

PROBLEM SET 3: EFFECTS OF HEAD START

DUE BY 11:59 PM ON FRIDAY 2/21

Head Start is an early-childhood development program run by the U.S. federal government. It provides health, nutrition, and education services to children from disadvantaged backgrounds.

The dataset contains a sample of children of NLSY '79 participants, some of whom participated in Head Start. All of the sample children have at least one sibling also in the sample. The variables are ordered as follows:

- *head_start* - *sibdiff* relate to head start participation.
- *mom_id* - *lnbw* were determined prior to Head Start participation. (Note: the PPVT is an early-childhood cognitive test.)
- *comp_score_5to6* - *comp_score_11to14* correspond to test scores in childhood.
- *repeat* - *fphealth* deal with outcomes in the teenage years and young adulthood.

For various reasons, some variables have missing data. You may decide how to deal with missing values on your own.

1. Summarize the data. What can you say about the backgrounds of children who participated in Head Start?
2. As a first step, estimate the association between Head Start participation and age 5-6 test scores using OLS. Make sure you estimate standard errors correctly. If we assumed Head Start participation is exogenous, what would we conclude about the effects of Head Start on test scores? Be sure to explain the magnitude of the estimated effect. Is it reasonable to assume that Head Start participation is exogenous?
3. Now estimate the same association using a random effects model (with mother random effects). How do the results compare with OLS? Does the comparison make you more or less confident that OLS or random effects can shed light on the causal effect of Head Start on test scores?
4. Now estimate a mother fixed effects model. Run regressions both with and without pre-Head Start control variables. Which control variables can you include, and which can't you include? Why? What do the results imply about the effects of Head Start on test scores? If the fixed effects results are different from those in your answer from question (2), explain why.

5. Some advocates for early-childhood education suggest that the effects of programs like Head Start are long-lasting. Carry out fixed effects analyses of test scores at later ages. Does Head Start participation have similar effects on test scores in later childhood, or do the effects fade out with age? Make sure you compare results using comparable test-score units. (Hint: you can convert test scores to standard deviations.)
6. Estimate fixed effects models of the effect of Head Start on longer-term outcomes besides test scores. Many of these outcomes are binary, but you may use linear models. Interpret your results.
7. Do the effects of Head Start participation on longer-term outcomes vary by race/ethnicity? By sex?
8. The Obama administration advocated expanding federal funding for early-childhood education programs, while the Trump administration has argued for cuts. Based on your results, which position seems better supported by evidence? Would you feel comfortable using your results to predict the effects of such an expansion? Why or why not?