Definition of the ATT = E[Y(1)|D=1, Post] – E[Y(0)|D=1, Post]

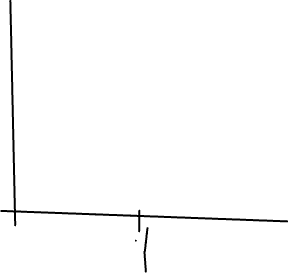
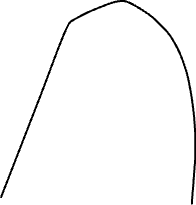
“Average Y(1) minus Average Y(0) *for the treatment group only in the post treatment period only”.*

This is all diff-in-diff can **ever at best estimate.**

Variance of treatment

Dbar(1-Dbar)

0.25



Share of time treated *in your dataset*

*Dbar=0.5*

When a treatment group is treated in the middle (0.5) of your panel dataset, it will get the largest weight of any group in the data *regardless of its estimated treatment effect.* OLS does that – OLS creates its own weights in order to minimize the sum of squared residuals to become the best linear predictor of the outcome. But what it is doing here is that it’s weighting up or down a group in that final TWFE calculation for no other reason than where it is in the data.

Google sheets for TWFE:

https://docs.google.com/spreadsheets/d/1onabpc14JdrGo6NFv0zCWo-nuWDLLV2L1qNogDT9SBw/edit?usp=sharing