Expeditious Ways to Derive Efficient Influence Functions for Causal Mediation Analysis and Multiple Robustness

Hun Lee

2022-08-31

Contents

Keywords: Efficient influence function; Regular and asymptotically linear (RAL) estimators; Natural (in)direct effect; Population intervention (in)direct effect; Stochastic intervention (in)direct effect; Double Robustness; Triple Robustness.

1. Introduction

Real-world data sets often fail to comply with parametric (distributional) model assumptions, and yet it is not uncommon that current scientific research and statistical practices overlook the violation of parametric model assumptions. Given convenient and interpretable features of parametric models, it is not surprising that parametric models are frequently used even when it is known that they are misspecified. However, the failure to use an appropriate model results in a biased estimate of a target parameter, leading to invalid confidence intervals and p-values.

To address the problems of violation of parametric model assumptions and misspecification of statistical models, increasing attention have been paid to developing nonparametric models and semiparametric models. In non/semi-parametric models, the type of estimators are often restricted to asymptotically linear estimators because most reasonable estimators belong to this class (Tsiatis, 2006), and an influence function should exist for an estimator for a parameter to be asymptotically linear. This paper focuses on deriving simple efficient influence functions and complex efficient influence functions for causal mediation analysis under a nonparametric model as well as their multiple robustness. The details of the methods for deriving effcient influence functions can be found in two recent journals (Hines et al., 2022; Kennedy, 2022). There are also many literature and resources available on the topic of parameter estimation in nonparametric and semiparametric models (Pfanzagl, 1990; Bickel et al., 1993; van der Vaart 2000; Tsiatis, 2006; van der Laan and Rubin, 2006; Kosorok, 2008; van der Laan and Rose, 2011; Vermeulen and Vansteelandt, 2011; Chernozhukov et al, 2018; Kennedy, 2018). Semiparametric approaches are also used to deal with data with missing not at random (MNAR) which refers to situations where the missingness of a varialble depends on unobserved variables (Sun et al., 2018; Malinsky et al., 2021).

Contemporary statistics also has seen increasing demand for tackling the problems that arise from high-dimensional data. This is because the large number of dimensions in data sets, also knonwn as curse of dimensionality, often causes large variance of the estimates and the problem of overfitting. To solve this problem, the methods of Bayesian nonparametric models also have been increasingly developed, such as BART (Hill, 2011; Hahn et al., 2020), Dirichlet Process (Chib and Hamilton, 2002; Karabatsos and Walker, 2012; Roy et al., 2018), Gaussian Process (Rasmussen, 2003; Ray and van der Vaart, 2020), the spike-and-slab prior (Antonelli et al., 2019), the Bayesian LASSO (Park and Casella, 2008; Mallick and Yi, 2014), the overall application of Bayesian nonparametric methods in high-dimensional setting (Oganisian and Roy, 2021; Linero and Antonelli 2022), the choice of priors in high dimensional regimes (Li et al, 2022), and Bayesian modeling with good frequentist properties in high dimensions (Antonelli et al, 2020).

This paper is motivated to show the usefulness of derivative rules with simple efficient influence functions as building blocks to derive efficient influence functions in an expeditious way. Many existing journals use the definition of pathwise differentiablility to get the efficient influences which can be computationally cumbersome. Thus, this paper aims to show an expeditious way to derive efficient influence functions for natural indirect effect (NIE), natural direct effect (NDE), population intervention indirect effect (PIIE), population intervention direct effect (PIDE), stochastic intervention indirect effect (SIIE), and stochastic intervention direct effect (SIDE) under a nonparametric model. Then, we will show the double or triple robustness of these efficient influence functions under model misspecificaiton. There are also literature and resources related to these topics (Tchetgen Tchetgen and Shpitser, 2012; Kennedy et al., 2017; Fulcher et al., 2019; Fulcher et al., 2020; Kennedy, 2019; Diaz et al., 2020; Hejazi et al., 2020; Xia and Chan 2021; Hejazi et al., 2022; van der Laan et al., 2022).

1.1 Defining Influence Function

An estimator $\hat{\Psi}_n$ for Ψ is called asymptotically linear if and only if there exists a q-dimensional random function $\varphi(X, \theta_0)$ such that

1. $E_{\theta_0}\{\varphi(X, \theta_0)\} = 0^{q\times 1}$,
2. $E_{\theta_0}\{\varphi(X, \theta_0)\varphi(X, \theta_0)^T\}$ is finite and non-singular.

and $\sqrt{n}(\hat{\Psi}_n - \Psi) = \dfrac{1}{\sqrt{n}} \sum_{i=1}^{n} \varphi(X, \theta_0) + o_p(1)$. The function $\varphi(X, \theta_0)$ is defined with respect to the true distribution $P(X; \theta_0)$ that generates the data, and the q-dimensional measurable random function $\varphi(X)$ is called the **influence function** of the estimator $\hat{\Psi}_n$. In nonparametric models, there is only one influence function, and that influence function is also the efficient influence function.

1.2 Defining a Parametric Submodel

Suppose $X_{1n}, X_{2n}, ..., X_{nn} \overset{i.i.d}{\sim} P(X; \theta_n)$, where θ_n is close to some fixed parameter θ^*. Then, an estimator $\hat{\Psi}_n(X_{1n}, ..., X_{nn})$ is regular if for each θ^*, $\sqrt{n}(\hat{\Psi}_n - \Psi_n)$ has a limiting distribution that does not depend on the local data generating process (LDGP). Informally, small perturbations of the true data generating distribution do not affect the limiting distribution of the estimator. Thus, we can use point mass contamination strategy under standard regularity conditions when deriving efficient influence functions for regular and asymptotically linear (RAL) estimators. To this end, we define a parametric submodel $\mathcal{P}_t = t\tilde{\mathcal{P}} + (1-t)\mathcal{P}$ which perturbs statistical functionals of the true observed data distribution $\mathcal{P}$ in the direction of $\tilde{\mathcal{P}}$. This parametric submodel is a smooth model $\mathcal{P}_t = \{\mathcal{P}_t : t \in [0, 1]\}$ that satisfies (i) $\mathcal{P}_t \subseteq \mathcal{P}$, and (ii) $\mathcal{P}_{t=0} = \mathcal{P}$.

1.3 A method to Derive Efficient Influence Functions

Now, we bring this idea to a directional derivative of pathwise differentiable functionals to derive the efficient influence function of Ψ of interest under a nonparametric model. Specifically, we compute the Gateaux derivative of the parameter at a chosen submodel in the direction of a point mass contamination, $\dfrac{d\Psi(\mathcal{P}_t)}{dt}\Big|_{t=0}$. Using the definition of pathwise differentiability,

$$\frac{d\Psi(\mathcal{P}_t)}{dt} = E[\varphi(O, \mathcal{P}_t)S_t(O)] = \mathcal{P}_t\{\varphi(O, \mathcal{P}_t)S_t(O)\}$$

where the score function $S_t(o)$ is the derivative of the log density with respect to t.

Then using the Riesz Representation Theorem (Hines et al., 2022),

$$\begin{aligned}
\frac{d\Psi(\mathcal{P}_t)}{dt} &= \mathcal{P}_t\{\varphi(O, \mathcal{P}_t)S_t(O)\} \\
&= \int \varphi(o, \mathcal{P}_t)S_t(o)d\mathcal{P}_t(o) \\
&= \int \varphi(o, \mathcal{P}_t)\{d\tilde{\mathcal{P}}(o) - d\mathcal{P}\} \\
&= (\tilde{\mathcal{P}} - \mathcal{P})\{\varphi(O, \mathcal{P}_t)\} \\
&= \tilde{\mathcal{P}}\{\varphi(O, \mathcal{P}_t)\}
\end{aligned}$$

At $t = 0$,

$$\frac{d\Psi(\mathcal{P}_t)}{dt}\Big|_{t=0} = (\tilde{\mathcal{P}} - \mathcal{P})\{\varphi(O, \mathcal{P})\} = \tilde{\mathcal{P}}\{\varphi(O, \mathcal{P})\},$$

since the efficient influence function $\mathcal{P}\{\varphi(O, \mathcal{P})\} = 0$ has mean zero.

Now, with the chosen submodel given by $\mathcal{P}_t(x) = t\mathbf{1}_{\tilde{x}}(x) + (1-t)\mathcal{P}(x)$ where $\mathbf{1}_{\tilde{x}}(x)$ is the Dirac delta function with respect to $\tilde{x}$, the result of the Gateaux derivative is the efficient influence function $\varphi(x; \mathcal{P})$ at observation x because the score for this submodel is $\dfrac{\mathbf{1}_{\tilde{x}}(x)}{d\mathcal{P}(x)} - 1$ at $t = 0$.

2. Deriving Efficient Influence Functions

2.1 Simple Efficient Influence Functions

In this section, we assume the data are discrete for computational convenience. Throughout, we will denote Y outcome, A exposure, Z mediator, and X covariates. Note that we still get the same result when we use integral instead of summation assuming the data are continuous.

Example 1. (probability density)

As a first simple example, consider the density at a given value x, $\Psi(\mathcal{P}) = f(x)$. Under the parametric submodel $\mathcal{P}_t(x) = t\mathbf{1}_{\tilde{x}}(x) + (1-t)\mathcal{P}(x)$, we readily get the following:

$$\Psi(\mathcal{P}_t) = t\mathbf{1}(X = x) + (1-t)f(x)$$

Taking a derivative with respect to t at t = 0 gives the efficient influence function

$$\begin{aligned}\frac{d\Psi(\mathcal{P}_t)}{dt}\Big|_{t=0} &= \mathbf{1}(X = x) - f(x)\\ &= \mathbf{1}(X = x) - \Psi(\mathcal{P})\end{aligned}$$

Checking if the expectation of the efficient influence function is zero,

$$\begin{aligned}E[\mathbf{1}(X = x) - f(x)] &= f(x) - \Psi(\mathcal{P})\\ &= \Psi(\mathcal{P}) - \Psi(\mathcal{P})\\ &= 0 \quad \square\end{aligned}$$

Example 2. (conditional probability density)

Consider next the conditional density at a given value z conditional on given values a and x, $\Psi(\mathcal{P}) = f(z|a, x)$.

Under the parametric submodel of example 1,

$$\begin{aligned}\Psi(\mathcal{P}_t) &= f_t(z|a, x)\\ &= \frac{f_t(z, a, x)}{f_t(a, x)}\end{aligned}$$

By the chain rule and the quotient rule for derivatives,

$$\begin{aligned}\frac{d\Psi(\mathcal{P}_t)}{dt}\Big|_{t=0} &= \frac{f_t'(z, a, x)}{f(a, x)} - \frac{f(z, a, x)f_t'(a, x)}{f(a, x)^2}\\ &= \frac{\mathbf{1}(Z = z, A = a, X = x) - f(z, a, x)}{f(a, x)} - \frac{f(z, a, x)}{f(a, x)^2}\{\mathbf{1}(A = a, X = x) - f(a, x)\}\\ &= \frac{\mathbf{1}(Z = z, A = a, X = x)}{f(a, x)} - f(z|a, x) - \frac{f(z|a, x)}{f(a, x)}\mathbf{1}(A = a, X = x) + f(z|a, x)\\ &= \frac{\mathbf{1}(A = a, X = x)}{f(a, x)}\{\mathbf{1}(Z = z) - f(z|a, x)\}\\ &= \frac{\mathbf{1}(A = a, X = x)}{f(a, x)}\{\mathbf{1}(Z = z) - \Psi(\mathcal{P})\}\end{aligned}$$

Checking if the expectation of the efficient influence function is zero,

$$\begin{aligned}
E\Big[\frac{\mathbf{1}(A=a,X=x)}{f(a,x)}\{\mathbf{1}(Z=z)-\Psi(\mathcal{P})\}\Big] &= \frac{f(z,a,x)}{f(a,x)}-\Psi(\mathcal{P}) \\
&= f(z|a,x)-\Psi(\mathcal{P}) \\
&= \Psi(\mathcal{P})-\Psi(\mathcal{P}) \\
&= 0 \quad \square
\end{aligned}$$

Example 3. (regression function 1)

Consider next the regression of Y on X for a given value x, $\Psi(\mathcal{P}) = E_{\mathcal{P}}(Y|X=x)$.

$$\Psi(\mathcal{P}_t) = \sum_y y\frac{f_t(y,x)}{f_t(x)}$$

By the chain rule and the quotient rule for derivatives,

$$\begin{aligned}
\frac{d\Psi(\mathcal{P}_t)}{dt}\Big|_{t=0} &= \sum_y y\Big\{\frac{f_t'(y,x)}{f(x)}-\frac{f(y,x)f_t'(x)}{f(x)^2}\Big\} \\
&= \sum_y y\Big\{\frac{1}{f(x)}\{\mathbf{1}(Y=y,X=x)-f(y,x)\}-\frac{f(y,x)}{f(x)^2}\{\mathbf{1}(X=x)-f(x)\}\Big\} \\
&= \sum_y y\Big\{\frac{\mathbf{1}(Y=y,X=x)}{f(x)}-f(y|x)-\frac{f(y|x)}{f(x)}\mathbf{1}(X=x)-f(y|x)\Big\} \\
&= \sum_y y\Big\{\frac{\mathbf{1}(X=x)}{f(x)}\{\mathbf{1}(Y=y)-f(y|x))\}\Big\} \\
&= \frac{\mathbf{1}(X=x)}{f(x)}\{Y-E[Y|x]) \\
&= \frac{\mathbf{1}(X=x)}{f(x)}\{Y-\Psi(\mathcal{P}))
\end{aligned}$$

Checking if the expectation of the efficient influence function is zero,

$$\begin{aligned}
E\Big[\frac{\mathbf{1}(X=x)}{f(x)}\{Y-\Psi(\mathcal{P}))\Big] &= \sum_x \Big[\frac{\mathbf{1}(X=x)}{f(x)}\{E[Y|x]-\Psi(\mathcal{P}))\Big]f(x) \\
&= \sum_x \Big[\frac{\mathbf{1}(X=x)}{f(x)}\{\Psi(\mathcal{P}-\Psi(\mathcal{P}))\Big]f(x) \\
&= 0 \quad \square
\end{aligned}$$

Example 4. (regression function 2)

Before moving on to complex examples, we finally consider the regression of Y on A and X for given values a and x, $\Psi(\mathcal{P}) = E_{\mathcal{P}}(Y|A=a,X=x)$.

$$\Psi(\mathcal{P}_t) = \sum_y y\frac{f_t(y,a,x)}{f_t(a,x)}$$

By the chain rule and the quotient rule for derivatives,

$$\frac{d\Psi(\mathcal{P}_t)}{dt}\Big|_{t=0} = \sum_y y\Big\{\frac{f_t'(y,a,x)}{f(a,x)} - \frac{f(y,a,x)f_t'(a,x)}{f(a,x)^2}\Big\}$$
$$= \frac{\mathbf{1}(A=a,X=x)}{f(a,x)}\{Y - E[Y|a,x]\}$$
$$= \frac{\mathbf{1}(A=a,X=x)}{f(a,x)}\{Y - \Psi(\mathcal{P})\}$$

Checking if the expectation of the efficient influence function is zero,

$$E\Big[\frac{\mathbf{1}(A=a,X=x)}{f(a,x)}\{Y - \Psi(\mathcal{P})\}\Big] = \sum_{a,x}\Big[\frac{\mathbf{1}(X=x)}{f(x)}\{E[Y|a,x] - \Psi(\mathcal{P})\}\Big]f(a,x)$$
$$= \sum_{a,x}\Big[\frac{\mathbf{1}(X=x)}{f(x)}\{\Psi(\mathcal{P} - \Psi(\mathcal{P}))\Big]f(a,x)$$
$$= 0 \quad \square$$

2.2 Complex Efficient Influence Functions

Now, our interest is in deriving the efficient influence function of certain causal effects under a nonparametric model, such as natural indirect effect (NIE), natural direct effect (NDE), population intervention indirect effect (PIIE), population intervention direct effect (PIDE), stochastic intervention indirect effect (SIIE), and stochastic intervention direct effect (SIDE).

The strategy is to pretend the data are discrete and treat infleunce functions $\mathbb{IF}$ as derivatives, allowing use of differentiation rules. Finally, we use the results of the previous exampls of the simple efficient influence functions as elementary units.

i.e.

$$\mathbb{IF}\{E[Y|A=a,x]\} = \frac{\mathbf{1}(A=a,X=x)}{f(a,x)}\{Y - E[Y|a,x]\}$$

Natural Indirect Effect

The NIE is the difference between the potential outcome under exposure value a and the potential outcome if exposure had taken value a but the mediator variable had taken the value that it would have under a^*:

$$\text{NIE} = E[Y\{a, Z(a)\} - Y\{a, Z(a^*)\}] = \Psi_1 - \Psi_2$$

Natural direct Effect

The NDE is the difference between the potential outcome if exposure had taken value a but the mediator variable had taken the value that it would have under a^* and the potential outcome under exposure value a^*:

$$\text{NDE} = E[Y\{a, Z(a^*) - Y\{a^*, Z(a^*)\}] = \Psi_2 - \Psi_3$$

Example 5. (Efficient Influence function of Ψ_1)

By the product rule and substituting the results of the simple efficient influence functions,

$$\begin{aligned}
\mathbb{IF}(\Psi_1) &= \mathbb{IF}\Big[\sum_x E[Y|A=a,x]f(x)\Big] \\
&= \sum_x \Big[\mathbb{IF}\{E[Y|a,x]\}f(x) + E[Y|a,x]\mathbb{IF}\{f(x)\}\Big] \\
&= \sum_x \Big[\frac{\mathbf{1}(A=a,X=x)}{f(a,x)}\{Y-E[Y|a,x]\}f(x) + E[Y|a,x]\{\mathbf{1}(X=x)-f(x)\}\Big] \\
&= \frac{\mathbf{1}(A=a)}{f(a|X)}\{Y-E[Y|a,X]\} + E[Y|a,X] - \Psi_1
\end{aligned}$$

Example 6. (Efficient Influence function of Ψ_2)

By the product rule and substituting the results of the simple efficient influence functions,

$$\begin{aligned}
\mathbb{IF}(\Psi_2) &= \mathbb{IF}\Big[\sum_{z,x} E[Y|z,a,x]f(z|a^*,x)f(x)\Big] \\
&= \sum_{z,x}\Big[\mathbb{IF}\{E[Y|z,a,x]\}f(z|a^*,x)f(x) \\
&\quad + \ E[Y|z,a,x]\mathbb{IF}\{f(z|a^*,x)\}f(x) \\
&\quad + \ E[Y|z,a,x]f(z|a^*,x)\mathbb{IF}\{f(x)\}\Big] \\
&= \sum_{z,x}\Big[\frac{\mathbf{1}(Z=z,A=a,X=x)}{f(z|a,x)f(a|x)f(x)}\{Y-E[Y|z,a,x]\}f(z|a^*,x)f(x) \\
&\quad + \ \frac{\mathbf{1}(A=a^*,X=x)}{f(a^*,x)}\{\mathbf{1}(Z=z)-f(z|a^*,x)\}E[Y|z,a,x]f(x) \\
&\quad + \{\mathbf{1}(X=x)-f(x)\}E[Y|z,a,x]f(z|a^*,x)\Big] \\
&= \frac{\mathbf{1}(A=a)}{f(Z|a,X)f(a|X)}\{Y-E[Y|Z,a,X]\}f(Z|a^*,X) \\
&\quad + \frac{\mathbf{1}(A=a^*)}{f(a^*|X)}\{E[Y|Z,a,X]-\sum_z E[Y|z,a,X]f(z|a^*,X)\} + \sum_z E[Y|z,a,X]f(z|a^*,X) - \Psi_2
\end{aligned}$$

Example 7. (Efficient Influence function of Ψ_3)

By the product rule and substituting the results of the simple efficient influence functions,

$$\begin{aligned}
\mathbb{IF}(\Psi_3) &= \mathbb{IF}\Big[\sum_x E[Y|A=a^*,x]f(x)\Big] \\
&= \frac{\mathbf{1}(A=a^*)}{f(a^*|X)}\{Y-E[Y|a^*,X]\} + E[Y|a^*,X] - \Psi_3
\end{aligned}$$

This example follows the same computational process as the example 5 as well as the result under different exposure value.

Population Intervention Indirect Effect

The PIIE is a novel measure of indirect effect corresponding to the effect of an intervention which changes the mediator from its natural value (i.e. its observed value) to the value that it would have had under exposure value a^* (Fulcher et al., 2018):

$$\text{PIIE} = E[Y\{A, Z(A)\} - Y\{A, Z(a^*)\}] = E[Y] - \Psi_4$$

Population Intervention Direct Effect

The PIDE is a novel measure of direct effect corresponding to the effect of an intervention which changes the exposure from its natural level to the value under intervention a^*, while keeping the mediator variable at the value that it would have under intervention a^* (Fulcher et al., 2018):

$$\text{PIDE} = E[Y\{A, Z(a^*)\} - Y\{a^*, Z(a^*)\}] = \Psi_4 - \Psi_3$$

Example 8. (Efficient Influence function of Ψ_4)

By the product rule and substituting the results of the simple efficient influence functions,

$$\begin{aligned}
\mathbb{IF}(\Psi_4) &= \mathbb{IF}\Big[\sum_{z,a,x} E[Y|z,a,x]f(z|a^*,x)f(a|x)f(x)\Big] \\
&= \sum_{z,a,x} \Big[\mathbb{IF}\{E[Y|z,a,x]\}f(z|a^*,x)f(a|x)f(x) \\
&\quad + E[Y|z,a,x]\mathbb{IF}\{f(z|a^*,x)\}f(a|x)f(x) \\
&\quad + E[Y|z,a,x]f(z|a^*,x)\mathbb{IF}\{f(a|x)\}f(x) \\
&\quad + E[Y|z,a,x]f(z|a^*,x)f(a|x)\mathbb{IF}\{f(x)\}\Big] \\
&= \sum_{z,a,x} \Big[\frac{\mathbf{1}(Z=z, A=a, X=x)}{f(z|a,x)f(a|x)f(x)}\{Y - E[Y|z,a,x]\}f(z|a^*,x)f(a|x)f(x) \\
&\quad + \frac{\mathbf{1}(A=a^*, X=x)}{f(a^*,x)}\{\mathbf{1}(Z=z) - f(z|a^*,x)\}E[Y|z,a,x]f(a|x)f(x) \\
&\quad + \frac{\mathbf{1}(X=x)}{f(x)}\{\mathbf{1}(A=a) - f(a|x)\}E[Y|z,a,x]f(z|a^*,x)f(x) \\
&\quad + \{\mathbf{1}(X=x) - f(x)\}E[Y|z,a,x]f(z|a^*,x)f(a|x)\Big] \\
&= \frac{f(Z|a^*,X)}{f(Z|A,X)}\{Y - E[Y|A,Z,X]\} \\
&\quad + \frac{\mathbf{1}(A=a^*)}{f(a^*|X)}\Big\{\sum_a E[Y|Z,a,X]f(a|X) - \sum_{z,a} E[Y|z,a,X)f(z|a^*,X)f(a|X)\Big\} \\
&\quad + \sum_z E[Y|z,A,X]f(z|a^*,X) - \Psi_4
\end{aligned}$$

Stochastic Intervention Indirect Effect

The SIIE is the difference between the potential outcome if exposure were drawn from a post-intervention distribution $g_\delta(a|x)$ and the potential outcome if exposure were drawn from a post-intervention distribution $g_\delta(a|x)$ while keeping the distribution of the mediator fixed:

$$\text{SIIE} = E[Y\{A_\delta, Z(A_\delta)\} - Y\{A_\delta, Z\}] = \Psi_5 - \Psi_6$$

where A_δ denotes a draw from user-specified $g_\delta(a|x)$ and δ is a user-given value (Kennedy, 2019; Hejazi et al., 2022).

Stochastic Intervention Direct Effect

The SIDE is the difference between the potential outcome if the potential outcome if exposure were drawn from a post-intervention distribution $g_\delta(a|x)$ while keeping the distribution of the mediator fixed and the mean of outcome.

$$\text{SIDE} = E[Y\{A_\delta, Z\} - Y\{A, Z\}] = \Psi_6 - E[Y]$$

Example 9. (Efficient Influence function of Ψ_5)

By the product rule and substituting the results of the simple efficient influence functions,

$$\begin{aligned}
\mathbb{IF}(\Psi_5) &= \mathbb{IF}\Big[\sum_{z,a,x} E[Y|a,x] g_\delta(a|x) f(z|a,x) f(x)\Big] \\
&= \sum_{z,a,x} \Big[\mathbb{IF}\{E[Y|a,x]\} g_\delta(a|x) f(z|a,x) f(x) \\
&\qquad + E[Y|a,x] g_\delta(a|x) \mathbb{IF}\{f(z|a,x)\} f(x) \\
&\qquad + E[Y|a,x] g_\delta(a|x) f(z|a,x) \mathbb{IF}\{f(x)\} \\
&= \sum_{z,a,x} \Big[\frac{\mathbf{1}(A=a, X=x)}{f(z|a,x) f(a|x) f(x)} \{Y - E[Y|a,x]\} g_\delta(a|x) f(z|a,x) f(x) \\
&\qquad + \frac{\mathbf{1}(A=a, X=x)}{f(a,x)} \{\mathbf{1}(Z=z) - f(z|a,x)\} E[Y|z,a,x] g_\delta(a|x) f(x) \\
&\qquad + \{\mathbf{1}(X=x) - f(x)\} E[Y|a,x] g_\delta(a|x) f(z|a,x) \\
&= \sum_z \frac{g_\delta(A|X)}{f(A|X)} \{Y - E[Y|A,X]\} f(z|A,X) \\
&\qquad + \frac{g_\delta(A|X)}{f(A|X)} \{\sum_z \mathbf{1}(Z=z) E[Y|A,X] - \sum_z E[Y|A,X] f(z|A,X)\} \\
&\qquad + \sum_{z,a} E[Y|a,X] g_\delta(a|X) f(z|a,X) - \Psi_5
\end{aligned}$$

Example 10. (Efficient Influence function of Ψ_6)

By the product rule and substituting the results of the simple efficient influence functions,

$$
\begin{aligned}
\mathbb{IF}(\Psi_6) &= \mathbb{IF}\Big[\sum_{z,a,x} E[Y|a,x]g_\delta(a|x)f(z|x)f(x)\Big] \\
&= \sum_{z,a,x}\Big[\mathbb{IF}\{E[Y|a,x]\}g_\delta(a|x)f(z|x)f(x) \\
&\quad + E[Y|a,x]g_\delta(a|x)\mathbb{IF}\{f(z|x)\}f(x) \\
&\quad + E[Y|a,x]g_\delta(a|x)f(z|x)\mathbb{IF}\{f(x)\} \\
&= \sum_{z,a,x}\Big[\frac{\mathbf{1}(A=a,X=x)}{f(z|a,x)f(a|x)f(x)}\{Y-E[Y|a,x]\}g_\delta(a|x)f(z|a,x)f(x) \\
&\quad + \frac{\mathbf{1}(X=x)}{f(x)}\{\mathbf{1}(Z=z)-f(z|x)\}E[Y|a,x]g_\delta(a|x)f(x) \\
&\quad + \{\mathbf{1}(X=x)-f(x)\}E[Y|a,x]g_\delta(a|x)f(z|x) \\
&= \sum_z \frac{g_\delta(A|X)}{f(A|X)}\{Y-E[Y|A,X]\}f(z|X) \\
&\quad + \sum_{z,a}\mathbf{1}(Z=z)E[Y|a,X]g_\delta(a|X) - \sum_{z,a}E[Y|a,X]g_\delta(a|X)f(z|X) \\
&\quad + \sum_{z,a}E[Y|a,X]g_\delta(a|X)f(z|X) - \Psi_6 \\
&= \sum_z \frac{g_\delta(A|X)}{f(A|X)}\{Y-E[Y|A,X]\}f(z|X) \\
&\quad + \sum_{z,a}\mathbf{1}(Z=z)E[Y|a,X]g_\delta(a|X) - \Psi_6
\end{aligned}
$$

3. Robustness of Influence Functions under Model Misspecification

With the premise that the influence functions are for asymptotically linear estimators under standard regularity conditions, and by the central by the central limit theorem and Slutsky's theorem we obtain that

$$\sqrt{n}(\hat{\Psi}-\Psi) \rightsquigarrow N(0, E[\varphi\varphi^T])$$

where $E[\varphi\varphi^T] < \infty$ and nonsingular as aforementioned in section 1.1. Now, we want to show these properties still hold under model misspecification by showing that $E[\varphi^{eff}] = 0$.

3.1 Double Robustness

Example 5. (continued).

The efficient influence function has expectation 0 if one of the following scenarios holds:

1.$E[Y|a,x]$ is correct.

2.$f(a|x)$ is correct.

1. $E[Y|a,x]$ is correctly specified and $\tilde{f}(a|x)$ misspecified

$$
\begin{aligned}
E[\varphi^{eff}] &= E\Big[\frac{\mathbf{1}(A=a)}{\tilde{f}(a|X)}\{Y - E[Y|a,X]\} + E[Y|a,X] - \Psi_1\Big] \\
&= 0 + \sum_x E[Y|a,x]f(x) - \Psi_1 \\
&= \Psi_1 - \Psi_1 \\
&= 0 \quad \square
\end{aligned}
$$

2. $f(a|x)$ is correctly specified and $\tilde{E}[Y|a,x]$ misspecified

$$
\begin{aligned}
E[\varphi^{eff}] &= E\Big[\frac{\mathbf{1}(A=a)}{f(a|X)}\{Y - \tilde{E}[Y|a,X]\} + \tilde{E}[Y|a,X] - \Psi_1\Big] \\
&= \sum_{a,x}\Big\{\frac{\mathbf{1}(A=a)}{f(a|x)}\{E[Y|a,x] - \tilde{E}[Y|a,x]\}\Big\}f(a|x)f(x) + \sum_x \tilde{E}[Y|a,x]f(x) - \Psi_1 \\
&= \sum_x E[Y|a,x]f(x) - \sum_x \tilde{E}[Y|a,x]f(x) + \sum_x \tilde{E}[Y|a,x]f(x) - \Psi_1 \\
&= \Psi_1 + 0 - \Psi_1 \\
&= 0 \quad \square
\end{aligned}
$$

Example 8. (continued).

The efficient influence function has expectation 0 if one of the following scenarios holds:

1.$E[Y|z,a,x]$ and $f(a|x)$ are correct.

2.$f(z|a,x)$ is correct.

1. $E[Y|z,a,x]$ and $f(a|x)$ are correctly specified and $\tilde{f}(z|a,x)$ misspecified

$$
\begin{aligned}
E[\varphi^{eff}] &= E\Big[\frac{\tilde{f}(Z|a^*,X)}{\tilde{f}(Z|A,X)}\{Y - E[Y|A,Z,X]\} \\
&\quad + \frac{\mathbf{1}(A=a^*)}{f(a^*|X)}\Big\{\sum_a E[Y|Z,a,X]f(a|X) - \sum_{z,a} E[Y|z,a,X)\tilde{f}(z|a^*,X)f(a|X)\Big\} \\
&\quad + \sum_z E[Y|z,A,X]\tilde{f}(z|a^*,X) - \Psi_4\Big] \\
&= 0 + \sum_{z,a',x}\frac{\mathbf{1}(a'=a^*)}{f(a^*|x)}\{\sum_a E[Y|z,a,x]f(a|x)\}f(z,a',x) \\
&\quad - \sum_{a',x}\frac{\mathbf{1}(a'=a^*)}{f(a^*|x)}\{\sum_{z,a} E[Y|z,a,x)\tilde{f}(z|a^*,x)f(a|x)\}f(a'|x)f(x) \\
&\quad + \sum_z\sum_{a,x} E[Y|z,a,x]\tilde{f}(z|a^*,x)f(a|x)f(x) - \Psi_4 \\
&= \sum_{z,a,x} E[Y|z,a,x]f(z|a^*,x)f(a|x)f(x) - \sum_{z,a,x} E[Y|z,a,x]\tilde{f}(z|a^*,x)f(a|x)f(x) \\
&\quad + \sum_{z,a,x} E[Y|z,A,X]\tilde{f}(z|a^*,X)f(a|x)f(x) - \Psi_4 \\
&= \Psi_4 + 0 - \Psi_4 \\
&= 0 \quad \square
\end{aligned}
$$

2. $f(z|a,x)$ is correctly specified and $\tilde{E}[Y|z,a,x]$ & $\tilde{f}(a|x)$ misspecified

$$
\begin{aligned}
E[\varphi^{eff}] &= E\Big[\frac{f(Z|a^*,X)}{f(Z|A,X)}\{Y-\tilde{E}[Y|A,Z,X]\} \\
&\quad + \frac{\mathbf{1}(A=a^*)}{\tilde{f}(a^*|X)}\Big\{\sum_a \tilde{E}[Y|Z,a,X]\tilde{f}(a|X) - \sum_{z,a}\tilde{E}[Y|z,a,X)f(z|a^*,X)\tilde{f}(a|X)\Big\} \\
&\quad + \sum_z \tilde{E}[Y|z,A,X]f(z|a^*,X) - \Psi_4\Big] \\
&= \sum_{z,a,x}\{E[Y|z,a,x]-\tilde{E}[Y|a,z,x]\}f(z|a^*,x)f(a|x)f(x) \\
&\quad + \sum_{z,a',x}\frac{\mathbf{1}(a'=a^*)}{\tilde{f}(a^*|x)}\{\sum_a \tilde{E}[Y|z,a,x]\tilde{f}(a|x)\}f(z|a',x)f(a'|x)f(x) \\
&\quad - \sum_{a',x}\frac{\mathbf{1}(a'=a^*)}{\tilde{f}(a^*|x)}\{\sum_{z,a}\tilde{E}[Y|z,a,x)f(z|a^*,x)\tilde{f}(a|X)\}f(a'|x)f(x) \\
&\quad + \sum_z\sum_{a,x}\tilde{E}[Y|z,a,x]f(z|a^*,x)f(a|x)f(x) - \Psi_4 \\
&= \Psi_4 - \sum_{z,a,x}\tilde{E}[Y|z,a,x]f(z|a^*,x)f(a|x)f(x) \\
&\quad + \sum_{z,a,x}\frac{\tilde{f}(a|x)}{\tilde{f}(a^*|x)}\tilde{E}[Y|z,a,x]f(z|a^*,x)f(a^*|x)f(x) - \sum_{z,a,x}\frac{\tilde{f}(a|x)}{\tilde{f}(a^*|x)}\tilde{E}[Y|z,a,x]f(z|a^*,x)f(a^*|x)f(x) \\
&\quad + \sum_{z,a,x}\tilde{E}[Y|z,a,x]f(z|a^*,X)f(a|x)f(x) - \Psi_4 \\
&= \Psi_4 + 0 + 0 - \Psi_4 \\
&= 0 \quad \square
\end{aligned}
$$

Example 10. (continued).

The efficient influence function has expectation 0 if one of the following scenarios holds:

1.$E[Y|a,x]$ is correct.

2.$f(z|x)$ and $f(a|x)$ are correct.

1. $E[Y|a,x]$ is correctly specified and $\tilde{f}(z|x)$ & $\tilde{f}(a|x)$ misspecified

$$
\begin{aligned}
E[\varphi^{eff}] &= E\Big[\sum_z \frac{g_\delta(A|X)}{\tilde{f}(A|X)}\{Y-E[Y|A,X]\}\tilde{f}(z|X) \\
&\quad + \sum_{z,a}\mathbf{1}(Z=z)E[Y|a,X]g_\delta(a|X) - \Psi_6\Big] \\
&= 0 + \sum_{z,a}\sum_{z',x}\mathbf{1}(z'=z)E[Y|a,x]g_\delta(a|X)f(z'|x)f(x) - \Psi_6 \\
&= \sum_{z,a,x}E[Y|a,x]g_\delta(a|X)f(z|x)f(x) - \Psi_6 \\
&= \Psi_6 - \Psi_6 \\
&= 0 \quad \square
\end{aligned}
$$

2. $f(z|x)$ and $f(a|x)$ are correctly specified and $\tilde{E}(Y|a,x)$ misspecified

$$
\begin{aligned}
E[\varphi^{eff}] &= E\Big[\sum_{z} \frac{g_\delta(A|X)}{f(A|X)}\{Y - \tilde{E}[Y|A,X]\}f(z|X) \\
&\qquad + \sum_{z,a} \mathbf{1}(Z=z)\tilde{E}[Y|a,X]g_\delta(a|X) - \Psi_6\Big] \\
&= \sum_{z}\sum_{a,x} \frac{g_\delta(a|x)}{f(a|x)} E[Y|a,x]f(z|x)f(a,x) - \sum_{z}\sum_{a,x} \frac{g_\delta(a|x)}{f(a|x)}\tilde{E}[Y|A,X]\}f(z|x)f(a,x) \\
&= \sum_{z,a}\sum_{z',a} \mathbf{1}(z'=z)\tilde{E}[Y|a,X]g_\delta(a|X)f(z'|x)f(x) - \Psi_6 \\
&= \Psi_6 - \sum_{z,a,x} \tilde{E}[Y|a,X]g_\delta(a|X)f(z|x)f(x) + \sum_{z,a,x} \tilde{E}[Y|a,X]g_\delta(a|X)f(z|x)f(x) - \Psi_6 \\
&= \Psi_6 + 0 - \Psi_6 \\
&= 0 \quad \square
\end{aligned}
$$

3.2 Triple Robustness

Example 6. (continued).

The efficient influence function has expectation 0 if one of the following scenarios holds:

1. $E[Y|z,a,x]$ and $f(a|x)$ are correct.
2. $E[Y|z,a,x]$ and $f(z|a,x)$ are correct.
3. $f(z|a,x)$ and $f(a|x)$ are correct.

1. $E[Y|z,a,x]$ and $f(a|x)$ are correctly specified and $\tilde{f}(z|a,x)$ misspecified

$$
\begin{aligned}
E[\varphi^{eff}] &= E\Big[\frac{\mathbf{1}(A=a)}{\tilde{f}(Z|a,X)f(a|X)}\{Y - E[Y|Z,a,X]\}\tilde{f}(Z|a^*,X) \\
&\quad + \frac{\mathbf{1}(A=a^*)}{f(a^*|X)}\{E[Y|Z,a,X] - \sum_{z} E[Y|z,a,X]\tilde{f}(z|a^*,X)\} \\
&\quad + \sum_{z} E[Y|z,a,X]\tilde{f}(z|a^*,X) \; - \Psi_2\Big] \\
&= 0 + \sum_{z,a',x} \frac{\mathbf{1}(a'=a^*)}{f(a^*|x)} E[Y|z,a,x]f(z|a',x)f(a'|x)f(x) \\
&\quad - \sum_{z}\sum_{a',x} \frac{\mathbf{1}(a'=a^*)}{f(a^*|x)} E[Y|z,a,x]\tilde{f}(z|a^*,x)f(a'|x)f(x) \\
&\quad - \sum_{z}\sum_{x} E[Y|z,a,x]\tilde{f}(z|a^*,x)f(x) - \Psi 2 \\
&= \sum_{z,x} E[Y|z,a,x]f(z|a^*,x)f(x) - \sum_{z,x} E[Y|z,a,x]\tilde{f}(z|a^*,x)f(x) + \sum_{z,x} E[Y|z,a,x]\tilde{f}(z|a^*,x)f(x) - \Psi_2 \\
&= \Psi_2 + 0 - \Psi_2 \\
&= 0 \quad \square
\end{aligned}
$$

2. <u>$E[Y|z,a,x]$ and $f(z|a,x)$ are correctly specified and $\tilde{f}(a|x)$ misspecified</u>

$$
\begin{aligned}
E[\varphi^{eff}] &= E\Big[\frac{\mathbf{1}(A=a)}{f(Z|a,X)\tilde{f}(a|X)}\{Y - E[Y|Z,a,X]\}f(Z|a^*,X) \\
&\quad + \frac{\mathbf{1}(A=a^*)}{\tilde{f}(a^*|X)}\{E[Y|Z,a,X] - \sum_z E[Y|z,a,X]f(z|a^*,X)\} \\
&\quad + \sum_z E[Y|z,a,X]f(z|a^*,X) \ - \Psi_2\Big] \\
&= 0 + \sum_{z,a',x} \frac{\mathbf{1}(a'=a^*)}{\tilde{f}(a^*|x)} E[Y|z,a,x]f(z|a',x)f(a'|x)f(x) \\
&\quad - \sum_z \sum_{a',x} \frac{\mathbf{1}(a'=a^*)}{\tilde{f}(a^*|x)} E[Y|z,a,x]f(z|a^*,x)f(a'|x)f(x) \\
&\quad + \sum_z \sum_x E[Y|z,a,x]f(z|a^*,x)f(x) - \Psi_2 \\
&= \sum_{z,,x} \frac{1}{\tilde{f}(a^*|x)} E[Y|z,a,x]f(z|a^*,x)f(a^*|x)f(x) - \sum_{z,,x} \frac{1}{\tilde{f}(a^*|x)} E[Y|z,a,x]f(z|a^*,x)f(a^*|x)f(x) \\
&\quad + \sum_{z,x} E[Y|z,a,x]f(z|a^*,x)f(x) - \Psi_2 \\
&= 0 + \Psi_2 - \Psi_2 \\
&= 0 \ \ \square
\end{aligned}
$$

3. <u>$f(z|a,x)$ and $f(a|x)$ are correctly specified and $\tilde{E}(Y|z,a,x)$ misspecified</u>

$$
\begin{aligned}
E[\varphi^{eff}] &= E\Big[\frac{\mathbf{1}(A=a)}{f(Z|a,X)f(a|X)}\{Y - \tilde{E}[Y|Z,a,X]\}f(Z|a^*,X) \\
&\quad + \frac{\mathbf{1}(A=a^*)}{f(a^*|X)}\{\tilde{E}[Y|Z,a,X] - \sum_z \tilde{E}[Y|z,a,X]f(z|a^*,X)\} \\
&\quad + \sum_z \tilde{E}[Y|z,a,X]f(z|a^*,X) \ - \Psi_2\Big] \\
&= \sum_{z,a,x} \frac{f(z|a^*,x)}{f(z|a,x)f(a|x)}\{E[Y|z,a,x] - \tilde{E}[Y|z,a,x]\}f(z|a,x)f(a|x)f(x) \\
&\quad + \sum_{z,a',x} \frac{\mathbf{1}(a'=a^*)}{f(a^*|x)} \tilde{E}[Y|z,a,x]f(z|a',x)f(a'|x)f(x) \\
&\quad - \sum_z \sum_{a',x} \frac{\mathbf{1}(a'=a^*)}{f(a^*|x)} \tilde{E}[Y|z,a,x]f(z|a^*,x)f(a'|x)f(x) \\
&\quad + \sum_z \sum_x \tilde{E}[Y|z,a,x]f(z|a^*,x)f(x) - \Psi_2 \\
&= \sum_{z,x} E[Y|z,a,x]f(z|a^*,x)f(x) - \sum_{z,x} \tilde{E}[Y|z,a,x]f(z|a^*,x)f(x) \\
&\quad + \sum_{z,x} \tilde{E}[Y|z,a,x]f(z|a^*,x)f(x) - \sum_{z,x} \tilde{E}[Y|z,a,x]f(z|a^*,x)f(x) \\
&\quad + \sum_{z,x} \tilde{E}[Y|z,a,x]f(z|a^*,x)f(x) - \Psi_2 \\
&= \Psi_2 + 0 + 0 - \Psi_2 \\
&= 0 \ \ \square
\end{aligned}
$$

Example 9. (continued).

The efficient influence function has expectation 0 if one of the following scenarios holds:

1. $E[Y|a,x]$ and $f(a|x)$ are correct.
2. $E[Y|a,x]$ and $f(z|a,x)$ are correct.
3. $f(z|a,x)$ and $f(a|x)$ are correct.

1. $E[Y|a,x]$ and $f(a|x)$ are correctly specified and $\tilde{f}(z|a,x)$ misspecified

$$
\begin{aligned}
E[\varphi^{eff}] &= E\Big[\sum_z \frac{g_\delta(A|X)}{f(A|X)}\{Y - E[Y|A,X]\}\tilde{f}(z|A,X) \\
&\qquad + \frac{g_\delta(A|X)}{f(A|X)}\{\sum_z \mathbf{1}(Z=z)E[Y|A,X] - \sum_z E[Y|A,X]\tilde{f}(z|A,X)\} \\
&\qquad + \sum_{z,a} E[Y|a,X]g_\delta(a|X)\tilde{f}(z|a,X) - \Psi_5\Big] \\
&= 0 + \sum_{z',a,x}\sum_z \frac{g_\delta(a|x)}{f(a|x)}\mathbf{1}(z'=z)E[Y|a,x]f(z'|a,x)f(a|x)f(x) \\
&\qquad - \sum_z \sum_{a,x} \frac{g_\delta(a|x)}{f(a|x)}\sum_z E[Y|a,x]\tilde{f}(z|a,x)f(a|x)f(x) \\
&\qquad + \sum_{z,a}\sum_x E[Y|a,x]g_\delta(a|x)\tilde{f}(z|a,x)f(x) - \Psi_5 \\
&= \sum_{z,a,x} E[Y|a,x]g_\delta(a|x)f(z|a,x)f(x) - \sum_{z,a,x} E[Y|a,x]g_\delta(a|x)\tilde{f}(z|a,x)f(x) \\
&\qquad + \sum_{z,a,x} E[Y|a,x]g_\delta(a|x)\tilde{f}(z|a,x)f(x) - \Psi_5 \\
&= \Psi_5 + 0 - \Psi_5 \\
&= 0 \quad \square
\end{aligned}
$$

2. $f(z|a,x)$ and $f(a|x)$ are correctly specified and $\tilde{E}(Y|a,x)$ misspecified

$$
\begin{aligned}
E[\varphi^{eff}] &= E\Big[\sum_z \frac{g_\delta(A|X)}{\tilde{f}(A|X)}\{Y - E[Y|A,X]\}f(z|A,X) \\
&\qquad + \frac{g_\delta(A|X)}{\tilde{f}(A|X)}\{\sum_z \mathbf{1}(Z=z)E[Y|A,X] - \sum_z E[Y|A,X]f(z|A,X)\} \\
&\qquad + \sum_{z,a} E[Y|a,X]g_\delta(a|X)f(z|a,X) - \Psi_5\Big] \\
&= 0 + \sum_{z',a,x}\sum_z \frac{g_\delta(a|x)}{\tilde{f}(a|x)}\mathbf{1}(Z=z)E[Y|a,x]f(z',a,x) - \sum_z\sum_{a,x}\frac{g_\delta(a|x)}{\tilde{f}(a|x)}E[Y|a,x]f(z|a,x)f(a|x) \\
&\qquad - \sum_{z,a}\sum_x E[Y|a,x]g_\delta(a|x)f(z|a,x)f(x) - \Psi_5 \\
&= \sum_{z,a,x}\frac{f(a|x)}{\tilde{f}(a|x)}E[Y|a,x]g_\delta(a|x)f(z|a,x)f(x) - \sum_{z,a,x}\frac{f(a|x)}{\tilde{f}(a|x)}E[Y|a,x]g_\delta(a|x)f(z|a,x)f(x) + \Psi_5 - \Psi_5 \\
&= 0 + \Psi_5 - \Psi_5 \\
&= 0 \quad \square
\end{aligned}
$$

3. $f(z|a,x)$ and $f(a|x)$ are correctly specified and $\tilde{E}(Y|a,x)$ misspecified

$$
\begin{aligned}
E[\varphi^{eff}] = E\Big[& \sum_z \frac{g_\delta(A|X)}{f(A|X)}\{Y - \tilde{E}[Y|A,X]\}f(z|A,X) \\
& + \frac{g_\delta(A|X)}{f(A|X)}\{\sum_z \mathbf{1}(Z=z)\tilde{E}[Y|A,X] - \sum_z \tilde{E}[Y|A,X]f(z|A,X)\} \\
& + \sum_{z,a} \tilde{E}[Y|a,X]g_\delta(a|X)f(z|a,X) - \Psi_5\Big] \\
= & \sum_z \sum_{a,x} \frac{g_\delta(a|x)}{f(a|x)} E[Y|a,x]f(z|a,x)f(a,x) - \sum_z \sum_{a,x} \frac{g_\delta(a|x)}{f(a|x)} \tilde{E}[Y|a,x]f(z|a,x)f(a,x) \\
& + \sum_{z',a,x} \sum_z \frac{g_\delta(a|x)}{f(a|x)} \mathbf{1}(z'=z)\tilde{E}[Y|a,x]f(z',a,x) - \sum_z \sum_{a,x} \frac{g_\delta(a|x)}{f(a|x)} \tilde{E}[Y|a,x]f(z|a,x)f(a,x) \\
& + \sum_{z,a} \sum_x \tilde{E}[Y|a,x]g_\delta(a|x)f(z|a,x)f(x) - \Psi_5 \\
= & \sum_{z,a,x} E[Y|a,x]g_\delta(a|x)f(z|a,x)f(x) - \sum_{z,a,x} \tilde{E}[Y|a,x]g_\delta(a|x)f(z|a,x)f(x) \\
& + \sum_{z,a,x} \tilde{E}[Y|a,x]g_\delta(a|x)f(z|a,x)f(x) - \sum_{z,a,x} \tilde{E}[Y|a,x]g_\delta(a|x)f(z|a,x)f(x) \\
& + \sum_{z,a,x} \tilde{E}[Y|a,x]g_\delta(a|x)f(z|a,x)f(x) - \Psi_5 \\
= & \Psi_5 + 0 + 0 - \Psi_5 \\
= & 0 \quad \square
\end{aligned}
$$

4. References

Antonelli, J., Parmigiani, G., and Dominici, F. (2019). High-dimensional confounding adjustment using continuous spike and slab priors. Bayesian analysis, 14(3), 805.

Antonelli, J., Papadogeorgou, G., and Dominici, F. (2022). Causal inference in high dimensions: A marriage between Bayesian modeling and good frequentist properties. Biometrics, 78(1), 100–114.

Sun, B., Liu, L., Miao, W., Wirth, K., Robins, J., and Tchetgen Tchetgen, E. J. (2018). Semiparametric Estimation with Data Missing Not at Random Using an Instrumental Variable. Statistica Sinica, 28(4), 1965–1983.

Bickel, P. J., Klaassen, C. A., Bickel, P. J., Ritov, Y., Klaassen, J., Wellner, J. A., and Ritov, Y. (1993). Efficient and adaptive estimation for semiparametric models, volume 4. Johns Hopkins University Press Baltimore.

Chernozhukov, V., Chetverikov, D., Demirer, M., Duflo, E., Hansen, C., Newey, W., and Robins, J. (2018). Double/debiased machine learning for treatment and structural parameters. The Econometrics Journal, 21(1):C1–C68.

Chib, S. and Hamilton, B. H. (2002). Semiparametric bayes analysis of longitudinal data treatment models. Journal of Econometrics, 110(1), 67–89.

Díaz, I., Hejazi, N. S., Rudolph, K. E.m, and van Der Laan, M. J. (2021). Nonparametric efficient causal mediation with intermediate confounders. Biometrika, 108(3), 627–641.

Fulcher, I. R., Shi, X., and Tchetgen Tchetgen, E.J. (2019). Estimation of Natural Indirect Effects Robust to Unmeasured Confounding and Mediator Measurement Error. Epidemiology, 30(6), 825-834.

Fulcher, I. R., Shpitser, I., Marealle, S., and Tchetgen Tchetgen, E. J. (2020). Robust inference on population indirect causal effects: the generalized front door criterion. Journal of the Royal Statistical Society. Series B, Statistical methodology, 82(1), 199–214.

Hejazi, N. S., van Der Laan, M. J., Janes, H. E., Gilbert P. B., and Benkeser D.C. (2020). Efficient nonparametric inference on the effects of stochastic interventions under two phase sampling, with applications to vaccine efficacy trials. Biometrics, 77, 1241–1253.

Hejazi, N. S., Rudolph, K. E., van Der Laan, M. J., and Díaz, I. (2022). Nonparametric causal mediation analysis for stochastic interventional (in)direct effects, Biostatistics, kxac002.

Karabatsos, G. and Walker, S. G. (2012). A bayesian nonparametric causal model. Journal of Statistical Planning and Inference, 142(4), 925–934.

Kennedy, E. H., Ma, Z., McHugh, M. D., and Small, D. S. (2017). Nonparametric methods for doubly robust estimation of continuous treatment effects. Journal of the Royal Statistical Society. Series B, Statistical methodology, 79(4), 1229–1245.

Kennedy, E. H. (2018). Semiparametric theory. Wiley StatsRef: Statistics Reference Online.

Kennedy, E. H. (2019). Nonparametric causal effects based on incremental propensity score interventions. Journal of the American Statistical Association, 114(526): 645-656.

Kennedy, E. H. (2022). Semiparametric doubly robust targeted double machine learning: a review. arXiv:2203.06469 [stat.ME].

Kosorok, M. R. (2008). Introduction to empirical processes and semiparametric inference. Springer.

Li, F., Ding, P., and Mealli, F. (2022). Bayesian causal inference: a critical review. arXiv:2206.15460 [stat.ME].

Linero, A. R. and Antonelli, J. L. (2022). The how and why of bayesian nonparametric causal inference. Wiley Interdisciplinary Reviews: Computational Statistics, page e1583.

Mallick, H. and Yi, N. (2014). A New Bayesian Lasso. Stat Interface, 7(4), 571-582.

Malinsky, D., Shpitser, I., and Tchetgen Tchetgen., E. J. (2021). Semiparametric Inference for Non-monotone Missing-Not-at-Random Data: the No Self-Censoring Model. Journal of the American Statistical Association, 0(0): 1-9.

Oganisian, A. and Roy, J. A. (2021). A practical introduction to bayesian estimation of causal effects: Parametric and nonparametric approaches. Statistics in Medicine, 40(2), 518– 551.

Hines, O., Dukes, O., Diaz-Ordaz,K., and Vansteelandt, S. (2022). Demystifying statistical learning based on efficient influence functions. The American Statistician, 76(3), 1–48.

Park, T. and Casella, G. (2008). The bayesian lasso. Journal of the American Statistical Association, 103(482), 681–686.

Pfanzagl, J. (1990). Estimation in semiparametric models. In Estimation in Semiparametric Models. Springer.

Rasmussen, C. E. (2003). Gaussian processes in machine learning. In Summer School on Machine Learning, pages 63–71. Springer.

Ray, K. and van der Vaart, A. (2020). Semiparametric bayesian causal inference. The Annals of Statistics, 48(5), 2999–3020.

Roy, J., Lum, K. J., Zeldow, B., Dworkin, J. D., Re III, V. L., and Daniels, M. J. (2018). Bayesian nonparametric generative models for causal inference with missing at random covariates. Biometrics, 74(4), 1193–1202.

Tchetgen Tchetgen, E. J. and Shpitser, I. (2012). Semiparametric theory for causal mediation analysis: efficiency bounds, multiple robustness, and sensitivity analysis. Annals of statistics, 40(3), 1816–1845.

Tsiatis, A., A. (2006). Semiparametric Theory and Missing Data. New York: Springer. van der Laan, M. J. and Rubin, D. (2006). Targeted Maximum Likelihood Learning. The International Journal of Biostatistics, 2(1).

van der Laan, M. J. and Rose, S. (2011). Targeted Learning. Springer Series in Statistics. Springer New York, New York, NY.

van der Laan, L., Zhang, W., and Gilbert P. B. (2022). Nonparametric estimation of the causal effect of a stochastic threshold-based intervention. Biometrics, doi: 10.1111/biom.13690.

van der Vaart, A. W. (2000). Asymptotic Statistics. Cambridge: Cambridge University Press. Vermeulen, K., and Vansteelandt, S. (2010).

Semiparametric Efficiency Proefschrift ingediend tot het behalen van de graad van Master in de Wiskunde, afstudeerrichting Toegepaste Wiskunde. https://libstore.ugent.be/fulltxt/RUG01/001/787/545/RUG01-001787545_2012_0001_AC.pdf

Xia, F. and Chan, K. C. G. (2021). Identification, Semiparametric Efficiency, and Quadruply Robust Estimation in Mediation Analysis with Treatment-Induced Confounding. Journal of the American Statistical Association, DOI: 10.1080/01621459.2021.1990765.