ECON6300/7320/8300
Advanced Microeconometrics
Linear regression basics

Christiern Rose

[1] University of Queensland

Practical 2

Introduction

- This class will review:
  - Data management (refer to T2 Introduction to STATA.pdf)
  - Descriptive analysis
  - Linear regression
  - Prediction and analysis of residuals
  - Individual and joint hypothesis tests (Wald, t, F)
  - Regression diagnostics
- We begin with a demonstration from Microeconometrics using STATA Chapter 3 looking at whether private health insurance reduces medical expenditures
- We move on to a practical looking at the gender gap in earnings of Australian clinicians

Demonstration - Introduction (1)

- This follows Chapter 3 of the course textbook
- We analyse data on medical expenditures of individuals aged 65+ who qualify for health care under the U.S. Medicare program
  - The data is from the Medical Expenditure Panel Survey
  - Medicare does not cover all medical expenditures.
  - Around 50% of individuals take out additional private cover to ensure against out-of-pocket expenses.

Demonstration - Introduction (2)

- Question: Is it worth it? By how much does private cover reduce medical expenditure?
  - Need to control for any factors which might determine medical expenditures and the propensity for individuals to take out private insurance
- Apply multiple regression to estimate the treatment effect controlling for observable factors

Practical - Earnings decomposition (1)

- We have earnings data for a sample of Australian GPs from the MABEL survey
- It is well known that female doctors earn significantly less on average than male doctors. We use will use the Oaxaca-Blinder decomposition to decompose the earnings gap

Practical - Earnings decomposition (2)

- The Oaxaca-Blinder decomposition is as follows. For two groups A, B let define the linear model:

$$Y_g = X_g'\beta_g + u_g, \quad \mathbb{E}[u_g] = 0$$

Then it can be shown that

$$\begin{aligned} R &= \mathbb{E}[Y_A] - \mathbb{E}[Y_B] \\ &= (\mathbb{E}[X_A] - \mathbb{E}[X_B])'\beta_B && \text{(Endowment effect)} \\ &+ \mathbb{E}[X_B]'(\beta_A - \beta_B) && \text{(Coefficient effect)} \\ &+ (\mathbb{E}[X_A] - \mathbb{E}[X_B])'(\beta_A - \beta_B) && \text{(Interaction effect)} \end{aligned}$$

Practical - Earnings decomposition (3)

1. Load the data into STATA
2. Describe and summarise the data for the pooled sample, males only and females only. What is the mean difference in earnings between males and females? The median?
3. Determine a suitable dependent variable for a regression which measures annual earnings
4. Regress your dependent variable on yhrs female expr exprsq fellow pgradoth pracsize childu5 visa. Interpret the results. Is there evidence of heteroskedasticity? Is the model correctly specified? Are there any outliers?
5. Perform the regression separately for males and females. Interpret the results.
6. Perform a single regression in which males and females have heterogeneous coefficients. Test equality of the coefficients for males and females.
7. Use the oaxaca command to perform the Oaxaca-Blinder decomposition. You may need to install it first with the line ssc install oaxaca. See also the STATA journal article.