Econ 272 / GSB 607 **Department of Economics**

Spring 2025 **Stanford University**

Problem Set II

Due: Sunday April 13th, 2025, 11pm.

Be concise but also clear what numbers you are reporting, and answer in full sentences. You should also hand in supporting code, but all the answers should be in a pdf or word file.

1. Load the data set census.mat, which has three variables, an outcome (log weekly earnings), a year of birth indicator, and an indicator for education being more than 12 years (the treatment we focus on). This is based on the public use part of of the census (a random sample from the full census).

   (a) Estimate the average effect of the education indicator on the outcome, and report the Neyman variance estimator.

   (b) Report bootstrap standard errors for the estimator in part (a).

   (c) Report the Liang-Zeger cluster-robust standard error based on clustering on year of birth.

   (d) Report the cluster bootstrap standard error based on bootstrapping the clusters.

   (e) What is the within year-of-birth correlation between log earnings (the outcome)?

   (f) What is the within year-of-birth correlation between the education indicator (the treatment)?

   (g) Is there another standard error you would report here?

2. Theoretical Calculations

   (a) Which standard error do you think should be reported for the average effect of the treatment?

   (b) Explain your choice.