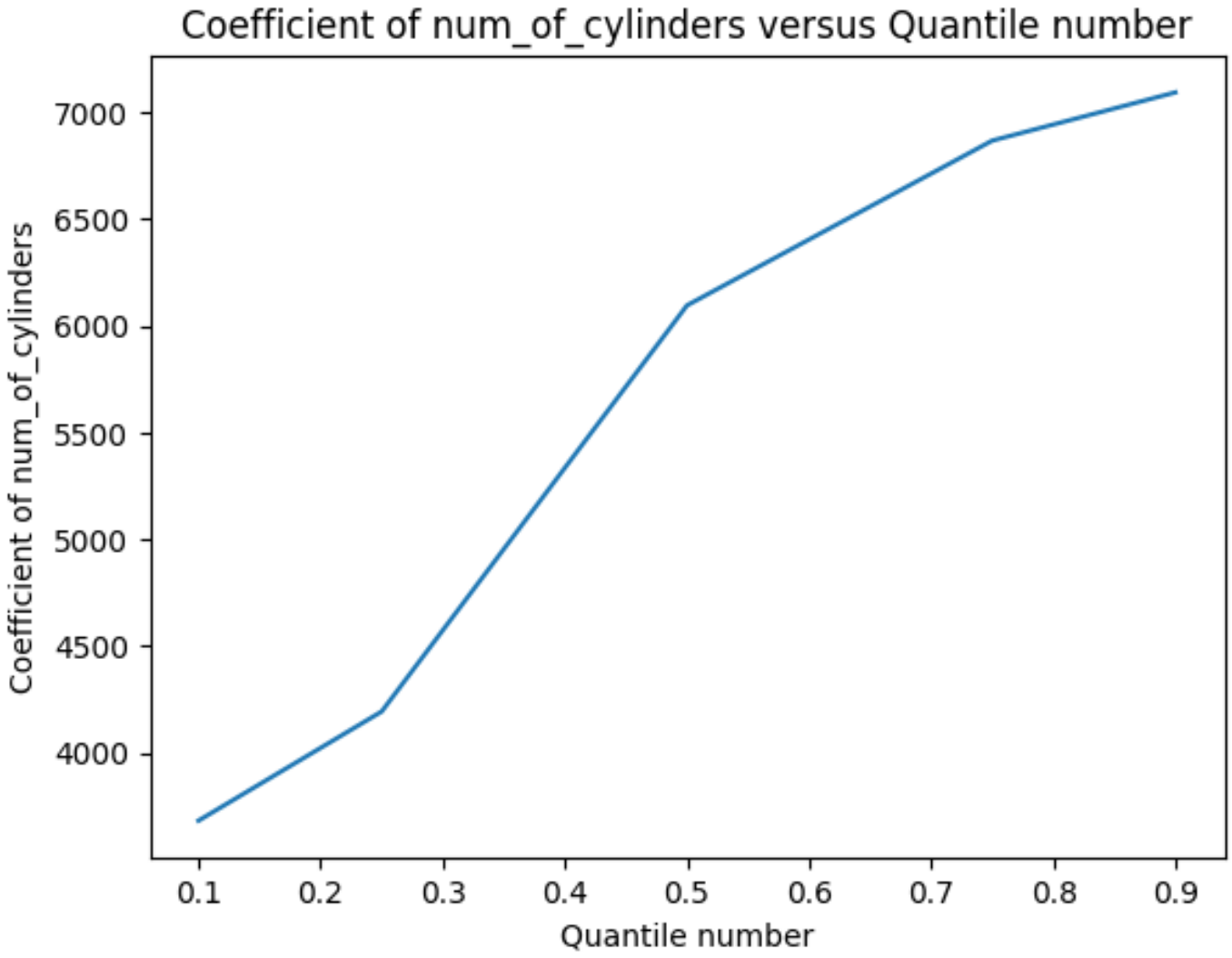
```
10th Quantile Regression Model

                         QuantReg Regression Results
==============================================================================
Dep. Variable:                  price   Pseudo R-squared:               0.3496
Model:                       QuantReg   Bandwidth:                      3120.
Method:                 Least Squares   Sparsity:                       8825.
Date:                Sun, 03 Mar 2024   No. Observations:                 201
Time:                        19:49:34   Df Residuals:                     199
                                        Df Model:                           1
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const       -3471.3929    459.912     -7.548      0.000   -4378.319   -2564.466
engine_size   100.4821      3.326     30.213      0.000      93.924     107.040
==============================================================================
```

Result 3

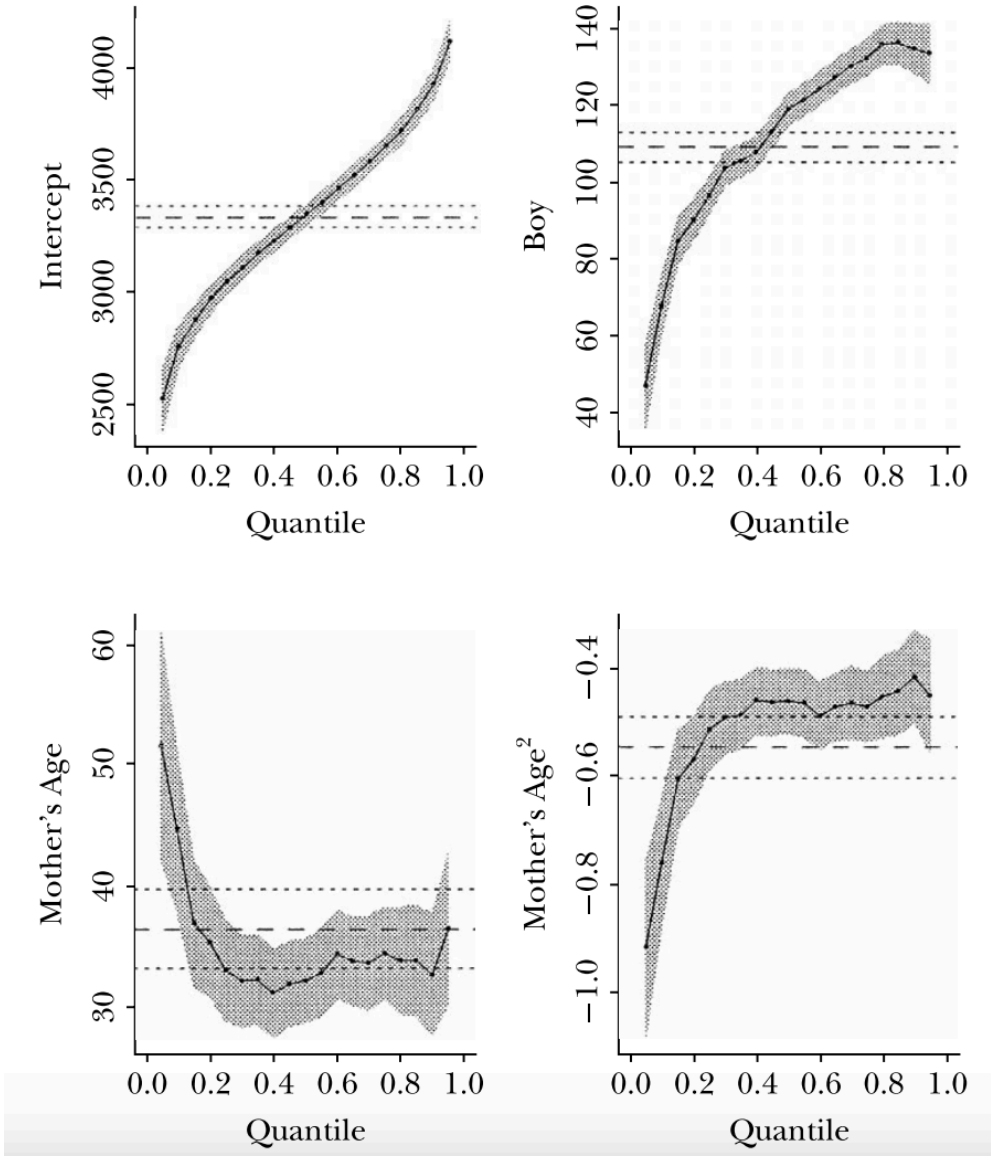

Price versus Number of Cylinders
10th percentile Model
25th percentile Model
50th percentile Model
75th percentile Model
90th percentile Model
OLS Model
price
70000
60000
50000
40000
30000
20000
10000
0
2
4
6
8
10
12
number_of_cylinders

Result 4: coefficient



Result: Koenker and Hallock (2001)

Determinants of Infant Birthweight



Result 5: conditional 10th quantile

```
10th Quantile Regression Model

                         QuantReg Regression Results
==============================================================================
Dep. Variable:                  price   Pseudo R-squared:                0.3631
Model:                       QuantReg   Bandwidth:                       2896.
Method:                 Least Squares   Sparsity:                        8920.
Date:                Sun, 03 Mar 2024   No. Observations:                  201
Time:                        20:53:02   Df Residuals:                      196
                                        Df Model:                            4
==============================================================================
                     coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
body_style      -259.3258    256.944     -1.009      0.314    -766.056     247.404
const          -6074.7999   1504.274     -4.038      0.000   -9041.440   -3108.160
engine_size      108.4548      8.736     12.414      0.000      91.226     125.684
make             -42.4867     24.497     -1.734      0.084     -90.799       5.825
num_of_cylinders 627.8831    514.367      1.221      0.224    -386.522    1642.288
==============================================================================
```

Restult 6: Comparasion

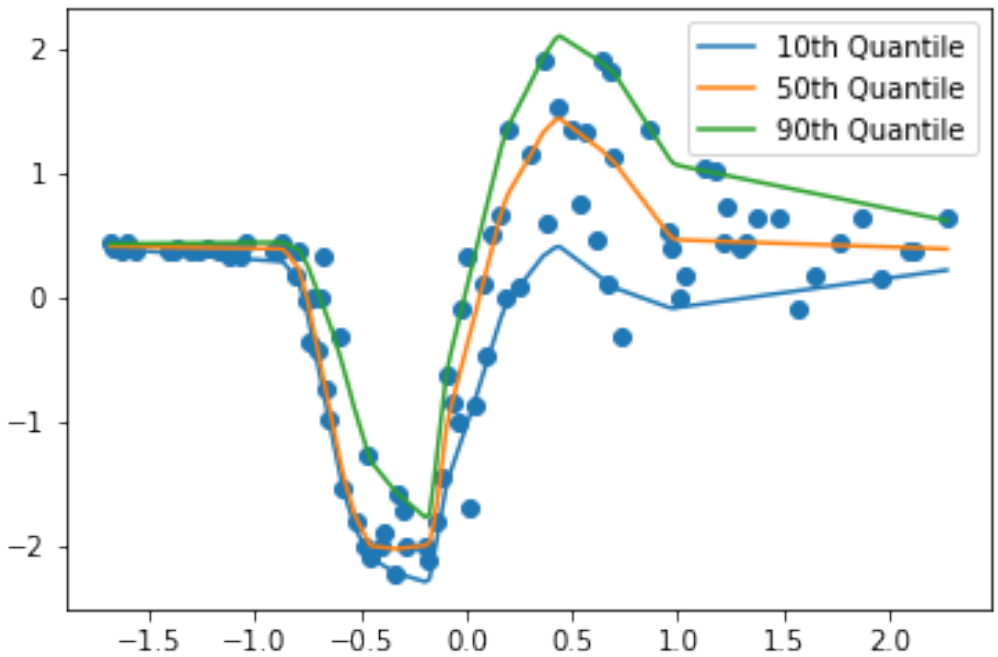
```
#conditional QR above
```

	coef	std err	t	P>|t|	[0.025	0.975]
engine_size	108.4548	8.736	12.414	0.000	91.226	125.684
engine_size	124.7681	9.389	13.289	0.000	106.252	143.284
engine_size	179.2992	11.475	15.625	0.000	156.668	201.931
engine_size	244.5579	11.010	22.213	0.000	222.846	266.270
engine_size	312.2262	15.627	19.980	0.000	281.408	343.044

```
#unconditional QR from the very beginning
```

	coef	std err	t	P>|t|	[0.025	0.975]
engine_size	100.4821	3.326	30.213	0.000	93.924	107.040
engine_size	127.4649	4.785	26.641	0.000	118.030	136.900
engine_size	171.9091	5.881	29.233	0.000	160.313	183.506
engine_size	195.0846	6.265	31.141	0.000	182.731	207.438
engine_size	227.5281	19.160	11.875	0.000	189.745	265.311

Non-linear QR



Non-linear QR

- Koenker R, Xiao Z. Quantile autoregression[J]. Journal of the American statistical association, 2006, 101(475): 980-990.
- Padilla O H M, Tansey W, Chen Y. Quantile regression with ReLU networks: Estimators and minimax rates[J]. Journal of Machine Learning Research, 2022, 23(247): 1-42.
- Chronopoulos I C, Raftapostolos A, Kapetanios G. Forecasting Value-at-Risk using deep neural network quantile regression[J]. Journal of Financial Econometrics, 2023.

Reference

- Koenker R, Bassett Jr G. Regression quantiles[J]. Econometrica: journal of the Econometric Society, 1978: 33-50.
- Koenker R, Hallock K F. Quantile regression[J]. Journal of economic perspectives, 2001, 15(4): 143-156.
- http://www.econ.uiuc.edu/ roger/research/rq/rq.html