

Turley, Falynn C.; Redden, David; Case, Janice L.; Katholi, Charles; Szychowski, Jeff; Dubay, Derek

Comparison of type I error rates and statistical power of different propensity score methods.

(English) [Zbl 07192578](#)

J. Stat. Comput. Simulation 88, No. 4, 769-784 (2018)

Summary: Propensity score analysis (PSA) is a technique to correct for potential confounding in observational studies. Covariate adjustment, matching, stratification, and inverse weighting are the four most commonly used methods involving propensity scores. The main goal of this research is to determine which PSA method performs the best in terms of protecting against spurious association detection, as measured by Type I error rate, while maintaining sufficient power to detect a true association, if one exists. An examination of these PSA methods along with ordinary least squares regression was conducted under two cases: correct PSA model specification and incorrect PSA model specification. PSA covariate adjustment and PSA matching maintain the nominal Type I error rate, when the PSA model is correctly specified, but only PSA covariate adjustment achieves adequate power levels. Other methods produced conservative Type I Errors in some scenarios, while liberal Type I error rates were observed in other scenarios.

MSC:

62-XX Statistics

Keywords:












propensity score; confounding; type I error; logistic regression; power










Software:

Binnor

Full Text: [DOI](#)

References:

- [1] Rosenbaum PR, Rubin DB. The central role of the propensity score in observational studies for causal effects. *Biometrika*. 1983;70:41-55. doi: 10.1093/biomet/70.1.41[Crossref], [Web of Science , [Google Scholar] · [Zbl 0522.62091](#)
- [2] Cochran WG. The effectiveness of adjustment by subclassification in removing bias in observational studies. *Biometrics*. 1968;24:295-313. Epub 1968/06/01. doi: 10.2307/2528036[Crossref], [PubMed], [Web of Science , [Google Scholar]
- [3] Rosenbaum PR. Model-based direct adjustment. *J Amer Statist Assoc*. 1987;82:387-394. doi: 10.1080/01621459.1987.10478441[Taylor & Francis Online], [Web of Science , [Google Scholar] · [Zbl 0622.62010](#)
- [4] Rubin DB. Estimating causal effects of treatments in randomized and nonrandomized studies. *J Educ Psychol*. 1974;66:688-701. doi: 10.1037/h0037350[Crossref], [Web of Science , [Google Scholar]
- [5] Hade EM, Lu B. Bias associated with using the estimated propensity score as a regression covariate. *Stat Med*. 2014;33:74-87. Epub 2013/06/22. doi: 10.1002/sim.5884[Crossref], [PubMed], [Web of Science , [Google Scholar]
- [6] Austin PC. An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate Behav Res*. 2011;46:399-424. Epub 2011/08/06. doi: 10.1080/00273171.2011.568786[Taylor & Francis Online], [Web of Science , [Google Scholar]
- [7] Austin PC. Type I error rates, coverage of confidence intervals, and variance estimation in propensity-score matched analyses. *Int J Biostat*. 2009;5:Article 13. doi: 10.2202/1557-4679.1146[Crossref], [PubMed], [Google Scholar]
- [8] Austin PC. A comparison of 12 algorithms for matching on the propensity score. *Stat Med*. 2014;33:1057-1069. Epub 2013/10/15. doi: 10.1002/sim.6004[Crossref], [PubMed], [Web of Science , [Google Scholar]
- [9] Austin PC. The performance of different propensity score methods for estimating marginal hazard ratios. *Stat Med*. 2013;32:2837-2849. Epub 2012/12/15. doi: 10.1002/sim.5705[Crossref], [PubMed], [Web of Science , [Google Scholar]
- [10] Schafer JL, Kang J. Average causal effects from nonrandomized studies: a practical guide and simulated example. *Psychol Methods*. 2008;13:279-313. doi: 10.1037/a0014268[Crossref], [PubMed], [Web of Science , [Google Scholar]
- [11] Sjolander A, Greenland S. Ignoring the matching variables in cohort studies – when is it valid and why?. *Stat Med*. 2013;32:4696-4708. Epub 2013/06/14. doi: 10.1002/sim.5879[Crossref], [PubMed], [Web of Science , [Google Scholar]
- [12] Rubin DB, Thomas N. Combining propensity score matching with additional adjustments for prognostic covariates. *J Amer Statist Assoc*. 2000;95:573-585. doi: 10.1080/01621459.2000.10474233[Taylor & Francis Online], [Web of Science , [Google Scholar]

- [13] Austin PC, Mamdani MM, Stukel TA, et al. The use of the propensity score for estimating treatment effects: administrative versus clinical data. *Stat Med.* 2005;24:1563-1578. doi: 10.1002/sim.2053[Crossref], [PubMed], [Web of Science , [Google Scholar]
- [14] Lunceford JK, Davidian M. Stratification and weighting via the propensity score in estimation of causal treatment effects: a comparative study. *Stat Med.* 2004;23:2937-2960. Epub 2004/09/08. doi: 10.1002/sim.1903[Crossref], [PubMed], [Web of Science , [Google Scholar]
- [15] Robins JM, Hernan MA, Brumback B. Marginal structural models and causal inference in epidemiology. *Epidemiology.* 2000;11:550-560. doi: 10.1097/00001648-200009000-00011[Crossref], [PubMed], [Web of Science , [Google Scholar]
- [16] Rosenbaum PR, Rubin DB. Reducing bias in observational studies using subclassification on the propensity score. *J Amer Statist Assoc.* 1984;79:516-524. doi: 10.1080/01621459.1984.10478078[Taylor & Francis Online], [Web of Science , [Google Scholar]
- [17] Rubin DB, Thomas N. Matching using estimated propensity scores: relating theory to practice. *Biometrics.* 1996;52:249-264. doi: 10.2307/2533160[Crossref], [PubMed], [Web of Science , [Google Scholar] · [Zbl 0881.62121](#)
- [18] Arah O. Variable selection and covariate balance in propensity score methods for comparative effectiveness research. [PowerPoint Presentation]. In press 2009. [Google Scholar]
- [19] International SUG, Institute SAS. SUGI 29 Proceedings: SAS Users Group International Conference; 2004 May 9-12; Palais Des Congrès de Montréal, Montréal, Canada; Also Includes the Papers from the SUGI 27 and SUGI 28 Proceedings: SAS Institute; 2004. [Google Scholar]
- [20] Austin PC. Variance estimation when using inverse probability of treatment weighting (IPTW) with survival analysis. *Stat Med.* 2016;35:5642-5655. doi: 10.1002/sim.7084[Crossref], [PubMed], [Web of Science , [Google Scholar]
- [21] Wolter K. Introduction to variance estimation. 1st ed. New York: Springer-Verlag; 1985. [Google Scholar] · [Zbl 0581.62009](#)
- [22] Rubin DB. Estimating causal effects from large data sets using propensity scores. *Ann Intern Med.* 1997;127:757-763. Epub 1998/02/12. doi: 10.7326/0003-4819-127-8_Part_2-199710151-00064[Crossref], [PubMed], [Web of Science , [Google Scholar]
- [23] Williamson E, Morley R, Lucas A, et al. Propensity scores: from naive enthusiasm to intuitive understanding. *Stat Methods Med Res.* 2012;21:273-293. Epub 2011/01/26. doi: 10.1177/0962280210394483[Crossref], [PubMed], [Web of Science , [Google Scholar] · [Zbl 1242.62124](#)
- [24] Newey W. Treatment effects [Course Materials for New Econometric Methods]. In press 2007. [Google Scholar]
- [25] Imbens G, National Bureau of Economic Research. Nonparametric estimation of average treatment effects under exogeneity: a review. Cambridge (MA): National Bureau of Economic Research; 2003. [Crossref], [Google Scholar]
- [26] Demirtas H, Amatya A, Doganay B. Binnor: an R package for concurrent generation of binary and normal data. *Comm Statist Simulation Comput.* 2014;43:569-579. doi: 10.1080/03610918.2012.707725[Taylor & Francis Online], [Web of Science , [Google Scholar] · [Zbl 1291.62077](#)

This reference list is based on information provided by the publisher or from digital mathematics libraries. Its items are heuristically matched to zbMATH identifiers and may contain data conversion errors. It attempts to reflect the references listed in the original paper as accurately as possible without claiming the completeness or perfect precision of the matching.