In-Class Activity: Residential Mobility Preferences

Quantile Regression and Bootstrap Inference

1 Objective

Estimate residential mobility preferences using stated-choice experimental data. Learn to: (1) estimate quantile regressions with panel data, (2) compute bootstrapped standard errors with clustering, and (3) calculate willingness-to-pay measures with uncertainty quantification.

2 Background

This exercise uses data from a Survey of Consumer Expectations (SCE) experiment where respondents evaluated hypothetical residential moves. Each scenario presents three alternatives with varying attributes (housing costs, crime, distance, etc.). The dependent variable is the **log odds ratio**:

$$\text{ratio}_{ij} = \ln\left(\frac{p_{ij}}{p_{i,\text{stay}}}\right)$$

where p_{ij} is the stated probability of choosing alternative j and $p_{i,\text{stay}}$ is the probability of staying.

3 Data Structure

- **Unit of observation**: Individual-scenario-alternative
- **Panel structure**: Multiple scenarios per individual (`scuid`)
- **Key variables**:
  - `ratio`: Log odds ratio (dependent variable)
  - `income`: Log income differential
  - `homecost`: Log housing cost differential
  - `crime`: Log crime rate differential
  - `dist`: Distance from current location
  - `moved`: Indicator for having moved

4 Step-by-Step Guide

4.1 Step 1: Load and Explore Data

Commands:

```
import delimited "https://raw.githubusercontent.com/OU-PhD-
    Econometrics/fall-2025/master/LectureNotes/11-SubjExp/
    SCEmobilityExample.csv", clear
describe
summarize
```

Questions:

1. How many individuals are in the sample?
2. How many scenarios does each individual evaluate on average?
3. What is the mean and median of the `ratio` variable?

4.2 Step 2: Visualize Choice Patterns

Command:

```
histogram ratio, bin(30)
```

Interpretation: The distribution of log odds ratios reveals preference intensity. Values near zero indicate indifference between moving and staying.

4.3 Step 3: Basic Quantile Regression

Theoretical foundation: We estimate the conditional median:

$$Q_{0.5}(\text{ratio}_{ij}|\mathbf{X}_{ij}) = \beta_0 + \beta_{\text{income}} \cdot \text{income}_{ij} + \beta_{\text{crime}} \cdot \text{crime}_{ij} + \ldots$$

Command:

```
xtset scuid
qreg ratio income homecost crime dist family size mvcost
    taxes norms schqual withincitymove copyhome moved
```

Questions:

1. What is the sign and magnitude of β_{income}?
2. Which attributes have the largest (in absolute value) coefficients?
3. Are the signs economically sensible?
4. How do the median regression estimates compare with OLS?

4.4 Step 4: Bootstrap Standard Errors

Why bootstrap? Standard errors must account for:

- Within-individual correlation (clustering)
- Non-normal sampling distribution of quantile estimators

Command:

```
bootstrap, cluster(scuid) reps(100): qreg ratio income
    homecost crime dist family size mvcost taxes norms
    schqual withincitymove copyhome moved
```

Questions:

1. How do bootstrapped SEs compare to standard SEs?
2. Why do we cluster at the individual (`scuid`) level?

4.5 Step 5: Calculate Willingness-to-Pay

WTP formula: For attribute x, willingness-to-pay is:

$$\text{WTP}_x = -\left[\exp\left(-\frac{\beta_x}{\beta_{\text{income}}} \cdot \Delta x\right) - 1\right] \cdot \text{Income}_{\text{median}}$$

where:

- Δx is the change in attribute (e.g., ln(2) for doubling crime)
- $\text{Income}_{\text{median}} = \$65{,}000$ (sample median)

Example - Crime WTP:

```
scalar b_inc = _b[income]
scalar b_crime = _b[crime]
scalar wtp_crime = -(exp(-b_crime/b_inc * ln(2)) - 1) * 65000
display "WTP to avoid doubling crime: $" wtp_crime
```

Questions:

1. What is the WTP to avoid a doubling of the crime rate?
2. What is the WTP per mile of distance reduction?
3. What is the non-pecuniary moving cost?

4.6 Step 6: Bootstrap WTP Estimates

Why? WTP is a nonlinear transformation of βs. Bootstrap provides correct inference.

Command:

```
bootstrap crime_wtp = (-(exp(-_b[crime]/_b[income]*ln(2))-1)*65000)
         dist_wtp = (-(exp(-_b[dist]/_b[income])-1)*65000)
         moved_wtp = (-(exp(-_b[moved]/_b[income])-1)*65000),
         cluster(scuid) reps(100):
    qreg ratio income homecost crime dist family size mvcost
        taxes norms schqual withincitymove copyhome moved
```

Questions:

1. What are the 95% confidence intervals for each WTP?
2. Are the WTP estimates statistically significant?

5 Discussion Questions

1. **Economic interpretation**: What does a WTP of $50,000 to avoid doubling crime mean in practice?
2. **Heterogeneity**: How might WTP differ by demographics (age, income, family status)?
3. **Policy relevance**: How could these estimates inform urban planning or housing policy?
4. **Limitations**: What assumptions underlie this analysis? When might they fail?

6 Extensions (Optional)

1. Estimate the model separately by homeownership status
2. Calculate WTP as percentages of income
3. Examine heterogeneity using quantile regression at different percentiles ($\tau = 0.25, 0.75$)

7 Key Takeaways

- Quantile regression is robust to outliers and reveals distributional effects
- Clustered bootstrap accounts for within-panel correlation
- WTP provides economically interpretable preference measures
- Nonlinear transformations require careful inference (bootstrap or delta method)