Fixed Effects Homework

Jennifer Hill, Ray Lu & Zarni Htet

Objective

The goal of this exercise is to simulate data that might be used to estimate a causal effects using a fixed effects regression.

Setting

The hypothetical setting is similar to the twins study discussed briefly in class. We are interested in estimating the effect of birthweight on future outcomes such as IQ. However we are concerned that birthweight is correlated with other characteristics of the children that are also associated with IQ. In particular, here we suppose that the mother's IQ is the sole confounder when identifying the effect of birthweight on subsequent IQ.

Question 1. God role: simulate data.

To explore what's going on here we want to simulate both the kind of data that the researchers for that study had available to them and the data that gods would have available.

Simulate data for 1000 observations (500 pairs of twins) for the following variables: index for the mother (momindex), child birth weight in thousands of grams (W), child IQ at age 5 (Y), mother's IQ at time of birth (X).

The data should adhere to the following model

$$X_k \sim N(100, 15^2) W_{ik} \mid X \sim N(3 + .02(X_{ik} - 100), .5) Y_{ik} | X_{ik}, W_{ik} \sim N(100 + .05(X_{ik} - 100) + 0W_{ik}, 5^2)$$

where k indexes mothers and i indexes children within mothers. Thus X_{ik} represents the appropriate X_k value for the ith child born to the kth mother (this is the same for both children of the same mother).

Question 2: God role

a) What is the effect of birthweight on IQ at age 5?
b) What role does X play in this scenario?

Question 3: Researcher role

The researcher only has available to her the mother ID, the birth weight of the child, and the IQ of the child at age 5. Given this what might a researcher to do estimate the effect of birth weight on child IQ? Try two appropaches: a) A regression of child IQ on birth weight. b) A regression of child IQ on birth weight controlling for mom fixed effects.

Question 4: God role

Comment on the difference between these estimates and the truth.

Question 5: Assumptions

a) How would you formalize the ignorability assumption required by the fixed effects strategy for identifying the causal effect?
b) Besides ignorability, what other assumptions are built into this fixed effects strategy?