

Propensity Score Matching

Ani Katchova

© 2013 by Ani Katchova. All rights reserved.

Propensity Score Matching Overview

- Treatment evaluation examples and definitions
- Propensity score methodology
 - Treated and control groups
 - Probit/logit models to estimate propensity
 - Matching methods for treated and controlled observations
 - Treatment effects estimation
- Assumptions in propensity score matching
- Difference-in-differences models

Propensity Score Matching

Treatment evaluation definition

- Treatment evaluation is the estimation of the average effects of a program or treatment on the outcome of interest.
- Comparison of outcomes between treated and control observations.

Treatment evaluation examples

- Effects of training programs on job performance
- Government programs targeted to help schools and their effect on student performance

Two types of studies

- controlled experiments (assignment into treated and control groups is random)
- observational studies (assignment into treated and control groups is not random)

Propensity score matching methodology

- Assign the observations into two groups: the treated group that received the treatment and the control group that did not.
 - Treatment D is a binary variable that determines if the observation has the treatment or not
 - $D=1$ for treated observations and $D=0$ for control observations
- Estimate a probit/logit model for the propensity of observations to be assigned into the treated group. Use x variables that may affect the likelihood of being assigned into the treated group.
 - The propensity score model is a probit/logit model with D as the dependent variable and x as independent variables.
$$p(x) = \text{prob}(D = 1|x) = E(D|x)$$
 - The propensity score is the conditional (predicted) probability of receiving treatment given pre-treatment characteristics x .

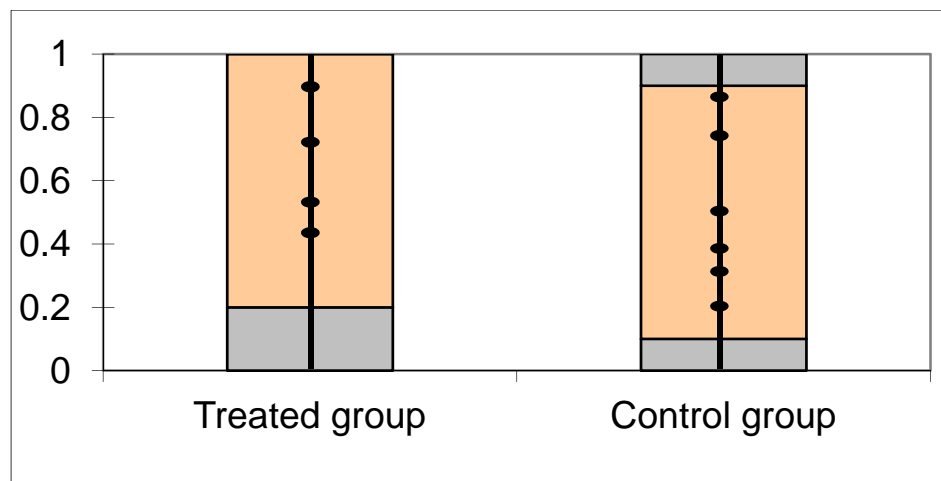
- Match observations from treated and control groups based on their propensity scores
 - Several matching methods are available: kernel, nearest neighbor, radius, stratification
- Calculate the treatment effects: compare the outcomes y between the treated and control observations, after matching

$$y = \begin{cases} y_1 & \text{if } D = 1 \\ y_0 & \text{if } D = 0 \end{cases}$$

- Counterfactual situation: compare the outcome of the treated observations with the outcome of the treated observations if they were not treated (find a close match using the control observations and use their outcome)

Matching methods explained

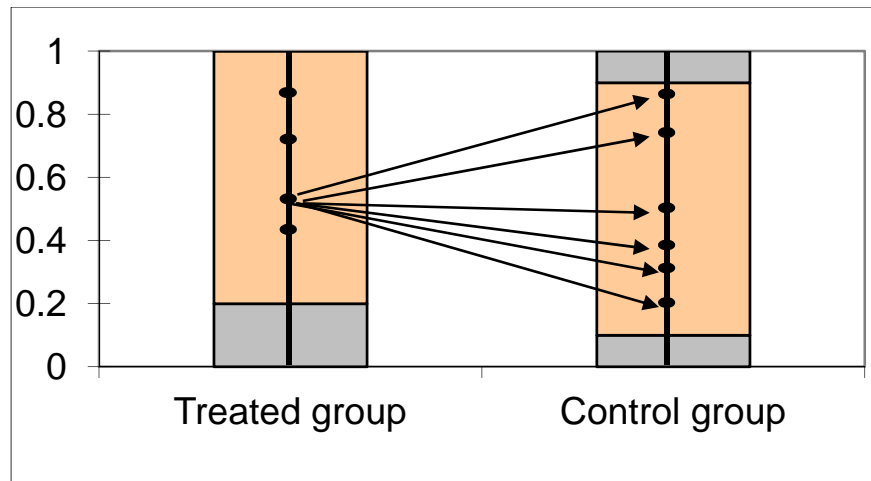
Propensity scores for treated and control groups



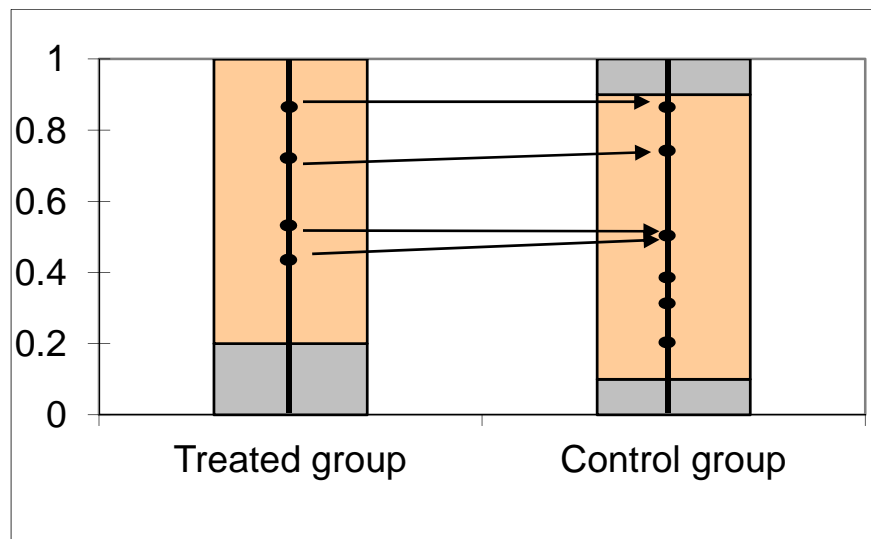
Matching methods: for each treated observation i , we need to find matches of control observation(s) j with similar characteristics.

- Matching with or without replacement
 - Matching without replacement - each control observation is used no more than one time as a match for a treated observation.
 - Matching with replacement – each control observation can be used as a match to several treated observations.

Kernel matching



Nearest neighbor matching



Nearest neighbor matching

- For each treated observation i , select a control observation j that has the closest x .

$$\min \| p_i - p_j \|$$

Radius matching

- Each treated observation i is matched with control observations j that fall within a specified radius.

$$\| p_i - p_j \| < r$$

Kernel matching

- Each treated observation i is matched with several control observations, with weights inversely proportional to the distance between treated and control observations.
- With matching based on propensity scores, the weights are defined as:

$$w(i, j) = \frac{K(\frac{p_j - p_i}{h})}{\sum_{j=1}^{n_0} K(\frac{p_j - p_i}{h})}$$

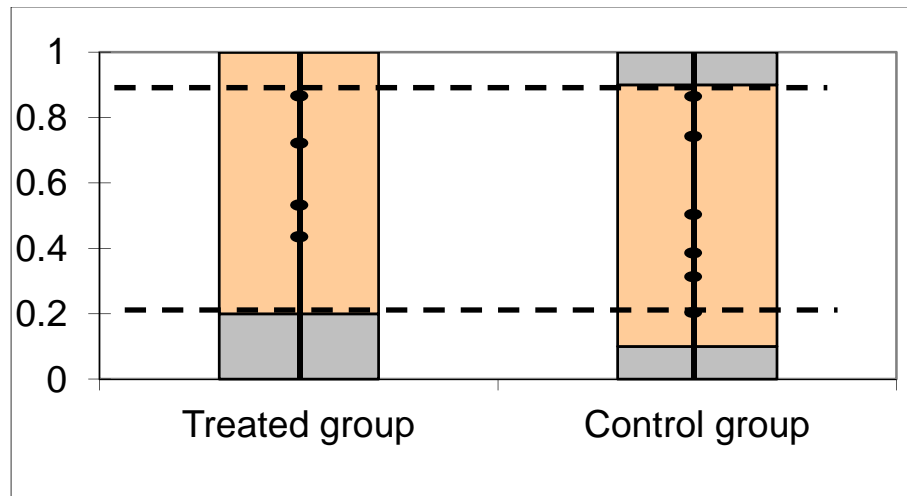
Here h is the bandwidth parameter.

Stratification or interval matching

- Compare the outcomes within intervals/blocks of propensity scores.

Matching with common support

- Restrict matching only based on the common range of propensity scores



Treatment effects

Average treatment effect (ATE)

- *ATE* is the difference between the outcomes of treated and control observations.

$$\Delta = y_1 - y_0$$

$$ATE = E(\Delta) = E(y_1|x, D = 1) - E(y_0|x, D = 0)$$

- A simple t-test between the outcomes for the treated and control groups.
- ATE is fine for random experiments but in observational studies, it may be biased if treated and control observations are not similar.

Average treatment effect on the treated (ATET)

- *ATET* is the difference between the outcomes of treated and the outcomes of the treated observations if they had not been treated.

$$ATET = E(\Delta|D = 1) = E(y_1|x, D = 1) - E(y_0|x, D = 1)$$

- The second term is a counterfactual so it is not observable and needs to be estimated.

Propensity score method

- After matching on propensity scores, we can compare the outcomes of treated and control observations.

$$ATET = E(\Delta|p(x), D = 1) = E(y_1|p(x), D = 1) - E(y_0|p(x), D = 0)$$

Empirical estimation

- Each treated observation i is matched j control observations and their outcomes y_0 are weighed by w .

$$ATET = \frac{1}{n_1} \sum_{i \in \{D=1\}} \left[y_{1,i} - \sum_j w(i,j) y_{0,j} \right]$$

Assumptions

- Partial equilibrium character (no general equilibrium effects)
 - Treatment does not indirectly affect the control observations.

Conditional independence assumption

- For random experiments, the outcomes are independent of treatment.

$$y_0, y_1 \perp D$$

- For observational studies, the outcomes are independent of treatment, conditional on x .

$$y_0, y_1 \perp D | x$$

- We need treatment assignment that ignores the outcomes.
- The treatment variable needs to be exogenous.

Unconfoundedness assumption

- Conditional independence of the control group outcome and treatment.
- Weaker assumption than the conditional independence assumption.

$$y_0 \perp D | x$$

Matching or overlap assumption

- For each value of x , there are both treated and control observations.
- For each treated observation, there is a matched control observation with similar x .

$$0 < \text{prob}(D = 1|x) < 1$$

Balancing condition

- Assignment to treatment is independent of the x characteristics, given the same propensity score.

$$D \perp x | p(x)$$

- The balancing condition is testable.

Difference-in-differences model

- The difference-in-differences model is applied when panel data on outcomes are available before (*b*) and after (*a*) the experiment occurs.
- The difference-in-differences model is an improvement over the one-period model.
- The difference-in-differences average treatment effect on the treated is specified as:

$$\begin{aligned} ATET &= E(\Delta_a - \Delta_b | D = 1) = E((y_{1a} - y_{0a}) - (y_{1b} - y_{0b}) | x, D = 1) = \\ &= E(y_{1a} - y_{1b} | x, D = 1) - E(y_{0a} - y_{0b} | x, D = 1) \end{aligned}$$

- The first term refers to the differences in outcomes before and after the treatment for the treated group. This term may be biased if there are time trends. The second term uses the differences in outcomes for the control group to eliminate this bias.
- To apply the difference-in-differences model: instead of the outcomes for the treated and control groups, we use the differences in outcomes after the treatment and before the treatment. The rest of the analysis is the same.