Flexible sensitivity analysis for causal inference in observational studies subject to unmeasured confounding

Sizhu Lu

Joint work with Professor Peng Ding

Department of Statistics, UC Berkeley

April 17, 2024

Ding group meeting

Overview

Background: causal inference in observational studies

Sensitivity analysis with unmeasured confounding

Extensions

Potential outcome framework

- Potential outcome framework.
- Potential outcomes: $Y_i(1)$ and $Y_i(0)$.
- Binary treatment: Z_i.
- Observed outcome: $Y_i = Y_i(Z_i) = Z_i Y_i(1) + (1 - Z_i) Y_i(0)$.
- Stable unit treatment values assumption.
- Super population regime: independently and identically distributed $\{X_i, Z_i, Y_i(1), Y_i(0) : i = 1, \dots, n\}$.

Parameter of interest

- Causal parameter of interest: $\tau = E\{Y(1) - Y(0)\}$, the average treatment effect, decomposes into

$$\begin{aligned} \tau &= [E(Y \mid Z=1)\mathrm{pr}(Z=1) + E\{Y(1) \mid Z=0\}\mathrm{pr}(Z=0)] \\ &\quad - [E\{Y(0) \mid Z=1\}\mathrm{pr}(Z=1) + E(Y \mid Z=0)\mathrm{pr}(Z=0)]\,. \end{aligned}$$

- Fundamental challenge of causal inference: to estimate the counterfactual means $E\{Y(1) \mid Z=0\}$ and $E\{Y(0) \mid Z=1\}$.
- Randomization leads to obvious identification, but in observational studies?

Identification under unconfoundedness

- Unconfoundedness assumption: $Z \perp\!\!\!\perp \{Y(1), Y(0)\} \mid X$ (Rosenbaum and Rubin, 1983).
- Under this assumption, $E\{Y(1) \mid Z=1, X\} = E\{Y(1) \mid Z=0, X\}$, thus τ is nonparametrically identified.
- Two identification formulas:

$$\begin{aligned}\tau &= E\{\mu_1(X) - \mu_0(X)\} \\ &= E\left\{\frac{ZY}{e(X)} - \frac{(1-Z)Y}{1-e(X)}\right\},\end{aligned}$$

  where
  - $\mu_1(X) = E(Y \mid Z=1, X)$ and $\mu_0(X) = E(Y \mid Z=0, X)$: conditional expectation of outcomes;
  - $e(X) = \mathrm{pr}(Z=1 \mid X)$: propensity score.
- Implicitly assume overlap: $0 < e(X) < 1$.

Estimation under unconfoundedness

- Estimators corresponding to the two identification formulas:

$$\hat{\tau}^{\text{reg}} = n^{-1} \sum_{i=1}^{n} \{\hat{\mu}_1(X_i) - \hat{\mu}_0(X_i)\},$$

$$\hat{\tau}^{\text{ht}} = n^{-1} \sum_{i=1}^{n} \left\{ \frac{Z_i Y_i}{\hat{e}(X_i)} - \frac{(1 - Z_i) Y_i}{1 - \hat{e}(X_i)} \right\},$$

where $\hat{e}(X_i)$ and $\hat{\mu}_z(X_i)$ are fitted propensity score and outcome models.

Estimation under unconfoundedness

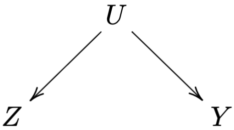
- Doubly robust estimator by combining both models (Bang and Robins, 2005):

$$\hat{\tau}^{\mathrm{dr}} = \hat{\tau}^{\mathrm{reg}} + n^{-1} \sum_{i=1}^{n} \left[\frac{Z_i \{Y_i - \hat{\mu}_1(X_i)\}}{\hat{e}(X_i)} - \frac{(1 - Z_i)\{Y_i - \hat{\mu}_0(X_i)\}}{1 - \hat{e}(X_i)} \right].$$

- Modifies $\hat{\tau}^{\mathrm{reg}}$ by inverse propensity score weighted residuals.
- Consistent to τ if either outcome models or propensity score model is correctly specified.

Sensitivity analysis

- Unconfoundedness assumption: untestable, cannot use data to validate.
- Existence of unmeasured confounding possibly overturns an observed association between the treatment and outcome.
- Hidden confounder U:



Source: Ding (2024).

- Sensitivity analysis: assess the impact of U; how strong the unmeasured confounding needs to be to overturn the observed association.

Sensitivity analysis

- Parametric models to assess the impact of U on the estimation of τ (Rosenbaum and Rubin, 1983; Lin et al., 1998; Imbens, 2003).
- Sensitivity analysis to test the sharp null hypothesis of no unit-level causal effects in matched-pair observational studies (Rosenbaum, 1987).
- E-value: sensitivity analysis for causal estimates based on risk ratios (Cornfield et al., 1959; Ding and VanderWeele, 2016; VanderWeele and Ding, 2017).
- Sensitivity analysis methods for the inverse propensity score weighting estimator (Zhao et al., 2019; Dorn and Guo, 2022).
- Useful for specific estimation or testing strategies.
- Deal with the standard estimators $\hat{\tau}^{\text{reg}}$, $\hat{\tau}^{\text{ht}}$ and $\hat{\tau}^{\text{dr}}$ simultaneously?

Identification challenge revisit: two extreme solutions

$$\begin{aligned}\tau &= [E(Y \mid Z=1)\mathrm{pr}(Z=1) + E\{Y(1) \mid Z=0\}\mathrm{pr}(Z=0)] \\ &\quad - [E\{Y(0) \mid Z=1\}\mathrm{pr}(Z=1) + E(Y \mid Z=0)\mathrm{pr}(Z=0)]\,.\end{aligned}$$

- Unconfoundedness assumption: $Z \perp\!\!\!\perp \{Y(1), Y(0)\} \mid X$.

$$\begin{aligned}E\{Y(1) \mid Z=1, X\} &= E\{Y(1) \mid Z=0, X\}, \\ E\{Y(0) \mid Z=1, X\} &= E\{Y(0) \mid Z=0, X\}.\end{aligned}$$

- Very restrictive—assumes the two treatment groups have **identical** conditional means.

Identification challenge revisit: two extreme solutions

$$
\begin{aligned}
\tau = & \; [E(Y \mid Z=1)\text{pr}(Z=1) + E\{Y(1) \mid Z=0\}\text{pr}(Z=0)] \\
& - [E\{Y(0) \mid Z=1\}\text{pr}(Z=1) + E(Y \mid Z=0)\text{pr}(Z=0)] \,.
\end{aligned}
$$

- Partial identification method:
  - Assume the potential outcomes are bounded between $[\ell, u]$.
  - $E\{Y(1)\}$ has lower bound

    $$E(Y \mid Z=1)\text{pr}(Z=1) + \ell\text{pr}(Z=0)$$

    and upper bound

    $$E(Y \mid Z=1)\text{pr}(Z=1) + u\text{pr}(Z=0).$$

  - Similar lower and upper bounds for $E\{Y(0)\}$, leading to bounds for τ.
  - Problem: bounds are always too wide to provide valuable causal information.
- Is there a midground?

Table of Contents

Sensitivity analysis with unmeasured confounding

- Define sensitivity parameters:

$$\frac{E\{Y(1) \mid Z=1,X\}}{E\{Y(1) \mid Z=0,X\}} = \varepsilon_1(X), \quad \frac{E\{Y(0) \mid Z=1,X\}}{E\{Y(0) \mid Z=0,X\}} = \varepsilon_0(X).$$

- $\varepsilon_1(X)$ and $\varepsilon_0(X)$: two sensitivity parameters.
- Quantifies the violation of the unconfoundedness assumption.
- $\varepsilon_1(X) = \varepsilon_0(X) = 1$: the unconfoundedness assumption.
- First fix them to obtain the corresponding estimators and then vary them within a range to obtain a sequence of estimators.

Identification and estimation under sensitivity analysis

Outcome regression

- With known $\varepsilon_1(X)$ and $\varepsilon_0(X)$,

$$
\begin{aligned}
E\{Y(1) \mid Z=0\} &= E\{\mu_1(X)/\varepsilon_1(X) \mid Z=0\}, \\
E\{Y(0) \mid Z=1\} &= E\{\mu_0(X)\varepsilon_0(X) \mid Z=1\}.
\end{aligned}
$$

- Estimator:

$$
\begin{aligned}
\hat{\tau}^{\text{proj}} &= n^{-1}\sum_{i=1}^{n}\{Z_i\hat{\mu}_1(X_i) + (1-Z_i)\hat{\mu}_1(X_i)/\varepsilon_1(X_i)\} \\
&\quad - n^{-1}\sum_{i=1}^{n}\{Z_i\hat{\mu}_0(X_i)\varepsilon_0(X_i) + (1-Z_i)\hat{\mu}_0(X_i)\}.
\end{aligned}
$$

Identification and estimation under sensitivity analysis

Inverse propensity score weighting

- With known $\varepsilon_1(X)$ and $\varepsilon_0(X)$,

$$
\begin{aligned}
E\{Y(1)\} &= E\left\{w_1(X)\frac{Z}{e(X)}Y\right\}, \\
E\{Y(0)\} &= E\left\{w_0(X)\frac{1-Z}{1-e(X)}Y\right\},
\end{aligned}
$$

  where

$$
w_1(X) = e(X)+\{1-e(X)\}/\varepsilon_1(X),\ w_0(X) = e(X)\varepsilon_0(X)+1-e(X).
$$

- Estimator:

$$
\hat{\tau}^{\mathrm{ht}} = n^{-1}\sum_{i=1}^{n}\hat{w}_1(X_i)\frac{Z_iY_i}{\hat{e}(X_i)} - n^{-1}\sum_{i=1}^{n}\hat{w}_0(X_i)\frac{(1-Z_i)Y_i}{1-\hat{e}(X_i)}.
$$

Identification and estimation under sensitivity analysis

- When $\varepsilon_1(X) = \varepsilon_0(X) = 1$, reduces to the unconfoundedness assumption. Recover previous results.
- Motivates a combined strategy: doubly robust estimation (Bang and Robins, 2005; Bickel et al., 1993).

Doubly robust

- Outcome regression

$$
\begin{aligned}
\hat{\tau}^{\text{proj}} &= n^{-1}\sum_{i=1}^{n}\{Z_i\hat{\mu}_1(X_i) + (1 - Z_i)\hat{\mu}_1(X_i)/\varepsilon_1(X_i)\} \\
&\quad - n^{-1}\sum_{i=1}^{n}\{Z_i\hat{\mu}_0(X_i)\varepsilon_0(X_i) + (1 - Z_i)\hat{\mu}_0(X_i)\}.
\end{aligned}
$$

- IPW

$$
\hat{\tau}^{\text{ht}} = n^{-1}\sum_{i=1}^{n}\hat{w}_1(X_i)\frac{Z_iY_i}{\hat{e}(X_i)} - n^{-1}\sum_{i=1}^{n}\hat{w}_0(X_i)\frac{(1 - Z_i)Y_i}{1 - \hat{e}(X_i)}.
$$

Efficient influence function

- Under our definition of sensitivity parameters, the efficient influence functions for $E\{Y(1)\}$ and $E\{Y(0)\}$ are respectively

$$\begin{aligned}\phi_1 &= w_1(X)\frac{Z}{e(X)}Y - \frac{\{Z-e(X)\}\mu_1(X)}{e(X)\varepsilon_1(X)} - E\{Y(1)\},\\ \phi_0 &= w_0(X)\frac{1-Z}{1-e(X)}Y - \frac{\{e(X)-Z\}\mu_0(X)\varepsilon_0(X)}{1-e(X)} - E\{Y(0)\},\end{aligned}$$

  so the efficient influence function for τ is $\phi_1 - \phi_0$.
- An estimator constructed based on the EIF:

$$\begin{aligned}\hat{\tau}^{\mathrm{dr}} &= n^{-1}\sum_{i=1}^{n}\left[\hat{w}_1(X_i)\frac{Z_iY_i}{\hat{e}(X_i)} - \frac{\{Z_i-\hat{e}(X_i)\}\,\hat{\mu}_1(X_i)}{\hat{e}(X_i)\varepsilon_1(X_i)}\right]\\ &\quad -n^{-1}\sum_{i=1}^{n}\left[\hat{w}_0(X_i)\frac{(1-Z_i)Y_i}{1-\hat{e}(X_i)} - \frac{\{\hat{e}(X_i)-Z_i\}\,\hat{\mu}_0(X_i)\varepsilon_0(X_i)}{1-\hat{e}(X_i)}\right].\end{aligned}$$

- Can be written as modifications of $\hat{\tau}^{\mathrm{proj}}$ and $\hat{\tau}^{\mathrm{ht}}$.

Double robustness and semiparametric efficiency

$$
\begin{aligned}
\hat{\tau}^{\mathrm{dr}} \;=\; & n^{-1}\sum_{i=1}^{n}\left[\hat{w}_1(X_i)\frac{Z_iY_i}{\hat{e}(X_i)} - \frac{\{Z_i-\hat{e}(X_i)\}\,\hat{\mu}_1(X_i)}{\hat{e}(X_i)\varepsilon_1(X_i)}\right] \\
& -n^{-1}\sum_{i=1}^{n}\left[\hat{w}_0(X_i)\frac{(1-Z_i)Y_i}{1-\hat{e}(X_i)} - \frac{\{\hat{e}(X_i)-Z_i\}\,\hat{\mu}_0(X_i)\varepsilon_0(X_i)}{1-\hat{e}(X_i)}\right].
\end{aligned}
$$

- Double robustness: consistent if either the propensity score or the outcome model is correctly specified.
- Semiparametric efficiency bound is achieved with
  1. consistency of both models,
  2. mild requirements on their convergence rates.

Implementation—calibration of sensitivity parameters

- Observed data do not provide information on sensitivity parameters.
- How can we make meaningful progress?
  - A standard strategy: leave one covariate out.
  - Pretending an observed covariate is an unmeasured confounder.
  - Assume ignorability $Z \perp\!\!\!\perp \{Y(1), Y(0)\} \mid X$ and calculate

$$\varepsilon_z(X_{-j}) = \frac{E\{Y(z) \mid Z = 1, X_{-j}\}}{E\{Y(z) \mid Z = 0, X_{-j}\}}, \quad (z = 0, 1).$$

  - Use the range of $\hat{\varepsilon}_z(X_{-j})$ to specify the range of $(\varepsilon_1(X), \varepsilon_0(X))$.

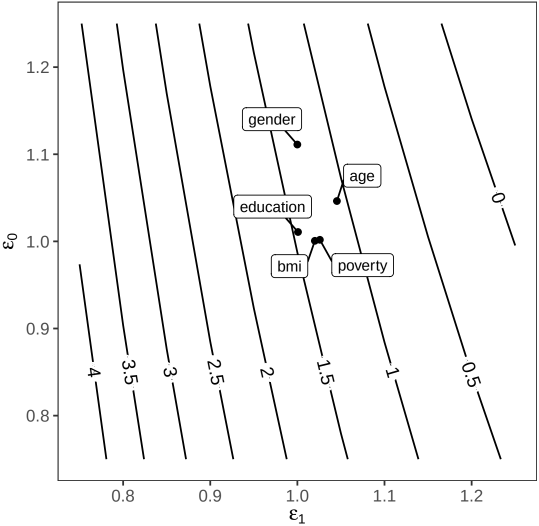
Sensitivity parameters: vary them with covariates or not

- Our theory allows for the dependence of sensitivity parameters on X.
  - Complicated in implementation.
  - Not easy to visualize the sensitivity analysis results.
- Practically, specify $(\varepsilon_1(X), \varepsilon_0(X)) = (\varepsilon_1, \varepsilon_0)$ independent of X.
  - Point estimates are monotonic in $\varepsilon_1, \varepsilon_0$: easy to interpret.
  - For non-negative outcomes: can also be interpreted as the worst-case result even with $(\varepsilon_1(X), \varepsilon_0(X))$ depending on X.
  - Formal result in the paper.

Demo with an example

- Re-analyze the observational study in Bazzano et al. (2003): whether cigarette smoking has a causal effect on homocysteine levels.
- Elevation of homocysteine level is a risk factor for cardiovascular disease.
- Data: the U.S. National Health and Nutrition Examination Survey 2005– 2006.
- Observed covariates: gender, age, education level, body mass index (BMI), and poverty.
- $\hat{\tau}^{\mathrm{dr}}$ for τ: 1.48 with a 95% confidence interval (0.78, 2.18).
- Unobserved confounders: genotype?

Demo with an example – visualization based on $\hat{\tau}^{\mathrm{dr}}$



- Point estimates as a function of $(\varepsilon_1, \varepsilon_0)$.
- Plot maximums of $\hat{\varepsilon}_z(X_{-j})$ for observed covariates.

Table of Contents

Extension to nonlinear causal parameters

- A more general class of nonlinear causal parameters $g(\mu_1, \mu_0)$.
- Previous results: $g(\mu_1, \mu_0) = \mu_1 - \mu_0$.
- Binary outcomes: the causal risk ratio and the causal odds ratio,
$$\mathrm{RR} = \frac{\mu_1}{\mu_0} \quad \text{and} \quad \mathrm{OR} = \frac{\mu_1/(1-\mu_1)}{\mu_0/(1-\mu_0)}.$$
- Plug-in estimators $g(\hat{\mu}_1^*, \hat{\mu}_0^*)$ for $* \in \{\text{pred, prod, ht, dr}\}$.

Extension to bias-corrected matching estimator

- Matching: impute the missing counterfactual $Y_i(0)$ for $Z_i = 1$ by finding M nearest neighbors of i in the control group.
- Use the average observed outcomes of the M nearest neighbors as the imputed value of the counterfactual $Y_i(0)$. Similar procedure for $Y_i(1)$.
- Matching-based estimator: average of the imputed individual treatment effect. Generally inconsistent (Abadie and Imbens, 2006).
- Abadie and Imbens (2011) proposed a bias-corrected version of the matching estimator by estimating the conditional outcome models $\mu_z(X)$ and combining it with the original matching estimator.

Extension to bias-corrected matching estimator

- Rewrite the bias-corrected matching estimator as (Lin et al., 2023):

$$\begin{aligned}\hat{\tau}_M^{\mathrm{bc}} &= \hat{\tau}^{\mathrm{reg}} + n^{-1}\sum_{i=1}^{n}\left\{1+\frac{K_M^1(X_i)}{M}\right\} Z_i \left\{Y_i - \hat{\mu}_1(X_i)\right\} \\ &\quad - n^{-1}\sum_{i=1}^{n}\left\{1+\frac{K_M^0(X_i)}{M}\right\}(1-Z_i)\left\{Y_i - \hat{\mu}_0(X_i)\right\},\end{aligned}$$

  where M is the fixed number of matches for each observation and $K_M^z(X_i)$ is the number of matched times of unit i in treatment group z, $z \in \{0, 1\}$.

Extension to bias-corrected matching estimator

- Under our sensitivity analysis framework, the bias-corrected matching estimator:

$$\begin{aligned}\hat{\tau}_M^{\mathrm{bc}} &= \hat{\tau}^{\mathrm{pred}} + n^{-1}\sum_{i=1}^{n}\frac{K_M^1(X_i)}{M}\frac{1}{\varepsilon_1(X_i)}Z_i\left\{Y_i - \hat{\mu}_1(X_i)\right\} \\ &\quad - n^{-1}\sum_{i=1}^{n}\frac{K_M^0(X_i)}{M}\varepsilon_0(X_i)(1-Z_i)\left\{Y_i - \hat{\mu}_0(X_i)\right\}.\end{aligned}$$

- Lin et al. (2023) views matching as a nonparametric method of estimating the propensity score. We essentially use $1 + K_M^1(X_i)/M$ and $1 + K_M^0(X_i)/M$ to estimate $1/e(X_i)$ and $1/\{1 - e(X_i)\}$, respectively.

Extension to multi-level treatment

- Observational studies with a multi-level treatment, $Z \in \{1, \ldots, K\}$.
- Each unit has K potential outcomes $\{Y(1), \ldots, Y(K)\}$ corresponding to the K treatment levels.
- Causal parameters of interest: comparisons of potential outcomes
$$\tau_c = \sum_{k=1}^{K} c_k E\{Y(k)\},$$
where $\sum_{k=1}^{K} c_k = 0$.
- For any two treatment levels k and l, the sensitivity parameters:
$$\varepsilon_{k,l}(X) = \frac{E\{Y(k) \mid Z = k, X\}}{E\{Y(k) \mid Z = l, X\}}.$$

Extension to multi-level treatment

- Identification:

$$\begin{aligned} E\{Y(k)\} &= \sum_{l=1}^{K} E\left\{1(Z=l)\frac{\mu_k(X)}{\varepsilon_{k,l}(X)}\right\} \\ &= \sum_{l=1}^{K} E\left\{\frac{e_l(X)}{\varepsilon_{k,l}(X)}\frac{1(Z=k)Y}{e_k(X)}\right\}, \end{aligned}$$

  where
  - $e_k(X) = \mathrm{pr}(Z = k \mid X)$: the generalized propensity score (Imbens, 2000; Imai and Van Dyk, 2004);
  - $\mu_k(X) = E\{Y \mid Z = k, X\}$: conditional outcome mean.
- Estimation:

$$\begin{aligned} \hat{\mu}_k^{\mathrm{dr}} &= \hat{\mu}_k^{\mathrm{reg}} + n^{-1}\sum_{i=1}^{n}\sum_{l=1}^{K}\frac{\hat{e}_l(X_i)1(Z_i=k)\{Y_i-\hat{\mu}_k(X_i)\}}{\varepsilon_{k,l}(X_i)\hat{e}_k(X_i)}, \\ \hat{\tau}_c^{dr} &= \sum_{k=1}^{K} c_k \hat{\mu}_k^{dr}. \end{aligned}$$

Other extensions in the paper

- Hajek-type weighting estimators.
- Average treatment effect on the treated:

$$\tau_{\mathrm{T}} = E\{Y(1) - Y(0) \mid Z = 1\} = E(Y \mid Z = 1) - E\{Y(0) \mid Z = 1\}.$$

- Survival outcomes under right-censoring:

$$\tau(t) = S_1(t) - S_0(t),$$

where $S_z(t) = \mathrm{pr}\{Y(z) > t\}$ denotes the potential survival functions for $z \in \{0, 1\}$.

- Controlled direct effect.
- Discussion on another sensitivity parameter, difference scale.

Summary

- **Flexible** sensitivity analysis framework.
- Simultaneously deal with weighting, outcome regression, and doubly robust estimators.
- Only requires simple modifications of the standard estimators.
- Extends to many other causal inference settings.
- Easy to implement – R package `saci`.

Thank you very much!

ArXiv: `https://arxiv.org/abs/2305.17643`

Abadie, A. and Imbens, G. W. (2006). Large sample properties of matching estimators for average treatment effects. *Econometrica* **74,** 235–267.

Abadie, A. and Imbens, G. W. (2011). Bias-corrected matching estimators for average treatment effects. *Journal of Business and Economic Statistics* **29,** 1–11.

Bang, H. and Robins, J. M. (2005). Doubly robust estimation in missing data and causal inference models. *Biometrics* **61,** 962–973.

Bazzano, L. A., He, J., Muntner, P., Vupputuri, S., and Whelton, P. K. (2003). Relationship between cigarette smoking and novel risk factors for cardiovascular disease in the united states. *Annals of Internal Medicine* **138,** 891–897.

Bickel, P. J., Klaassen, C. A., Ritov, Y., and Wellner, J. A. (1993). *Efficient and Adaptive Estimation for Semiparametric Models*. Baltimore: Johns Hopkins University Press.

Cornfield, J., Haenszel, W., Hammond, E. C., Lilienfeld, A. M., Shimkin, M. B., and Wynder, E. L. (1959). Smoking and lung cancer: recent evidence and a discussion of some questions. *Journal of the National Cancer Institute* **22,** 173–203.

Ding, P. (2024). *A First Course in Causal Inference*. New York: Chapman & Hall.

Ding, P. and VanderWeele, T. J. (2016). Sensitivity analysis without assumptions. *Epidemiology* **27,** 368–377.

Dorn, J. and Guo, K. (2022). Sharp sensitivity analysis for inverse propensity weighting via quantile balancing. *Journal of the American Statistical Association* **00,** 1–13.

Imai, K. and Van Dyk, D. A. (2004). Causal inference with general treatment regimes: Generalizing the propensity score. *Journal of the American Statistical Association* **99,** 854–866.

Imbens, G. W. (2000). The role of the propensity score in estimating dose-response functions. *Biometrika* **87,** 706–710.

Imbens, G. W. (2003). Sensitivity to exogeneity assumptions in program evaluation. *American Economic Review* **93,** 126–132.

Lin, D. Y., Psaty, B. M., and Kronmal, R. A. (1998). Assessing the sensitivity of regression results to unmeasured confounders in observational studies. *Biometrics* **54,** 948–963.

Lin, Z., Ding, P., and Han, F. (2023). Estimation based on nearest neighbor matching: from density ratio to average treatment effect. *Econometrica* **91,** 2187–2217.

Rosenbaum, P. R. (1987). Sensitivity analysis for certain permutation inferences in matched observational studies. *Biometrika* **74,** 13–26.

Rosenbaum, P. R. and Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika* **70,** 41–55.

VanderWeele, T. J. and Ding, P. (2017). Sensitivity analysis in observational research: introducing the e-value. *Annals of Internal Medicine* **167,** 268–274.

Zhao, Q., Small, D. S., and Bhattacharya, B. B. (2019). Sensitivity analysis for inverse probability weighting estimators via the percentile bootstrap. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **81,** 735–761.

More equivalent forms for $\hat{\tau}$ with outcome regression

$$
\begin{aligned}
\hat{\tau}^{\text{pred}} &= n^{-1}\sum_{i=1}^{n}\{Z_iY_i + (1-Z_i)\hat{\mu}_1(X_i)/\varepsilon_1(X_i)\} \\
&\quad - n^{-1}\sum_{i=1}^{n}\{Z_i\hat{\mu}_0(X_i)\varepsilon_0(X_i) + (1-Z_i)Y_i\},
\end{aligned}
$$

and

$$
\begin{aligned}
\hat{\tau}^{\text{proj}} &= n^{-1}\sum_{i=1}^{n}\{Z_i\hat{\mu}_1(X_i) + (1-Z_i)\hat{\mu}_1(X_i)/\varepsilon_1(X_i)\} \\
&\quad - n^{-1}\sum_{i=1}^{n}\{Z_i\hat{\mu}_0(X_i)\varepsilon_0(X_i) + (1-Z_i)\hat{\mu}_0(X_i)\}.
\end{aligned}
$$

More equivalent forms for $\hat{\tau}^{\mathrm{dr}}$

$$
\begin{aligned}
\hat{\tau}^{\mathrm{dr}} &= \hat{\tau}^{\mathrm{ht}} - n^{-1} \sum_{i=1}^{n} \check{Z}_i \left\{ \frac{\hat{\mu}_1(X_i)}{\hat{e}(X_i)\varepsilon_1(X_i)} + \frac{\hat{\mu}_0(X_i)\varepsilon_0(X_i)}{1-\hat{e}(X_i)} \right\} \\
&= \hat{\tau}^{\mathrm{pred}} + n^{-1} \sum_{i=1}^{n} \left\{ \frac{1-\hat{e}(X_i)}{\varepsilon_1(X_i)} \frac{Z_i \check{Y}_i}{\hat{e}(X_i)} - \hat{e}(X_i)\varepsilon_0(X_i) \frac{(1-Z_i)\check{Y}_i}{1-\hat{e}(X_i)} \right\} \\
&= \hat{\tau}^{\mathrm{proj}} + n^{-1} \sum_{i=1}^{n} \left\{ \hat{w}_1(X_i) \frac{Z_i \check{Y}_i}{\hat{e}(X_i)} - \hat{w}_0(X_i) \frac{(1-Z_i)\check{Y}_i}{1-\hat{e}(X_i)} \right\},
\end{aligned}
$$