MEI LT Problem Set 3[1]

Timothee Carayol

February 11, 2010

[1] Available on http://personal.lse.ac.uk/carayolt/ec402.htm

Overview

- Today's exercise covers fixed effects, Difference-in-Differences and Propensity Score Matching estimators.
- Three ways to deal with non-experimental situations (i.e. situations where we do not have a randomly assigned comparison group to which we can compare the outcome of the treatment group).

Overview

- Widely used benchmark: the NSW data.
  - Was a randomized experiment in the 70's (even though those were quite rare at the time).
  - Provided work experience to recipients with weak labour force attachment.
  - Several microeconometricians tried to replicate the results *without* using the experimental control group, but instead trying to construct another (nonexperimental) control group from national datasets.
  - Lalonde (1986), Dehejia and Wahba (1999 and 2002), Smith and Todd (2001), Angrist and Pischke (2009) used this data to evaluate the performance, e.g., of propensity score matching.
  - We will do the same thing today.

Question 1

- $Y_{it} = x'_{it}\beta + \epsilon_{it}$.
- Assume ϵ_{it} comprises a time-specific part δ_t and and error term μ_{it} such that $E(\delta_t) = 0$ and $E(\mu_i t | x_{it}, \delta_t) = 0$.
- Essentially:
  - Random effect is *GLS* using ϵ_{it} as the error term and using its decomposition to find Ω. Problem: for *GLS* to be consistent, we need some sort of *A*3, i.e. need to have δ_t uncorrelated with x_{it} (let's call this A3RE here).
  - Fixed effect identifies the effect *within* time period (here) or individual (in most cases involving panel data). Then we do *not* need the time-specific error term to be uncorrelated with the observables: in effect, we *control* for it.
- So: FE consistency requires fewer assumptions than RE. Can be the case that FE is consistent but not RE; but not the opposite. However, if RE is consistent, then it is efficient, contrary to FE.

Question 1 - A

- To check whether δ_t is indeed uncorrelated with x_{it} (A3RE), i.e. whether RE estimation is consistent: can run both RE and FE and compare estimates.
- If A3RE is correct, both RE and FE will be consistent, so estimates should not differ "too much".
- There is a formal test to see how much is "too much": the Hausman test. (This test can be used in any context where we have two estimators which are both consistent under H_0, one of which is efficient under H_0.)
- Can show that under $H_0 : A3RE$, $H = (\beta_{RE} - \beta_{FE})'(a\hat{V}ar(\beta_{FE}) - a\hat{V}ar(\beta_{RE}))^{-1}(\beta_{RE} - \beta_{FE}) \rightarrow_d \chi^2(K)$.

Question 1 - B

- Difference from mean: substract the averaged (by time period) equation from the original equation.
- Averaged: $\overline{Y_t} = \overline{x_t}'\beta + \delta_t + \overline{\mu_t}$.
- Difference from the mean: $Y_{it} - \overline{Y_t} = (x_{it} - \overline{x_t})'\beta + (\mu_{it} - \overline{\mu_t})$.
- This is then a well behaved model which we can estimate consistently with OLS.
- Keep in mind that when we use differences from the mean, we lose all the "between" variations: we only look at variations "within" a time period (here) or an individual (usually with panel data). RE, by contrast, uses all the variations.
- For it to make sense to difference from the mean, our initial specification needs to be correct; in particular:
  - The β's must not differ accross individuals or time.
  - The time-specific error component must take the same value for all individuals at each t.

Question 1 - C

- How to check whether β does not vary with t? (i.e. the effects of the different variables do not change over the time)
- Can run one regression for each t (or one big regression allowing for different coefficients for different periods) and then test whether $\beta_1 = \beta_2 = .. = \beta_T$.

Question 2 - Difference in Differences

- A - Checking whether the experimental assignment really was random.
  - We cannot reject equality of the observables between treatment and control group for most variables.
  - Not so clear for the “no degree” variable. Seems to be more dropouts in the control group.
  - Overall, looks acceptably good.

Question 2 - Difference in Differences

- B - Replicating the experiment using the randomly assigned control group (from the original experiment).
  - In both cases: some positive effect of the treatment, although not significant at the 5% level.
  - Since it is randomly assigned, there should not be any OVB, so omitting the observables should not bias the results.
  - However, it will increase the R2 (or equivalently reduce the residual variance), hence making the estimator for the treatment effect (slightly) more precise.

Question 2 - Difference in Differences

- C,D - Difference-in-Differences
  - Let us pretend we do not have a randomly assigned control group (as would typically be the case using non-experimental data).
  - Then we need to find a clever identification strategy that does not rely on randomization.
  - Diff-in-diff is one such strategy.
    - We no longer need the individuals to look the same in the treatment and control group, nor their counterfactual y_0 and y_1 to have the same expected value.
    - It only requires the following assumption: if the treatment group had *not* been treated, the difference in outcomes would have remained the same over the period considered. I.e., the difference in the evolution ("difference-in-difference") is due solely to the effect of the treatment.
  - Stata: find a much smaller treatment effect. Why?

Question 2 - Difference in Differences

- E - Checking our diff-in-diff assumption by looking at the pretrend (74 to 75).
- If our identifying assumption is true (i.e., that in the absence of treatment both groups have the same trend in outcomes), then we should find nothing here.
- We do find a significant difference in the pretrends across groups, i.e. diff-in-diff is probably biased.

Question 3 - Propensity Score Matching

- Alternatives: need to rely on CIA, i.e. assignment is random conditional on a set of observables.
- We could match on all "suspicious" observables X_i and compare outcomes for individuals with "similar" X_i's.
- Alternatively, can rely on propensity score, i.e. $p(X_i) = E(D_i|X_i)$.
- Fondamental theorem:
  - Suppose the CIA holds, i.e $\{y_{0i}, y_{1i}\}$ is independent on D_i conditional on X_i. Then it is also the case that $\{y_{0i}, y_{1i}\}$ is independent on D_i conditional on $p(X_i)$.
  - (This answers section B of this question.)
- Idea of propensity score matching: simply match individuals on the basis of the $p(X_i)$ instead of on the full X_i's.

Question 3 - Propensity Score Matching

- A - Let $M = \{0, .1, .2, .., .9\}$. For $m \in M$ and individual i, denote $p_{m,i}$ a dummy that takes value 1 iff $m \leqq p(X_i) \leqq m + .1$.
- Then, for example, if $p(X_i) = .85$, then $p_{8,i} = 1$.
- Then we could run $y_i = \sum_{m \in M} \pi_m p_{m,i} + \sum_{m \in M} \theta_m (p_{m,i} D_i) + \epsilon_i$.
- In that case, the θ_m's would identify the treatment effect for individuals with the considered propensity score.

Question 3 - Propensity Score Matching

- C - The specification in A is quite flexible, as it allows for different treatment effects across propensity scores.

Question 4 - Propensity Score Matching with Data

- A, B, C, D - Stata