ECON 7130 - Microeconomics III

Spring 2016

Notes for Lecture #5

Today:

- Difference-in-Differences (DD) Estimators
- Difference-in-Difference-in-Differences (DDD) Estimators (Triple Difference)

Difference-in-Difference Estimation

- Main idea of using Difference-in-Differences:
  - Look at effects of treatment by comparing two groups, before and after treatment
  - Like individual FE, difference out time invariant heterogeneity - this time between treatment and control groups
  - Individual FE like DD, but difference means of same unit over time
  - DRAW GRAPH with treatment and control groups with trends - show non DD and DD estimates
- Basic, non-regression setup:

$$\text{Treatment effect} = (\bar{y}_{B,2} - \bar{y}_{B,1}) - (\bar{y}_{A,2} - \bar{y}_{A,1})$$

  - where, B is treated group
  - A is control group
  - Treatment happens in period 2
- For DD stuff, can usually do a lot with just analyzing means - regressions just add control for other covariates that may confound results
- Key assumptions:
  - "Common trends" - absent treatment, treatment and control groups would have continued pre-treatment trends
    - Can test DD using data from more periods and plot the two time series to check parallel trend assumption
  - Treatment and control groups are comparable - differences between two would remain same, absent treatment
    - This is where judgement is used - will need to be able to argue that control is appropriate
    - Use alternative control groups [not as convincing as potential control groups are many]
    - Compare to FE - same person - groups can change composition (e.g. selection into treatment)
    - Has lead to "sharper" estimators like RD - more likely thought to give better control group
- Causality - DD gives ideal setting for Granger Causality type test
  - Include leads and lags of dummies for time of treatment
  - e.g., $Y_{it} = \gamma_s + \lambda_t + \underbrace{\sum_{\tau=0}^{m} \delta_{-\tau} D_{s,t-\tau}}_{\text{Lagged effects of treatment}} + \underbrace{\sum_{\tau=1}^{q} \delta_{+\tau} D_{s,t+\tau}}_{\text{Anticipatory effects of treatment}} + X_{ist}'\beta + \varepsilon_{ist}$
    - If treatment causal, then dummies for leads should show no effect of treatment

- SE
  - Binary covariates define groups within which errors are potentially correlated (e.g., cities, states, years, states after treatment)
  - Thus often want to use clustered standard errors
    * Good practice is to try clustering over different groups (e.g. state or time) and report the clustering that gives the largest standard errors
    * This a more serious problem if the number of periods in data is large (because more serial correlation)
- Stata: `reg` with dummies for post treatment, treatment group, and interactions of treatment group and post treatment dummies

<u>DD: Example 1, Finkelstein, "*EZ-tax*: Tax Salience and Tax Rates" (*QJE*, 2009):</u>

- Idea: Test of tax salience.
  - Do less salient taxes have as big an impact on demand?
  - Do less salient taxes result in bigger government?
- Natural experiment: the introduction of electronic toll collection (ETC) on U.S. roads, bridges, and tunnels
- Questions:
  - Does the introduction of ETC mean road use less sensitive to tolls?
  - Does the introduction of ETC result in more increases in tolls? (i.e., larger government)
- Data:
  - Hand collected data on tolls, traffic, and the timing of introduction of toll collections in 123 of the 183 sites with tolls in place in 1985.
  - She conducts her own survey of Mass Pike drivers and uses a survey of NY/NJ commuters
- Basic Model:
  - $y_{it} = \gamma_t + \beta_1 \text{ETC Adopt}_{it} + \beta_2 \text{ETC}_{it} + \varepsilon_{it}$, where
  - $y_{it} = \Delta log(\text{Minimum Toll})$
  - ETC=1 if electronic in place in year t
  - ETC Adopt=1 if adopted this year
  - γ_t are year controls
  - Since dep variable is in differences, year controls capture average growth rates by year, and the β's capture deviations from those growth rates.
- Coefficient of interest is β_2 - how tolls increase with ETC
- Identification:
  - Difference-in-Differences: comparing changes across areas with and without ETC.
  - Key assumption: ETC are exogenous
    * Places with ETC implemented are on same trend line as places w/o ETC
    * ETC implementation is not correlated with changes in toll setting relative to its norm.

- Does some analysis with collection location FE → ID off changes within location (but not in main spec)

- Results:
  - Elasticities of toll use smaller in presence of ETC (but small elasticities anyway)
  - Evidence is compelling that tax rates rise when less salient tax is created (tolls rise with ETC introduction)
    * Installing ETC leads to 75% more increase in tolls
    * Idea: once you use electronic payment you no longer pay attention to the toll amount
  - Political economy result:
    * Baseline assumption is that legislators do not want to increase taxes in election years.
    * If less salient taxes do not change behavior (people are less aware of the tax changes) then there should be less of an election year effect with the less salient tax
    * Indeed, under ETC there is less of an election year effect. Tolls less sensitive to electoral cycle when ETC in place.

DD: Example 2, Eissa, "Taxation and Labor Supply of Married Women: The Tax Reform Act of 1986 as a Natural Experiment" (NBER WP, 1995):

- Never published (not sure why), but great teaching paper and also very influential paper
- Question: How responsive are married women to changes in tax rates?
- Natural experiment: Tax Reform Act of 1986 (TRA86)
  - Lowered marginal rates for many - especially high income
  - Reduced the number of tax brackets
- Data: March CPS: 1984-1986 and 1990-1992 (TRA86 phased in by 1988)
- Basic Model (employment): $P(LFP_{it} = 1) = \alpha_0 + \alpha_1 X_{it} + \alpha_2 High_{it} + \alpha_3 Post86_t + \alpha_4 (High * Post86)_{it}$
  - $High_{it} = 1$ if in the 99th percentile
  - X_{it}= age, educ, # kids, young kids, race, central city, year & state fixed effects
  - α_4 is main coeff of interest - effect of TRA86 on work incentives
- Identification:
  - DD - comparing LFP differences between high and low income groups before and after TRA86
  - Control groups are 75th and 90th percentile (observations of people with income within +/- 5 percentage points of these percentiles)
    * Tradeoff: 90th better control but gets some treatment
  - Key assumptions:
    * TRA86 exogenous to female labor supply (unlikely a problem here)
    * Treatment and control (high and lower income) have similar employment trends absent TRA86
      · May not be a good assumption
      · Really depends how close you think two groups are (more power couples now?)
- Results:
  - Large response for participation, less for hours

- Consistent w/ lit showing greater responsiveness on participation margin than hours margin (Mroz, Hausman)

Difference-in-Difference-in-Difference Estimation

- Main idea of using Triple Difference:
  - Difference out trends that may differentially affect treatment and control groups in DD estimator
  - Kind of like a robust DD - if not different, then it's like a robustness test
- Why use DDD?
  - In principle, can create a DDD as the difference between actual DD and placebo DD (DD between 2 control groups).
  - However, DDD of limited interest in practice because
    1. If $\text{DD}_{placebo} \neq 0$, DD test fails, hard to believe DDD removes bias
    2. If $\text{DD}_{placebo} = 0$, then DD=DDD but DDD has higher s.e.
- Stata: Same as DD, just more interactions. DDD estimate will be a triple interaction

DDD: Example 1, Chetty, Looney, and Kroft, "Salience and Taxation: Theory and Evidence" (*AER*, 2009):

- Question: Does the effect of a tax depend upon whether it is included in the posted price?
- Experiments (we'll focus on the first):
  1. Field Experiment: Post tax inclusive prices on large number of products in grocery stores
  2. Natural Experiment: Excise tax included in posted price, sales tax not. Variation in taxes across states and time.
- Data:
  - 750 products (3 product categories)
  - 3 week period for treatment
  - Large grocery store, national chain
  - 2 control stores
  - Scanner data from all three stores over 65 week period
- Basic Model: $Y = \alpha + \beta_1 TT + \beta_2 TS + \beta_3 TC + \gamma_1 TT * TC + \gamma_2 TT * TS + \gamma_3 TS * TC + \delta TT * TC * TS + \xi X + \varepsilon$ where:
  - Y = log sales
  - TT = Treatment Time
  - TS = Treatment Store
  - TC = Treatment Category
  - δ is coeff of interest and equals DDD estimate using means if no covariates included
- Identification:
  - DDD - changes in demand for treated products relative to changes in demand for untreated products

    - DD_{TS} = -2.14 units is the "within treatment store" difference-in-difference estimate of the impact of posting tax inclusive prices
    - DD_{CS} = -0.06 units is the "within control store" DD of the sales trends of the treatment and control categories (statistically zero - validates assumption of common trends)
    - $DDD = DD_{TS} - DD_{CS}$ - within store and within product trends are difference out
    - Nicely done with analysis of means and then DDD with regression
  - Key assumptions:
    - For DD: Common trends: sales of the treatment and control products would have evolved similarly absent our intervention
    - For DDD: no shock during our experimental intervention that differentially affected sales of only the treatment products in the treatment store

- Results:
  - Result is that consumers seeming under-respond to taxation.
    - Making sales tax more salient reduced demand by 7.6%
    - A 10 percent tax increase reduces demand by the same amount as a 3.5 percent price increase
  - This lack of response implies that taxation is less distortionary that it would be if agents fully responded.

DDD: Example 2, Ravallion, Galasso, Lazo, and Philipp, "What Can Ex-Participants Reveal about a Program's Impact?" (*JHR*, 2005):

- Question (Methodological): Can we see impact of treatment even if we don't observe a pre-treatment period?
- Idea:
  - Match initial participants with non-participants (to get rid of effects of selection into program)
  - Match leavers and stayers (to get rid of selection effect of staying in program)
  - Calculate the DDD using the DD between matched stayers and leavers
  - Propensity score matching helps with selection on observables
  - DDD helps with selection on unobservables
- Experiment: Argentina's Trabajar Program: gov't work program for poor, unemployed
- Data: Survey of Trabajar Participants, Permanent Household Survey (twice yearly, cross-section)
- Identification:
  - Matching - assuming results in comparable groups for participation/not
  - Assumption that earnings trends similar for those who did and didn't drop out (absent difference in treatment) (I think)
- Triple Difference Estimator
  - DD_{treat} = Change in continuing participants' outcome - Change in ex-participants' outcome
  - $DD_{control}$ = Change in control group matched to participants - Change in control group matched to ex participants
  - $DDD = DD_{treat}$ - $DD_{control}$ = impact of program participation
  - Propensity Score Matching (PSM) controls for heterogeneity based on observables

- DD estimates control for heterogeneity based on unobservable differences in treatment and control groups
- Still lacking control for unobs heterogeneity that cause selection into program, but authors argue they can sign this bias

- Results:
  - Trabajar Program had significant impact on workers' earnings