

Assignment 13: DiD

Megha Joshi

Part 1: Identifying the Scenario

I want to examine the effect of [House Bill 5](#) that was passed by Texas in 2013 which involved changing requirements for graduation. The requirements included not requiring high school students to take Algebra II and requiring students to take college and career preparatory classes. The policy was implemented for all Texas schools. I want to examine the effects of this policy on high school graduation, college enrollment, persistence and completion, and taking developmental courses in college.

It would be infeasible to randomize some high schoolers to take Algebra II in high school and some not to for example. There would be a lot of compliance issues because Algebra II is a fundamental course. Districts, parents, and students may not agree with randomizing students. It may also be difficult to convince students who are in need of college and career prep courses to be randomized to not receive it. And because this policy is about high school graduation requirements and outcomes are long-term, we cannot offer control students delayed treatment either. Furthermore, this is a state-level policy where everyone gets the treatment. Any RCT, before the actual policy gets implemented, would have been done on a much smaller scale, making external validity questionable.

Part 2: Designing Observational Study

1. Counterfactual group identification

1. The counterfactual groups will be another state (Oklahoma) similar to Texas that did not implement any policy changes regarding graduation. Oklahoma is a different state though. Ideally we would have part of Texas as counterfactual. The population may be different and behave in different ways over time.
2. For Texas, we will use the state longitudinal data housed in the [Education Research Center](#). This database tracks students from elementary school to high school through college and workforce. For the control group, we will try to find similar datasets. If we cannot find student-level data, perhaps we can use district level data with outcomes being graduation rate, college enrollment rate etc. We will collect

data for multiple year (2008 to 2018 cohorts with the treatment being implemented in 2014-2015 year)

3. One key assumption is that both treatment and control group follow similar trends before the intervention. After the policy is implemented, we assume treatment group would follow similar trend as control group if the policy had not been implemented, and the difference between treatment and control group would be same as before the policy. Any difference in the trend is due to the treatment. If the trend is not parallel, then we cannot calculate difference in difference and estimate the causal effect. To examine parallel trends we will plot raw output data for the treatment and comparison groups. However we can't test if the potential outcome of treated group if the policy had not implemented keeps that trend from pre intervention. For my particular example, we can try to extrapolate from data from past - to see if graduation rates, college enrollment rates etc. stay stable over time.