Propensity Score Analysis

Hsueh-Sheng Wu

CFDR Workshop Series

August 17, 2020

BGSU
Center for Family and Demographic Research

Outline

- What is propensity score analysis
- Four ways to define the treatment effect of T
- What does propensity score analysis do?
- Methods on controlling confounding variables
- Procedure of conducting propensity score analysis
- Selecting covariates for propensity score analysis
- Assessing the equivalence between treatment and untreated groups
- Inverse probability weighting
- Stata examples
- Conclusions

What Is Propensity Score Analysis

- Social scientists usually use observational data where random assignment of treatment is not possible and independent and outcome variables of interest are usually correlated to other confounding variables.
- When these confounding variables are not account for, the true relation between the independent variable and the outcome variable will be inflated or reduced.
- The propensity score for a subject is the estimated probability that the subject would be in the treatment group, P(T=1)
- The use of propensity score control creates equivalent (balanced) treatment and control groups in terms of confounding variables and help identify the unbiased relation between the independent and outcome variables.

Four Ways to Define the Treatment Effect of T

- The individual causal effect of T : $Y_i(1) - Y_i(0)$, which is impossible to estimate since each respondent has either the value of $Y_i(1)$ or $Y_i(0)$
- The average of possible outcomes on the treated = $E[Y_i(1)]$ and the average of possible outcomes on the untreated = $E[Y_i(0)]$
- The average treatment effect of T (ATE) = $E[Y_i(1) - Y_i(0)] = E[Y_i(1)] - E[Y_i(0)]$
- The average treatment effect of T on the treated (ATET) = $E[Y_i(1) - Y_i(0)]$ for people with a value of 1 of T
- If respondents 1-4 are very different from respondents 5-8 and these differences lead to their receiving job training and income in 2020, then the ATE and ATET will not provide an accurate estimate of the treatment effect of T

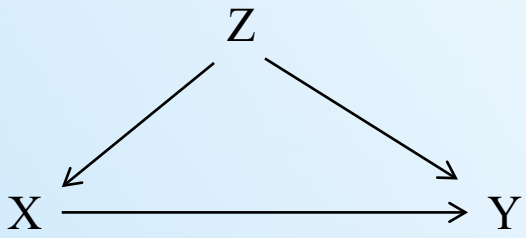
Table 1. A sample data

ID	Gender	income 2019	Job Training	Income 2020	Income 2020	Income 2020
i	X_{i1}	X_{i2}	T_i	Y_i	$Y_i(1)$	$Y_i(0)$
1	M	9690	1	10090	10090	?
2	F	7140	1	7540	7540	?
3	M	9735	1	10135	10135	?
4	F	8129	1	8529	8529	?
5	M	14500	0	205	?	14700
6	F	11995	0	206	?	12195
7	M	15906	0	207	?	16106
8	F	12990	0	208	?	13190

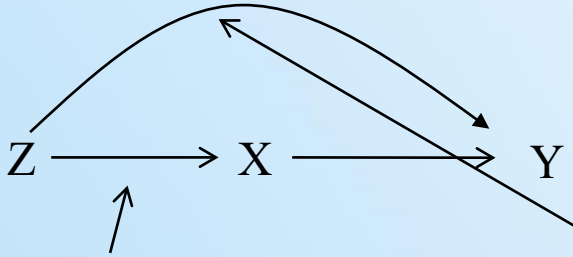
BGSU Center for Family and Demographic Research

What Does Propensity Score Analysis Do?

Graph 1. The X-Y relation is confounded by Z



Graph 2. The X-Y relation can be accurately estimated if researchers can control for the Z-X relation, Z-Y Relation, or both.



Propensity Score technique breaks the relation between Z and X through matching, stratification, and inverse probability weighting

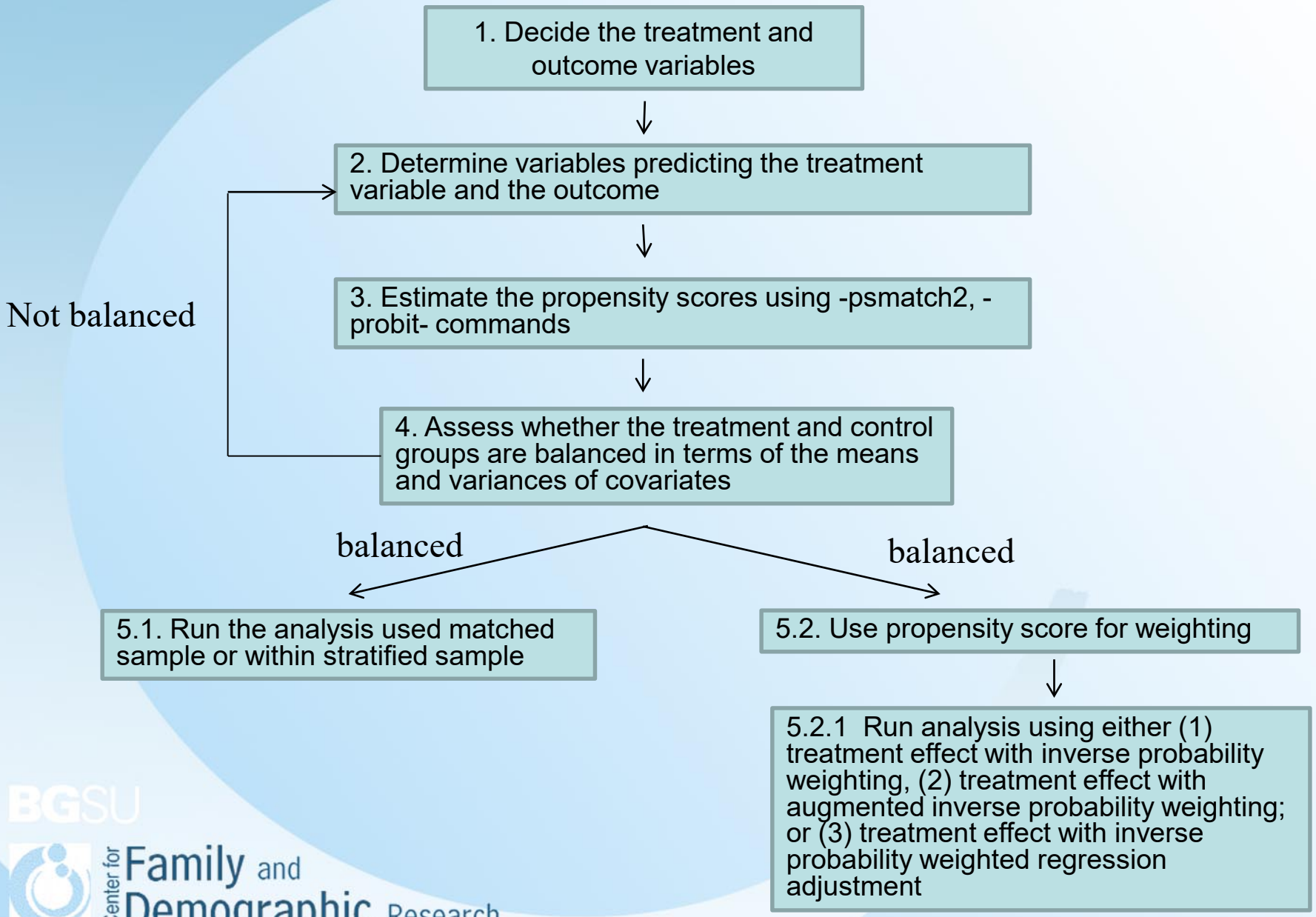
Regression adjustment breaks the relation between Z and Y by adding covariates of Y

BGSU Center for Family and Demographic Research

Methods on Controlling Confounding Variables

- Exact matching method
  - Match each respondent in the treated group with a respondent in the untreated group that have the same values of all variables
  - It is difficult to find an exact match for respondents when many covariates are involved.
- Stratification
  - Divide the sample into different homogenous strata and conduct analysis with each of them
  - The analysis results across stratum can be aggregated to obtain the average result of the whole sample
  - Hard to stratify on many covariates simultaneously
  - It will also reduce sample size
- Regression adjustment
  - Include possible confounding variables into the analysis
  - May not work if confounding variables have different functional forms for the treated group than for the untreated group

Procedure of Conducting Propensity Score Analysis



Selecting Covariates for Propensity Score Analysis

- When calculating propensity score, be sure to include covariates predicting both the independent variable and the outcome variable
- The more covariates you used, the more likely you control for confounding covariates. However, it also means that it is harder to find matches for some respondents.
- Elizabeth Stuart (2011) suggests:
  - With large samples, it is better to include more covariates than less
  - With small samples, it is better to include mainly variables related to the outcome variables
  - Do not include covariates that mediate the relationship between the independent variable and the outcome variable
  - Do not include covariates that perfectly predict the treatment assignment

How to Estimate Propensity Score

- Propensity score can be calculated using -psmatch2- and -probit- command
- -psmatch2- allows researchers to quickly identify which respondents are matched with which respondents

Examples:

1. allows one person to be matched to multiple treated respondents

psmatch2 mbsmoke mmarried c.mage##c.mage fbaby medu, out(bweight)

2. allows a people to be matched to only one treated respondents

psmatch2 mbsmoke mmarried c.mage##c.mage fbaby medu, out(bweight) noreplacement

- -probit- command

probit mbsmoke mmarried c.mage##c.mage fbaby medu

predict ps_probit

Different Types of Propensity Matching

- One-to-one matching: pick one match for each treated unit
- k-nearest neighbors matching: pick k matches for each treated unit
- Radius matching: researchers choose a threshold and all controls with a distance smaller than threshold are used to match the treated unit
- Kernel matching: like radius matching, but controls with smaller distances are given larger weights
- Mahalanobis matching: The Mahalanobis distance (MD) is the distance between two points in multivariate space. The centroid of this multivariate space can be thought of as an overall mean for multivariate data. Mahalanobis distance measures the distance between an individual and the centroid. The larger the MD, the further away from the centroid the data point is.

Assessing the Equivalence between Treatment and Untreated Groups

Rubin (2001) proposed three conditions:

- The difference in the means of the propensity scores in the two groups being compared must be smaller than 0.5 standard deviation
- The ratio of the variances of the propensity score in the two groups must be close to one
- The ratio of the variances of the residuals of the covariates after adjusting for the propensity score must be close to one

Inverse Probability Weighting

- Propensity score: p(x)=P(T=1|X=x)
- Inverse probability weight are w(x)=1/p(x) for treated individuals and w(x)=1/(1-p(x)) for untreated respondents
- The higher the propensity score a respondent has, the smaller weights the respondent gets.
- Stata –teffects- command has three inverse probability weighting estimation options:
  - *Treatment effect with inverse-probability weighting* uses weighted means rather than simple unweighted means to control the effects of confounders on the treatment
  - *Treatment effect with augmented inverse-probability weighting* includes an augmentation term to correct the estimator when the treatment model is mis-specified.
  - *Treatment effects with inverse-probability-weighted regression adjustment* uses inverse-probability weights to correct the estimator when the regression model is mis-specified.

Conclusions

- Propensity score analysis is a usual technique for controlling for confounding variables when observational data are used.
- The propensity score is estimated, rather than actually observed, thus the mis-specification of propensity score model can lead to inaccurate estimate of propensity score.
- Create balanced treated and untreated groups is a process of specifying the model, assess the equivalence between groups, and re-specifying the models.
- Researchers probably need to try several different matching methods before finding equivalent treatment and control groups.
- Propensity score can be used to calculate the inverse probability weights for respondents. In addition, the treatment effect with augmented inverse probability weighting works even when the treatment model is mis-specified. In addition, the treatment effects with inverse probability weighted regression adjustment works even when the regression adjustment model is mis-specified.
- Both -teffects- and -psmatch2- can be used for propensity score analysis, but with different features. Researchers need to choose one that better fits their research needs.
- If you have any questions, please contact me at wuh@bgsu.edu.