Problem Set 1
Econ 672
December 20, 2022

Problem 1

From the dataset "calculate_ate_sdo.dta", we will need to calculate the Simple Difference in Outcomes (SDO), the Average Treatment Effect (ATE), The Average Treatment Effect on the Treated (ATT), and the Average Treatment Effect on the Untreated (ATU). We will assume that we know potential outcomes Y^0 and Y^1 (but in reality we never observe both at the same time for the same unit). Our treatment of interest is D and we want to find its treatment effects on Y. Each unit has a unique identifier called "id".

1. Calculate the SDO and ATE. How bias is the SDO from the ATE? Is it upward or downward biased?
2. Calculate the ATT and ATU. Does $ATE = (\pi) \cdot ATT + (1 - \pi) \cdot ATU$, where $\pi = \frac{N_T}{N}$?
3. Calculate the selection bias.
4. Calculate the heterogeneous treatment bias.
5. Does $SDO = ATE + Selection\ bias + Heterogenous\ Treatment\ Bias$?

Problem 2

A small set data from the 2017 4th Quarter of the Public Use Workforce Innovation and Opportunity Act (WIOA) Individual Performance Records can be found on ELMS called "pirl_small.dta". (This file is too large for GitHub). This file has been limited to all individuals with 4 quarters of data after exiting the WIOA programs and particular covariates of interest.

We want to assess the treatment effect of training on earnings. Some individuals receive training, while others do not receive training. One-stop shops assigned training treatment based upon many factors, including individuals with "Significant Barriers to Employment" (SBE). This means that individuals are endogenous sorted into treatment and the independence assumption is violated.

Utilize a propensity score strategy to identify the causal effect of training on earnings. You will need to utilize the public-use data dictionary to code the variables into meaningful covariates of interest. https://www.dol.gov/sites/dolgov/files/eta/performance/pdfs/PY2017/WIOA%20Performance%20Records%20Public%20Use%20File%20Record%20Layout%20PY2017Q4.pdf. (Note: do not run effects before trimming your data because Stata may freeze).

6. Calculate the simple difference in outcomes for training and no-training.
7. Test your covariates of interest to see if they correlate with training.
8. Estimate a maximum likelihood model based on your covariates of interest on training.
9. Predict your estimated propensity score from your model.
10. Provide common support tests from two different methods discussed in class from the propensity scores
11. Trim you data and retest the common support. Do you have common support?
12. What is the estimated ATE and ATT using K-nearest neighbors with teffects?
13. What is the estimated ATE and ATT using Inverse Probability Weights with teffects?