

PubPol 713 Reading Questions 7

Nathan Mather

February 5, 2019

1 Question 1

I find the set up in the Scott-Clayton to be more applicable to what we could use to estimate the effect of the Michigan compulsory attendance law on student outcomes. My understanding is that we have Michigan data only and are investigating only the impact of the law change in Michigan. This more closely matches the set up of Scott-Clayton since Fitzpatrick and Jones have multiple states with multiple treatments. Scott-Claytons primary estimation strategy on page 627 does not include any control group. So, we could estimate the equation comparable to (2)

$$y_{ij} = \alpha + \beta(after_t) + X_i\delta + v_t + \epsilon_{it}$$

Where y_i is any outcome of interest. Perhaps years of schools, college attendance, test scores, or earnings. $after_t$ is an indicator for if a cohort is impacted by the law change or not.

We could also set up a comparable IV estimation to use the compulsory schooling law as an instrument for estimating the returns to a high school education on those that are compelled to stay. We would estimate

$$YS_{it} = \lambda + \gamma(after_t) + X_i\phi + \eta_t + \mu_{it}$$

$$y_{it} = \alpha + \beta(\widehat{YS_{it}}) + X_i\delta + v_i + \epsilon_{it}$$

Where YS_{it} is years of schooling. I we receive data on a valid control group after the law is in place we could also run a proper diff in diff like Scott-Clayton does as a robustness check on age 631. Such a

control group could potentially be students from schools along the boarder just outside Michigan. There are issues with this comparison but it would also be a good robustness check. We could estimate

$$y_{it} = \alpha + \gamma(after_t) + \psi(after_t * MI) + \phi(MI) + X_i\zeta + \eta_t + \epsilon_{it}$$

2 Question 2

I am still a bit confused about what exactly our goal is going to be in the Michigan Study. Are we trying to determine the impact of the law on student outcomes, or use the law to determine the impact of schooling on outcomes? I am going to go with the second but a regression discontinuity could work with either.

We could set up a regression discontinuity with date of birth as the running variable. The treatment variable would be years of schooling. We would expect the discontinuity to occur around December 1st of 1998 as these students are the first to have the law applied to them. This will be fuzzy since not all students impacted by the law will actually attend school longer. However, there may be a significant jump in years of schooling around this time. The outcome variable could be something like earnings later in life. We could use the following set of equations

$$YS_{it} = \lambda + \psi(above_i) + \gamma(DOBdist_i * below_i) + \phi(DOBdist_i * above_i) + X_i\psi + \epsilon_i$$

$$y_{it} = \alpha + \beta(\widehat{YS}_{it}) + \zeta(DOBdist_i * below_i) + \pi(DOBdist_i * above_i) + X_i\delta + \epsilon_i$$

Where $above_i$ and $below_i$ indicate that a student is above or below the birthrate cutoff respectively. $DOBdist_i$ indicates the distance of student i's date of birth from the cutoff. YS_{it} is years of schooling and y_{it} is an outcome of interest like earnings. X_i is a vector of controls.