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.
   1. Isolate the numbers that represent the variation *between* Zac and Skylar in their vacation-taking.
   2. Isolate the variation *within* Zac and within Skylar in their vacation-taking.
   3. (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?
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 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?
3. Classify each of the following forms of variation as “between variation”, “within variation”, or a combination of both.
   1. (Individual = person) How a child’s height changes as they age.
   2. (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.
   3. (Individual = city) Overall, Paris, France has more restaurants than Paris, Texas.
   4. (Individual = genre) The average pop music album sells more copies than the average jazz album
   5. (Individual = genre) Miles Davis’ Kind of Blue sold very well for a jazz album.
   6. (Individual = genre) Michael Jackson’s Thriller, a pop album, sold many more copies than Kind of Blue, a jazz album.
4. Why does the process of taking each observation relative to its individual-level mean have the effect of “controlling for individual”?

*How is it Performed?*

1. 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.
2. 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.
3. 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?
4. Which of the following explains why random effects is likely to do a better job of estimating the individual-level effects than fixed effects, if its assumptions hold?
   1. Because it makes the assumption that the individual effects are unrelated to the other predictors, which breaks that back door and thus reduces bias.
   2. Because random effects allows some amount of between variation into the model, and some of the real individual effect is that between variation.
   3. 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.
   4. 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)