Econ 270 / GSB 603   **Department of Economics**

Fall 2025   **Stanford University**

PROBLEM SET III

DUE: SUNDAY, OCTOBER 12th, 2025, 6PM.

Be concise but clear as to what numbers you are reporting, and answer in full sentences. You should also hand in supporting code, but all answers should be in a PDF or Word document.

Use the experimental data from the Riverside GAIN experiment data posted on the course website under `riverside_2025.txt`. There are eight variables, the latter five of which are pre-treatment variables:

- the binary treatment indicator (1 for the individuals who were assigned to the training program, and 0 for those who were not);
- earnings one year after the program, in thousands of dollars;
- earnings four years after the program, in thousands of dollars;
- a binary indicator for high school graduation;
- a binary indicator for being female;
- age, in years;
- a binary indicator for having at least one child under the age of six; and
- a binary indicator for being single.

There are 5,419 observations in the sample. We are interested in analyzing the effect of the program on the two outcome variables (earnings one and four years after the program).

1. (a) In the last problem set you estimated the average treatment effect and constructed a 90% confidence interval for the average effect. Construct a new 90% confidence interval using the plain vanilla boostrap.

   (b) Construct a 90% confidence interval using the Bayesian bootstrap.

   (c) Construct a 90% confidence interval by bootstrapping the t-statistic, using the 0.05 and 0.95 quantiles of the bootstrap distribution of the t-statistic. Report these quantiles.

   (d) Construct a 90% confidence interval using the subsampling with the square root of the sample size as the subsample size.

   (e) Estimate the average effect of the treatment adjusting whether individuals have

a high school degree or not. Motivate your choice of estimator and report a 95% confidence interval based on your choice of bootstrap. Describe in detail how you did the bootstrap.

2. Theoretical Calculations

    (a) Suppose you do an experiment on a population of 100 individuals, 50 old and 50 young. You randomly pick 50 individuals to be in the treatment group. You selected 27 old and 23 young for the treatment group. You have the choice between two estimators. One the difference in means between the treatment and control group, and the average of the difference between treated and controls in the old and the difference between the treated and controls in the young group. Which estimator would you choose and why?

    (b) Show that both estimators are unbiased for the average effect of the treatment.