Homework for Chapter 16: Fixed Effects

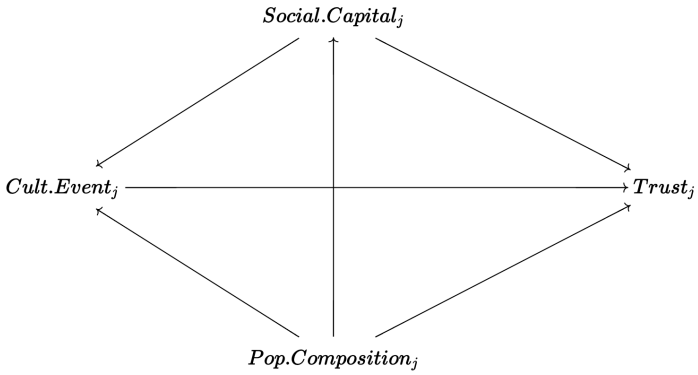
How Does It Work?

1. You observe the number of vacations taken by Zac and Skylar in 2012, 2013, and 2014. In those years, Zac took 3, 7, and 5 vacations, respectively. Skylar took 2, 6, and 10.
    a. Isolate the numbers that represent the variation *between* Zac and Skylar in their vacation-taking.
    We get the between-variation by comparing individual means. In this case, the mean for Zac is 5, and 6 for Skylar, so the variation between them is 1.
    b. Isolate the variation *within* Zac and within Skylar in their vacation-taking.
    For the within variation, we compare each individual to their mean. That is 2 + 2 + 0 for Zac and 4 + 0 + 4 for Skylar (in absolute numbers).
    c. (Difficult!) We perform a fixed effects analysis of the effect of vacations on happiness. A vacation increases Zac's happiness by 1 "happiness point," but it increases Skylar's happiness by 2 "happiness points." Will our fixed effects estimate likely give us an answer closer to 1, closer to 2, or exactly 1.5?
    Since regression is a variance-weighted estimator and there is more variation in Skylar than Zac (double the amount) the estimate will be between higher than 1.5 and lower than 2.
2. You are interested in the effect of cultural events on the levels of trust in a city. Perhaps big events like concerts bring people together and they can trust each other more. You plan to look at the relationship between trust and the number of events in a given year, with fixed effects for city. Draw a causal diagram for this research question with at least four back door paths. Which paths will be closed by fixed effects, and which will remain open?

In the DAG I included two variables, the amount of social capital (in the sense of Putnam) and the population composition (e.g. as measured by a ratio between natives and first or second-generation immigrants, or age composition). These two confounders open four back door paths:

- Pop.Com -> Cult.Events -> Trust
- Pop.Comp -> Soc.Captial -> Cult.Events -> Trust
- Soc.Capital -> Cul.Events -> Trust
- Soc.Captial <- Pop.Comp -> Cult.Events -> Trust

Because there are no colliders we could condition on social capital and population composition to get the direct effect of cultural events on trust. But assuming these are unobservable and we want to close these back door paths using city-level fixed effects we would have to consider whether any of these confounders are likely to change over the time period under consideration. I would argue the answer is no. My prior would be that "trust" in population changes by cohort replacement. Thus, any effect of cultural events would take time to materialize in substantive changes in trust. The population composition and social capital might follow a similar pace of change. If that were true, fixed effects would not help much. If, instead, we assume that a particularly important cultural event can have a quick substantial impact on trust then using fixed effects would close those back doors.



Classify each of the following forms of variation as "between variation", "within variation", or a combination of both.

a. (Individual = person) How a child's height changes as they age. Within.
b. (Individual = person) In a data set tracking many people over many years, the variation in the number of children a person has in a given year. Within.
c. (Individual = city) Overall, Paris, France has more restaurants than Paris, Texas. Between.
d. (Individual = genre) The average pop music album sells more copies than the average jazz album Between.
e. (Individual = genre) Miles Davis' Kind of Blue sold very well for a jazz album. Within.
f. (Individual = genre) Michael Jackson's Thriller, a pop album, sold many more copies than Kind of Blue, a jazz album. Between.

3. Why does the process of taking each observation relative to its individual-level mean have the effect of "controlling for individual"?

   Standardizing is a way to compare values across variables with different scales. The comparison then becomes a relative as opposed to an absolute one, i.e. we compare changes in a variable relative to the mean (and standard deviation) of that variable. Subtracting individual means for each observation does the same thing as mean-centering. It treats individuals as variables with a distribution centered on their mean. Comparing between individuals then becomes a relative comparison.

   How is it Performed?

   By subtracting individual means from each value.

4. You are interested in the effect of cultural events on the levels of trust in a city. You run a regression of trust levels (on a 0-100 scale) on the number of cultural events with city fixed effects and get a coefficient on cultural events of 3.6. Assume that there are still some back doors open, so do not interpret the result causally. Interpret the 3.6, explaining it in an English sentence.

For a particular city, an additional cultural event is on average related to trust levels that are 3.6 points higher than the city mean.

5. You are interested in the effect of cultural events on the levels of trust in a city. You run a regression of trust levels (on a 0-100 scale) on the number of cultural events with city and year fixed effects and get a coefficient on cultural events of 2.4. Assume that there are still some back doors open, so do not interpret the result causally. Interpret the 2.4, explaining it in an English sentence.
   For a given city and year an additional cultural event is related to a trust levels that are, on average, 2.4 points higher than the expected value of that city-year pair.

6. Two-way fixed effects with terms for both individual and time are often referred to as "controlling for individual and time effects". Why might a researcher want to do this rather than just taking individual fixed effects and adding a linear/polynomial/etc. term for time?

   Including fixed effects for time is simply a more flexible way to describe the functional form of the relationship between time and the outcome. It also adds more complexity (parameters) to the model, so it is a trade-off. If we think that that relationship is linear or polynomial we should include time as a linear effect, if we don't we can use fixed effects. In general, using fixed effects is more conservative than a linear effect so we might prefer them for that reason.

7. Which of the following explains why random effects are likely to do a better job of estimating the individual-level effects than fixed effects, if its assumptions hold?
   a. Because it makes the assumption that the individual effects are unrelated to the other predictors, which breaks that back door and thus reduces bias.
   b. Because random effects allow some amount of between variation into the model, and some of the real individual effect is that between variation.
   c. Because it uses the information from the entire data set to estimate each individual effect, rather than relying on only a few observations per individual.
   d. It won't. Enforcing Durbin-Wu-Hausman makes both methods produce the same estimates anyway.

Coding (which includes any how-the-pros-do-it questions)