Analysis Challenge 8

Difference-in-Differences

Using the Congressional data below ([read more about the data](https://github.com/milesdwilliams15/Teaching/blob/main/DPR%20201/Data/Files%20for%20Data%20Exercises/13_README.txt)), compare the estimates of the effect of Republican party membership on conservative voting behavior using three different designs:

1. Selection on observables or SOO
2. Regression discontinuity or RD
3. Difference-in-differences or DD

Here’s the code to read in the data and open the packages you’ll need:

library(tidyverse)  
library(estimatr)  
congress\_dt <- read\_csv("https://raw.githubusercontent.com/milesdwilliams15/Teaching/main/DPR%20201/Data/CongressionalData.csv")

The outcome variable is cvp. This stands for “conservative vote probability” and is a score given to each member of Congress in each term summarizing their roll-call voting. It meausres the probability that each member voted conservatively across all bills relative to the median member in that session. That means it is a *relative* probability and can take either positive or negative values. Postive means that a member voted more conservatively than the median member of Congress in a given term.

The causal variable of interest is republican which is 1 if the member was a Republican and 0 if they were a Democrat.

For the SOO design, you’ll estimate a multiple regression model that controls for rep\_twoparty\_pres. This variable is the average Republican two-party vote share for US President in congressional districts from 2004 and 2008. You can think of it as a proxy (indirect measure) of the political ideology of a district.

For the RD design, you’ll use rep\_voteshare as your running variable. Note that you’ll need to adjust for the fact that the discontinuity is at the 0.5 vote share value.

Finally, for the DD design, you can follow the example from our Monday lecture notes.

After you estimate all three designs, compare your estimates of the effect of party membership on member voting. Do these research designs provide similar or very different estimates?

**1 extra bonus point if you compare the results using a coefficient plot!**

Upload a knitted .Rmd file to Canvas that includes your code and comparison of the results.