

Shu Cong
s8cong@uwaterloo.ca

Dept. Mathematics
University of Waterloo



Does National Supported Work(NSW) Job Training Program Works?

Joint work with mentor Xiaoya Wang

July, 2024

Outline

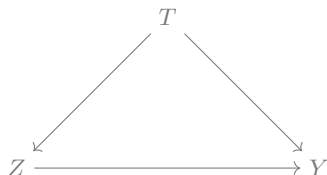
- 1 Introduction
 - Background
 - Project purpose
- 2 Model settings
 - Notations
 - Model settings
- 3 Methodology
 - Propensity score method
 - Propensity score matching
- 4 Data Application
 - Dataset
 - Love plot
 - Table
 - Analyze
- 5 Conclusion

Background

- ❖ The National Supported Work Demonstration (NSW) job-training program was designed to help disadvantaged workers lacking basic job skills move into the labor market by giving them work experience and counseling in a sheltered environment in the mid-1970s.
- ❖ Lalonde (1986) is interested in evaluating the effects of the NSW program using econometric methods [Lalonde, 1986].

Project purpose

- ❖ Research question:
 - ❖ Whether or not joining the in the NSW Job Training Program helps with real earnings in 1978 with addressing confounding issue
- ❖ In this project, we aim to do causal analysis and evaluate the causal effect of NSW program.



Notations

- ❖ T -treatment: whether or not a person joined NSW Job Training Program
- ❖ Y -outcome: earnings in 1978
- ❖ Z -confounders: age, education, race, etc.

Model settings

- Response model:

$$Y = \alpha_0 + \alpha_1 T + \alpha_2 Z_1 + \dots + \alpha_9 Z_8 + E. \quad (1)$$

$$E \sim N(0, \sigma_Y^2)$$

- Propensity score model:

$$\log \left(\frac{\pi}{1 - \pi} \right) = \beta_0 + \beta_1 Z_1 + \beta_2 Z_2 + \dots + \beta_8 Z_8 \quad (2)$$

with $\pi = P(T = 1 \mid Z_1, Z_2, Z_3, Z_4, Z_5, Z_6, Z_7, Z_8)$

Propensity score method [Rosenbaum and Rubin, 1983]

❖ Definition

The propensity score $\pi(Z)$ is defined as the conditional probability of receiving the treatment given the observed covariates:

$$\pi(Z) = P(T = 1 | Z)$$

where:

- ❖ T is a binary indicator of treatment assignment (1 if the unit receives the treatment, 0 otherwise).
- ❖ Z represents the observed covariates.

Propensity score matching

❖ Estimate Propensity Scores

- ❖ logistic regression:

$$\log\left(\frac{\hat{\pi}}{1-\hat{\pi}}\right) = \hat{\beta}_0 + \hat{\beta}_1 Z_1 + \hat{\beta}_2 Z_2 + \dots + \hat{\beta}_8 Z_8$$

❖ Matching

- ❖ 1:1 nearest neighbor (NN) matching [Rosenbaum and Rubin, 1983]:

- ▶ One by one, each treated unit is paired with an available control unit that has the closest propensity score to it.
- ▶ Any remaining control units are left unmatched and excluded from further analysis.

- ❖ Full matching [Hansen, 2004, Stuart and Green, 2008]:

- ▶ Match every treated unit to at least one control
- ▶ Match every control to at least one treated unit

Propensity score matching

❖ Check Balance



$$\text{SMD} = \frac{\bar{X}_T - \bar{X}_C}{\sqrt{\frac{S_T^2 + S_C^2}{2}}}$$

where S_T^2, S_C^2 are the sample variance for the treated and control group.

- ❖ Use **summary(match, un = FALSE)** in R to assess covariate balance post-matching, ensuring mean differences are near zero and standardized mean differences (SMD) are less than 0.1 for good balance.

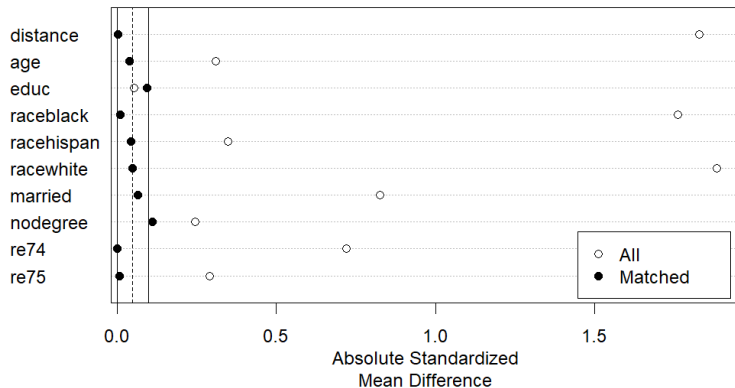
❖ Estimate Treatment Effect

- ❖ Average Treatment Effect (ATE):

- ▶ We can run a regression of the outcome on the treatment and covariates in the matched sample (i.e., including the matching weights)
- ▶ We estimate the treatment effect using g-computation as implemented in `marginaleffects::comparisons()`

Dataset: nsw_mixture

- ❖ Data from the National Supported Work Demonstration (NSW) job training program, where those treated were guaranteed a job for 9-18 months.
- ❖ A data frame with 445 rows and 11 variables.
- ❖ Confounders (Every binary variables)
 - ❖ **age** Age in years
 - ❖ **educ** Years of education
 - ❖ **black** Race: Black
 - ❖ **hisp** Ethnicity: Hispanic
 - ❖ **marr** Married
 - ❖ **nodegree** Has no degree
 - ❖ **re74** Real earnings in 1974
 - ❖ **re75** Real earnings in 1975
- ❖ Treatment: treat In the National Supported Work Demonstration Job Training Program
- ❖ Outcome: Real earnings in 1978



Term	Contrast	Estimate	Std. Err	z	$\Pr(> z)$	S	2.5%	97.5%
treat	$\mu_1 - \mu_0$	1977	704	2.81	0.00501	7.6	596	3357

Table: Estimate Table

Analyze

- ❖ **Propensity Score Matching:** We used full matching on the propensity score estimated via probit regression to achieve adequate balance, with all standardized mean differences for covariates below 0.1.
- ❖ **Balance Achievement:** Full matching utilized all treated and control units, ensuring no units were discarded, and achieving standardized mean differences for squares and two-way interactions below 0.15.
- ❖ **Treatment Effect:** The estimated average treatment effect on 1978 earnings was \$1977 (SE = 704, $p = 0.00501$), indicating a significant positive impact of the treatment on earnings.

Conclusion

- ❖ Method: Propensity Score Matching (PSM)
- ❖ The analysis suggests that while the job training program might have had a positive effect on earnings, the evidence is strong enough to be statistically significant at the 0.05 level. The confidence interval also confirms the job training program helps on earnings. Further research or additional data might be needed to draw more definitive conclusions.
- ❖ Limitations & Future Work:
 - ❖ Reliance on observed covariates; potential unobserved confounders
 - ❖ Advanced causal methods (e.g., instrumental variables, difference-in-differences)
 - ❖ Long-term impact analysis

References

- Ben B Hansen. Full matching in an observational study of coaching for the sat. *Journal of the American Statistical Association*, 99(467):609–618, 2004.
- Robert J. Lalonde. Evaluating the econometric evaluations of training programs. *American Economic Review*, 76:604–620, 1986.
- Paul R Rosenbaum and Donald B Rubin. The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1):41–55, 1983.
- Elizabeth A Stuart and Kerry M Green. Using full matching to estimate causal effects in nonexperimental studies: examining the relationship between adolescent marijuana use and adult outcomes. *Developmental psychology*, 44(2):395, 2008.