

Econ 270 / GSB 603

Fall 2025

Department of Economics

Stanford University

PROBLEM SET VI

DUE: TUESDAY, NOVEMBER 18th, 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.

1. Use the data from the Angrist-Krueger paper, with observations on education, log earnings, state of birth and year of birth.

- (a) Estimate the model

$$Y_i = \alpha + \beta X_i + \varepsilon_i$$

by least squares, where Y_i is log earnings and X_i is years of education.

- (b) Calculate the OLS standard errors.
- (c) Calculate the clustered standard errors.
- (d) Calculate both OLS and clustered standard errors by bootstrapping.
- (e) Estimate by least squares the model

$$Y_i = \alpha + \beta_1 X_i + \beta_2 \bar{X}_{C_i} + \varepsilon_i,$$

where \bar{X}_c is average years of education for all individuals in state c .

- (f) Calculate the OLS standard errors for both β_1 and β_2 .
 - (g) Calculate the clustered standard errors.
 - (h) Calculate both OLS and clustered standard errors by bootstrapping.
2. Theoretical Calculations

- (a) Consider the model

$$Y_i = \alpha + \beta X_i + \eta_{C_i} + \nu_i,$$

with the cluster components η_c independent with $\mathcal{N}(0, \sigma_\eta^2)$ and ν_i iid with $\mathcal{N}(0, \sigma_\nu^2)$ distributions. Suppose that $N_c = N/C$ for all clusters. Suppose that the X_i are constant within clusters. Show that the GLS estimator for known σ_η^2 and known σ_ν^2 has the same variance as the OLS estimator.