Lecture 7 – Matching

Intuition

The basic idea with matching is to find two similar people on the basis of covariates–one that was treated and one that wasn't. Then compare the outcomes of the treated person to the control person.

What is the assumption that is necessary for identification? Within matched observations, it is as good as random who is treated.

Benefits and Drawbacks

Common Support–often treated and control students have very different characteristics. Matching methods inherently restrict the characteristics so that they are similar across treatment and control.

Most people don't think that it's as good as random who is treated after matching. For instance, consider a set of identical twins. They have the same family characteristics, grew up in the same neighborhood, went to the same school, etc. Let's say one twin gets a bacehlor's degree and another does not. Is that likely random? Or is it the twin who is more interested in education/more motivated/etc. Is motivation okay to omit from regression of the effect of education on outcomes (like earnings)?

What does matching feel like given the assumptions necessary for a causal interpretation?

Types of Matching

Propensity Score–You estimate a logit for treatment and get a predicted probability of treatment. Next you match treated observations with non treated observations who have similar value of propensity score. You calculate the difference for matched observations and aggregate them up.

Nearest Neighbor Matching–You define a metric that allows you to find observations that are most similar to each treated observation. Then you take the difference between the treated observation and the nearest not treated observation. You aggregate these up and get an estimate of the treatment effect

Coarsened Exact Matching–Similar to nearest neighbor but you bin up the covariates and find exact matches on bins. For example, one bin is age 18-25, etc.

Deming (2009) has a nice discussion of this in Section 3 Part B (p. 118)