Homework 1

Jasen Lo

SS154

1 Data

Write 1-2 short paragraphs describing the data structure (provide background information about the March CPS and the survey methodology) and a research question you will explore using this data set. For example, the effect of schooling on earnings. Relate your question to the principles of causal inference.

Background Information, Survey Methodology & Data Structure

The Current Population Survey (CPS) is a statistical survey conducted jointly by the U.S. Census Bureau and the U.S. Bureau of Labor Statistics (BLS) and serves as the primary source of labour force statistics for the population of the United States. The CPS is conducted continuously, as data is collected monthly. The March CPS surveys the housed U.S population in March but includes an additional supplementary survey - the March Annual Social and Economic Supplement (ASEC), which provides data income, poverty, and health insurance coverage in the United States. The March CPS is a cross-sectional dataset instead of a panel dataset because the sampled households vary from each successive survey due to the sampling methodology, which uses a probability sample to select the housing units numbered approximately 60,000 required to meet reliability requirements as detailed in Chapter 3 & 4 of the "Current Population Survey, design and methodology" (CPS, 2006).

Research Question

Unemployment compensation in the United States usually pays eligible workers up to $450 per week, with criteria and compensation amounts varying by state (Kagan, 2018). My research question explores the effect of the duration of unemployment on the unemployment compensation received (public assistance or welfare received in the CPS) and if demographic factors such as race are associated with the likelihood of receiving unemployment benefits.

- The variables that I will use in the CPS are:
  - paw_val: Public assistance or welfare value received in 2017.
  - a_wkslk: Duration of unemployment in days.
  - prdtrace: Race.

Importing Data: I cleaned the data by removing those who were not unemployed, or not recieving any public assistance. There were also 8 observations who were unemployed for 99 days, which I also removed for the abnormal nature of these entries:

---

2 Descriptive Statistics

Present summary statistics (mean, standard deviation, sample size etc.) in a few tables and appropriate graphs that describe your data. Choose the 2-3 graphs or tables (tables+graphs=2-3) that best help describe the data.

3 Casual Effect

Write 1-2 short paragraphs about your summary statistics and explain whether it is possible to disentangle any causal effects.

4 Statistical Inference

Finally, carry out a hypothesis test for a population mean using the outcome variable and independent variable. Make sure to show each step (use five components), interpret the conclusion, and append the Stata code.

1. Determine the null and alternative hypotheses.

2. Specify the test statistic and its distribution if the null hypothesis is true.

3. Select and determine the rejection region.

4. Calculate the sample value of the test statistic.

5. State your conclusion.

HC Appendix

References