

Effects of propensity score overlap on the estimates of treatment effects

Yating Zheng & Laura Stapleton

University of Maryland, College Park

Background

- An important problem of casual inference is how to estimate treatment effects in observational studies.
- In recent years, propensity score (PS) methods have been widely used because they can:
 1. easily construct matched sets on multiple covariates (Dehejia & Wahba, 2002).
 2. avoid the model extrapolation problems in regression (Arpino & Cannas, 2015).
- A potential problem: PS methods require a sufficient overlap of the propensity score distributions between the treatment and control groups.
- Previous studies seldom explored how the lack of overlap would influence the estimates of treatment effects.

Research Question

- How does propensity score overlap influence the estimates of treatment effects, in terms of bias and variance?

Theoretical Framework

- Definition: propensity score is the probability of receiving the treatment given the observed covariates (Rosenbaum & Rubin, 1983).
- PS methods include: matching, weighting and sub-classification.
- Key idea: make the treatment and control group as comparable as possible in the baseline covariates to facilitate an unbiased estimate of the treatment effect.
- Steps of propensity score weighting (PSW):
 1. identify baseline covariates that could potentially bias estimates of the treatment effect
 2. calculate propensity scores using logistic regression (or a non-parametric approach) on the baseline covariates
 3. change the weights of subjects based on estimated propensity scores weighting method: weighting by the odds (WBO)

$$w_i = T_i + (1 - T_i) \frac{e_i}{1 - e_i}$$
 where w_i is the weight for subject i , T_i is treatment indicator, e_i is the estimated propensity score
 4. balance check
 5. estimate the treatment effects (Stuart, 2010).

- Overlap: the intersection of the propensity score distributions over the union of the distributions.

Design

- Simulation study
- Factors manipulated:
 - overlap rate: 0.1, 0.4, 0.7
 - sample size (treat group): 100, 500, 1000
 - sample size ratio (treat: control): 1:1, 1:3, 1:9
- 27 conditions, each replicated 1000 times
- Data generation:
 1. search for covariate parameters to achieve specified overlap rates
 2. generate covariates for treatment and control groups separately from normal and binomial distributions
 3. set the true treatment effect and calculate outcome values using a regression

Analysis

fit a logistic regression model with covariates as predictors and treatment assignment as outcome

use the fitted model to obtain estimated propensity scores

calculate overlap rate using estimated propensity scores

check the consistency between estimated overlap rates and specified overlap rates

weight the subjects using WBO

estimate treatment effects and compare with true treatment effect

Results

- As overlap rate and sample size increases, the average bias decreases. The trend is consistent across different ratios (Figure 1).
- At low overlap rates with small sample sizes, the estimates are really inaccurate. E.g., at the overlap rate of 0.1 with treatment group size of 100 and control group size of 100, the average bias is about 84% of the true treatment effect (Table 1).
- Variance of the treatment effect estimates decreases with the increase of overlap rate and sample size (Table 2).

Sample size ratio (Treat: Control)	Specified overlap rate	Sample size of the treatment group		
		100	500	1000
1:1	0.1	8.40E-01	4.11E-01	3.59E-01
	0.4	3.61E-02	7.27E-03	-3.03E-03
	0.7	3.05E-03	6.41E-04	1.11E-04
1:3	0.1	5.93E-01	3.48E-01	2.04E-01
	0.4	4.91E-03	2.99E-03	-2.10E-03
	0.7	1.55E-03	-7.60E-04	4.21E-04
1:9	0.1	2.00E-01	6.63E-02	3.17E-02
	0.4	-2.22E-03	2.83E-04	-1.02E-04
	0.7	1.51E-03	-3.23E-04	-4.03E-05

Sample size ratio (Treat: Control)	Specified overlap rate	Sample size of the treatment group		
		100	500	1000
1:1	0.1	4.53E-02	3.03E-02	2.30E-02
	0.4	3.56E-03	6.05E-04	3.40E-04
	0.7	5.23E-04	7.40E-05	3.56E-05
1:3	0.1	3.68E-02	1.77E-02	1.70E-02
	0.4	7.77E-04	1.35E-04	7.09E-05
	0.7	8.90E-05	1.23E-05	6.37E-06
1:9	0.1	1.83E-02	9.74E-03	5.80E-03
	0.4	8.13E-05	1.44E-05	7.31E-06
	0.7	1.28E-05	1.92E-06	8.98E-07

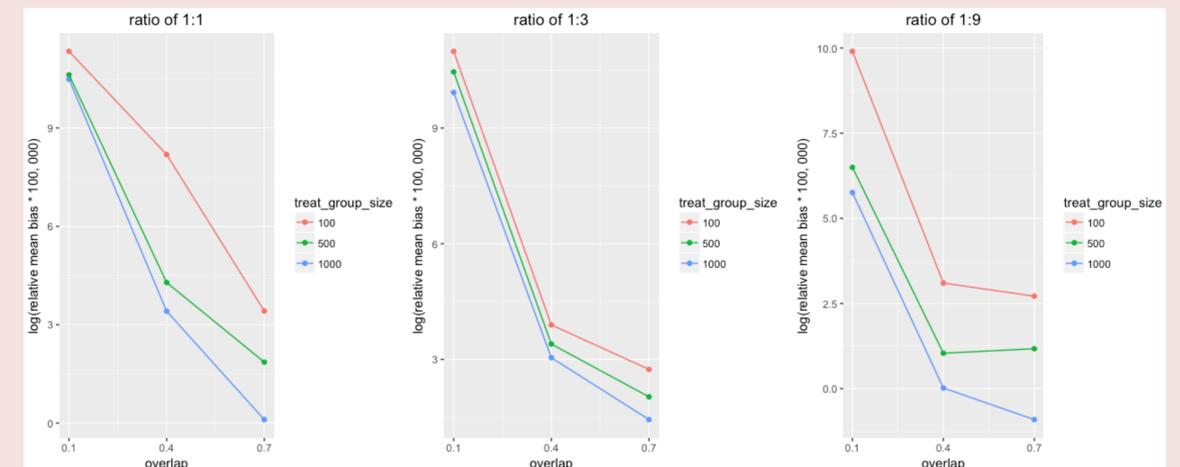


Figure 1. Relationships between bias, sample size and overlap rate.

Discussions

Findings:

- At small overlap rates with small sample sizes, the estimates of PSW are inaccurate and unstable. Researchers should be cautious about using PSW in these conditions.
- As sample size and overlap rate increase, the estimates become more accurate and reliable.

Limitations:

- The research design did not separate the effects of ratio and sample size, making it impossible to examine the effects of sample size ratio on the estimates of treatment effects at different overlap rates.

Future directions:

- Explore the effects of propensity score overlap on the treatment effect estimates using other PS methods.