

Problem on Matching

Course Instructor : Dr. Rahul Makhijani

Consider the data for 8 fast food restaurants that were part of a study of the effect of raising the minimum wage in NJ. The **treatment** group is the 2 restaurants in NJ and the **control** group is a set of 6 restaurants in PA (where the minimum wage was not raised). The outcome Y_i^{obs} is the number of people employed (including part time employees) at the end of the year. There are two covariates – X_{i1} , the identify of the fast food chain (Burger King or Kentucky Fried Chicken) and X_{i2} , employment at the end of the year prior to the increase in the minimum wage.

Observation i	Treatment T_i	Rest.Chain X_{i1}	Init.Empl. X_{i2}	Final.Empl. Y_i^{obs}
1	NJ	BK	22.5	30.0
2	NJ	KFC	14.0	12.5
3	PA	KFC	13.8	17.0
4	PA	BK	26.5	18.5
5	PA	BK	20.0	19.5
6	PA	BK	13.5	21.0
7	PA	BK	32.5	26.5
8	PA	KFC	21.0	23.0

1. We want to use matching to estimate the effect of raising the minimum wage assuming that unconfoundedness holds. We will match a single control unit with each treatment unit (without replacement). Our distance measure is $D(i, j) = 100 * \mathbb{I}(X_{i1} \neq X_{j1}) + |X_{i2} - X_{j2}|$ where the indicator I is 1 if the two units are different chains and 0 if they are the same chain. Identify the matches for the 2 treatment units.
2. What is the estimate of the average treatment effect on the treated units (ATT)?
Average Treatment Effect of the Treated (ATT) is the average of the individual treatment effects of those treated (hence not the entire population).
3. An alternative estimand is the average treatment effect (ATE). Which estimand makes more sense here? Briefly justify your answer.

Source: STAT 265 course by Prof. Hal Stern