SUPPLEMENTARY MATERIAL FOR *DOUBLY ROBUST TREATMENT EFFECT ESTIMATION WITH INCOMPLETE ATTRIBUTES*

By Imke Mayer[¶], Julie Josse[**] Stefan Wager[††] and TraumaBase® Group (Tobias Gauss, Jean-Denis Moyer)[‡‡]

*École des Hautes Études en Sciences Sociales[¶], École Polytechnique[¶**], INRIA Saclay[**], Stanford University[††] and Beaujon Hospital[‡‡]*

1. Proofs.

1.1. *Consistency of AIPW.*

Proof. In order to show the double robustness of $\hat{\tau}_{AIPW}$, as given in (4), let us rewrite it by rearranging the terms:

$$\begin{aligned}\hat{\tau}_{AIPW} &= \frac{1}{n}\sum_{i=1}^{n}\frac{W_iY_i}{\hat{e}(X_i)} - \frac{W_i-\hat{e}(X_i)}{\hat{e}(X_i)}\hat{\mu}_1(X_i) - \frac{1}{n}\sum_{i=1}^{n}\frac{(1-W_i)Y_i}{1-\hat{e}(X_i)} + \frac{W_i-\hat{e}(X_i)}{1-\hat{e}(X_i)}\hat{\mu}_0(X_i)\\ &=: \hat{\mu}_{1,AIPW} - \hat{\mu}_{0,AIPW}.\end{aligned}$$

First note that by the law of large numbers, $\hat{\mu}_{1,AIPW}$ and $\hat{\mu}_{0,AIPW}$ respectively estimate $\mathbb{E}[Y_i(1)] + \eta_1$ and $\mathbb{E}[Y_i(0)] + \eta_0$ where η_1 and η_0 are given by

$$\eta_1 \triangleq \mathbb{E}\left[\frac{W_i - e(X_i)}{e(X_i)}(Y_i(1) - \mu_1(X_i))\right],\ \eta_0 \triangleq \mathbb{E}\left[\frac{W_i - e(X_i)}{1-e(X_i)}(Y_i(0) - \mu_0(X_i))\right].$$

Indeed we have that

$$\begin{aligned}\mathbb{E}\left[\frac{W_iY_i}{e(X_i)} - \frac{W_i-e(X_i)}{e(X_i)}\mu_1(X_i)\right] &= \mathbb{E}\left[\frac{W_iY_i(1)}{e(X_i)} - \frac{W_i-e(X_i)}{e(X_i)}\mu_1(X_i)\right]\\ &= \mathbb{E}[Y_i(1)] + \mathbb{E}\left[\frac{W_i-e(X_i)}{e(X_i)}(Y_i(1)-\mu_1(X_i))\right],\end{aligned}$$

[¶]E-mail: imke.mayer@ehess.fr
[**]E-mail: julie.josse@polytechnique.edu
[††]E-mail: swager@stanford.edu

where the first equality results from SUTVA: $W_iY_i = W_i(W_iY_i(1) + (1 - W_i)Y_i(0)) = W_iY_i(1) + W_i(1 - W_i)Y_i(0)$. And similar for the derivation of η_0.

Now, the double robustness can easily be shown by considering these two terms:

- If the propensity model $e(x)$ is correctly specified but the outcome model $(\mu_0(x), \mu_1(x))$ is mis-specified we have

$$\begin{aligned}\eta_1 &= \mathbb{E}\left[\mathbb{E}\left[\frac{W_i - e(X_i)}{e(X_i)}(Y_i(1) - \mu_1(X_i))\right] \mid Y_i(1), X_i\right] \\ &= \mathbb{E}\left[\frac{\mathbb{E}[W_i \mid Y_i(1), X_i] - e(X_i)}{e(X_i)}(Y_i(1) - \mu_1(X_i))\right] \\ &= \mathbb{E}\left[\frac{\mathbb{E}[W_i \mid X_i] - e(X_i)}{e(X_i)}(Y_i(1) - \mu_1(X_i))\right] = 0.\end{aligned}$$

  We use the unconfoundedness assumption to go from the second to the third line and the definition of the propensity score for the last equality.
- If the propensity model $e(x)$ is mis-specified but the outcome model $(\mu_0(x), \mu_1(x))$ is correctly specified we have

$$\begin{aligned}\eta_1 &= \mathbb{E}\left[\mathbb{E}\left[\frac{W_i - e(X_i)}{e(X_i)}(Y_i(1) - \mu_1(X_i))\right] \mid W_i, X_i\right] \\ &= \mathbb{E}\left[\mathbb{E}\left[\frac{W_i - e(X_i)}{e(X_i)}\left(Y_i(1) - \mathbb{E}[Y_i \mid W_i = 1, X_i]\right) \mid W_i, X_i\right]\right] \\ &= \mathbb{E}\left[\frac{W_i - e(X_i)}{e(X_i)}\left(\mathbb{E}[Y_i(1) \mid W_i, X_i] - \mathbb{E}[Y_i \mid W_i = 1, X_i]\right)\right] \\ &= \mathbb{E}\left[\frac{W_i - e(X_i)}{e(X_i)}\left(\mathbb{E}[Y_i(1) \mid X_i] - \mathbb{E}[Y_i(1) \mid X_i]\right)\right] = 0,\end{aligned}$$

  where we use SUTVA and unconfoundedness to go from the third to the fourth line.

Analogously we obtain in both cases of mis-specification that $\eta_0 = 0$, proving the double robustness of $\hat{\tau}_{AIPW}$. □

1.2. *Treatment Effect Estimation with Missing Attributes.* Here we prove the balancing property of the generalized propensity score (6).

PROOF. We note that the distribution of W is fully specified by its mean. Therefore we need to prove that:

$$\mathbb{E}[W_i|\{Y_i(0), Y_i(1)\}, X_i^*] = \mathbb{E}[W_i|X_i^*] \Rightarrow \mathbb{E}[W_i|\{Y_i(0), Y_i(1)\}, e^*(X_i^*)] = \mathbb{E}[W_i|e^*(X_i^*)]$$

a) By the law of total expectation we have:

$$\mathbb{E}[W_i \,|\, e^*(X_i^*)] = \mathbb{E}[\mathbb{E}[W_i|X_i^*, e^*(X_i^*)] \,|\, e^*(X_i^*)] = \mathbb{E}[\mathbb{E}[W_i|X_i^*] \,|\, e^*(X_i^*)] = e^*(X_i^*)$$

b) And again using the law of total expectation we have the following:

$$\begin{aligned}
&\mathbb{E}[W_i \,|\, \{Y_i(0), Y_i(1)\}, e^*(X_i^*)] \\
&\quad = \mathbb{E}[\mathbb{E}[W_i \,|\, \{Y_i(0), Y_i(1)\}, X_i^*, e^*(X_i^*)] \,|\, \{Y_i(0), Y_i(1)\}, e^*(X_i^*)] \\
&\quad = \mathbb{E}[\mathbb{E}[W_i|\{Y_i(0), Y_i(1)\}, X_i^*] \,|\, \{Y_i(0), Y_i(1)\}, e^*(X_i^*)] \\
&\quad = \mathbb{E}[\mathbb{E}[W_i \,|\, X_i^*] \,|\, \{Y_i(0), Y_i(1)\}, e^*(X_i^*)] \qquad \text{(assuming (7))} \\
&\quad = \mathbb{E}[e^*(X_i^*) \,|\, \{Y_i(0), Y_i(1)\}, e^*(X_i^*)] = e^*(X_i^*)
\end{aligned}$$

□

2. Simulation study on synthetic data.

2.1. *Simulation details.* In order to test generate data under the CIT/CIO assumption, we proceed as follows:

- **CIT**: $W \sim X \odot R$ (where $R_{ij} = \mathbb{1}_{\{X_{ij} \text{ is observed}\}}$ and $\odot$ = Hadamard product).
  Example: for fixed $\alpha \in \mathbb{R}^4$ and $\tau \in \mathbb{R}$:
  $r^i = (1,1,0,0,0,1,0,0,0,1) \Rightarrow \text{logit}(\mathbb{P}(W_i = 1|X_i^{obs} = x_i^{obs}, R_i = r_i)) = \alpha_0 + \alpha_1 x_{i1} + \alpha_2 x_{i2} + \alpha_6 x_{i6} + \alpha_{10} x_{i10}$

- **CIO**: $Y \sim X \odot R$.
  Example: for fixed $\beta \in \mathbb{R}^4$ and $\tau \in \mathbb{R}$:
  $r^i = (1,1,0,0,0,1,0,0,0,1) \Rightarrow \mathbb{E}(Y_i|X_i^{obs} = x_i^{obs}, R_i = r_i, W_i = w_i) = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_6 x_{i6} + \beta_{10} x_{i10} + \tau w_i$

- **¬CIT**: $\text{logit}(\mathbb{P}(W_i = 1|X_i = x_i)) = \alpha_0 + \alpha^T x_i$.

- **¬CIO**: $\mathbb{E}(Y_i|X_i = x_i, W_i = w_i) = \beta_0 + \beta^T x_i + \tau w_i$.

2.2. *Simulation results for non-parametric estimator under an hierarchical data-generating model.* The hierarchical data-generating model used in Section 5.2 can be modified in order to allow for correlation between covariates by defining the code-depending Gaussian parameters as

$$(\mu(c), \Sigma(c)) = (U(V \tanh(Wc + a) + b), U \exp(\gamma^T (Wc + a) + \delta) I_p U^T),$$

for some randomly generated orthonormal matrix U.

The difference in terms of bias and variability between the AIPW-type estimators and their IPW-type equivalent is clear in this scenario. However the difference in terms of bias w.r.t. the different unconfoundedness assumptions is less apparent. More precisely, *mia.grf* and *mean.grf* seem to approximate the true treatment effect τ for large sample sizes ($n \geq 500$) similarly in both scenarios (first and second line in Figure 1a and 1b). These observations require further investigations in the future.

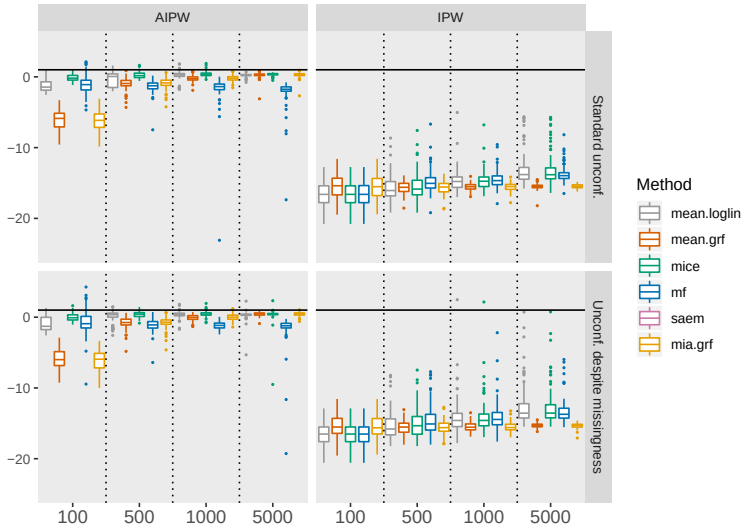
3. Simulation study on IHDP data. Similar conclusions drawn as for the AIPW-type estimators can be drawn for the inverse propensity weighted estimators as we see in Figure 2.

4. Details on the medical application (Traumabase).

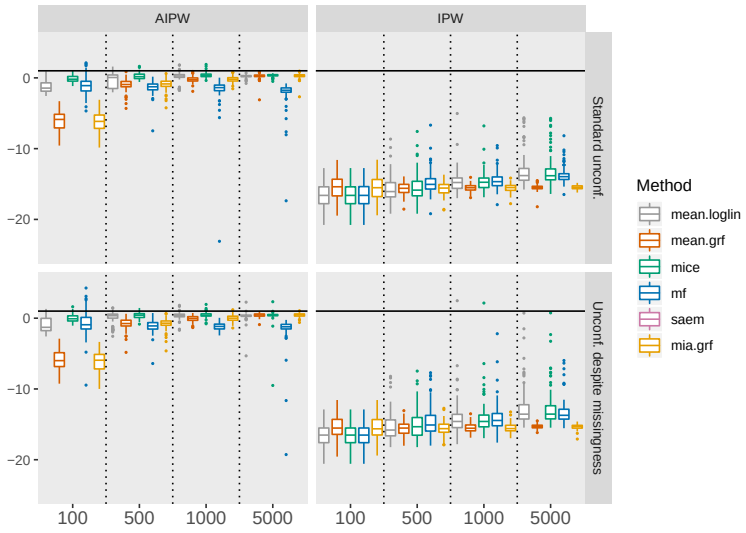
4.1. *Definition of the variables of the Traumabase used in the analysis.* Here we provide the names and short descriptions of the variables we use in our causal analysis. The moment at which the variable is first available is given in parentheses (ph = pre-hospital phase, h = hospital phase).

List of confounders:.

- *Trauma.center* (categorical): name of the trauma center. (ph/h)
- *SBP.ph*, *DBP.ph*, *HR.ph* (continuous): systolic and diastolic arterial pressure and heart rate during pre-hospital phase (SBP.ph = min(SBP.min, SBP.MICU), etc.); MICU = mobile intensive care unit. (ph)
- *Cardiac.arrest.ph* (categorical): cardiac arrest during pre-hospital phase. (ph)
- *HemoCue.init* (continuous): prehospital capillary hemoglobin concentration (the lower, the more the patient is probably bleeding and in shock); hemoglobin is an oxygen carrier molecule in the blood. (ph)
- *SpO2.min* (continuous): peripheral oxygen saturation, measured by pulse oxymetry, to estimate oxygen content in the blood ($95 - 100\%$: considered normal; $< 90\%$ critical and associated with considerable trauma, danger and mortality). (ph)

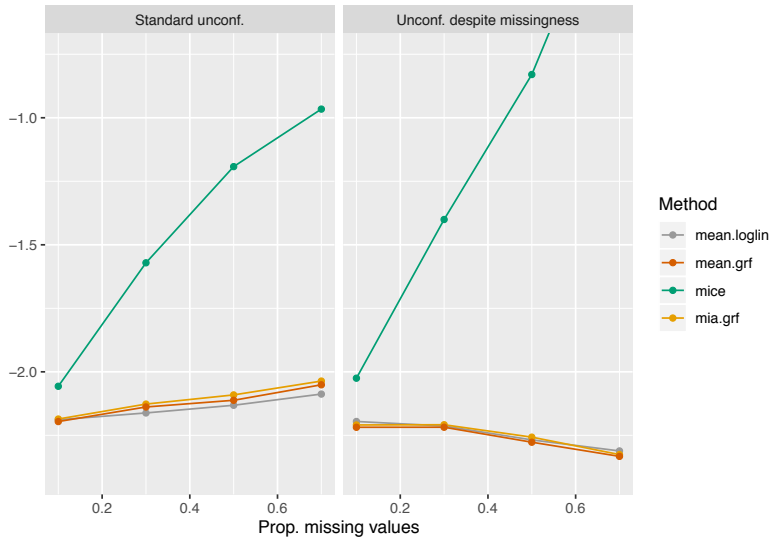


(a) MCAR (with 30% missing values in $X_{\cdot,1:10}$)

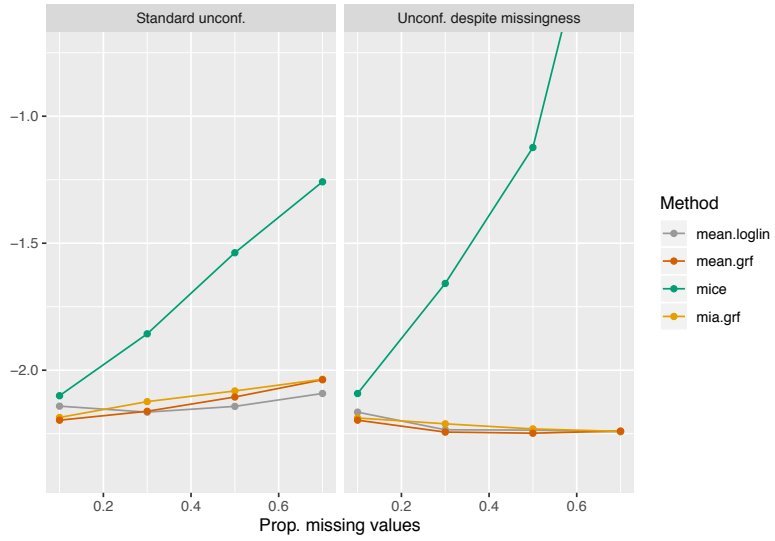


(b) MNAR (with 30% missing values in $X_{\cdot,1:5}$)

Fig 1: Estimated average treatment effect $\hat{\tau}$. **Hierarchical data-generating model with dense covariance matrices** for confounders.



(a) MCAR



(b) MNAR

Fig 2: RMSE of estimated average treatment effect $\hat{\tau}_{IPW}$ on IHDP dataset; (200 simulations for varying proportion of missing values; y-axis in log-scale).

- *Vasopressor.therapy* (continuous): treatment with catecholamines in

case of physical or emotional stress increasing heart rate, blood pressure, breathing rate, muscle strength and mental alertness. (ph)
- *Cristalloid.volume* (continuous): total amount of prehospital administered cristalloid fluid resuscitation (volume expansion). (ph)
- *Colloid.volume* (continuous): total amount of prehospital administered colloid fluid resuscitation (volume expansion). (ph)
- *Shock.index.ph* (continuous): ratio of heart rate and systolic arterial pressure during pre-hospital phase. (ph)
- *AIS.external* (discrete, range: $[0, 6]$): Abbreviated Injury Score for external injuries, here it is assumed to be a proxy of information available/visible during pre-hospital phase. (ph/h)
- *Delta.shock.index* (continuous): Difference of shock index between arrival at the hospital and arrival on the scene. (h)
- *Delta.hemoCue* (continuous): Difference of hemoglobin level between arrival at the hospital and arrival on the scene. (h)

List of predictors of mortality and that are not associated with treatment assignment.

- *Anticoagulant.therapy* (categorical): oral anticoagulant therapy before the accident. (ph)
- *Antiplatelet.therapy* (categorical): anti-platelet therapy before the accident. (ph)
- *GCS(.init)* (discrete, range: $[3, 15]$): Initial Glasgow Coma Scale (GCS) on arrival on scene of enhanced care team and on arrival at the hospital ($GCS = 3$: deep coma; $GCS = 15$: conscious and alert). (ph & h)
- *GCS.motor(.init)* (discrete, range: $[1, 6]$): Initial Glasgow Coma Scale motor score ($GCS.motor = 1$: no response; $GCS.motor = 6$: obeys command/purposeful movement). (ph & h)
- *Osmotherapy.ph* (categorical): pupil dilation indicating brain herniation. (ph & h)
- *Osmotherapy, Improv.anomaly.osmo* (categorical): administration of osmotherapy to alleviate compression of the brain (either Mannitol or hypertonic saline solution); change of pupil anomaly after administration of osmotherapy. (ph & h)
- *Medcare.time.ph* (continuous): total duration of prehospital care team engaged (arrival on scene to arrival at hospital). (h)
- *FiO2* (discrete, range: $[0, 5]$): inspired concentration of oxygen on ventilatory support (the higher the more critical; $Ventilation = 0$: no ventilatory support). (h)
- *Temperature.min* (continuous): Minimal body temperature. (h)

- *TCD.PI.max* (continuous): pulsatility index (PI) measured by echodoppler sonographic examen of blood velocity in cerebral arteries ($PI > 1.2$: indicates altered blood flow maybe due to traumatic brain injury). (h)
- *IICP* (categorical): at least one episode of increased intracranial pressure; mainly in traumatic brain injury; usually associated with worse prognosis. (h)
- *EVD* (categorical): external ventricular drainage (EVD); mean to drain cerebrospinal fluid to reduce intracranial pressure. (h)
- *Decompressive.craniectomy* (categorical): surgical intervention to reduce intracranial hypertension. (h)
- *Neurosurgery.day0* (categorical): neurosurgical intervention performed on day of admission. (h)
- *AIS.head, AIS.face* (discrete, range: $[0, 6]$): Abbreviated Injury Score, describing and quantifying facial and head injuries ($AIS = 0$: no injury; the higher the more critical).(h)
- *ISS* (discrete, range: $[0, 108]$): Injury Severity Score, sum of squares of top three AIS scores. (h)
- *IGS.II* (continuous): Simplified Acute Physiology Score. (h)

4.2. *ATE estimation on the Traumabase using overlap weights.* An often raised concern with many medical observational data sets is the potential violation of the overlap assumption. For instance some patients might never get the treatment due to infrastructural circumstances or due to recommendations followed strictly by the entire medical staff. The overlap assumption however is needed for consistency of the treatment effect estimations and states that every patient has a non-zero probability of being in either treated or control group. Another way of describing this assumption is that the treatment groups are sufficiently comparable, otherwise the attempt of drawing causal inferences is doomed to failure from the beginning.

Given the important level of heterogeneity among trauma patients, especially among patients with traumatic brain injury, and the multi-level and multi-actor nature of the data, it cannot be ruled out that the treatment groups have only small overlap. When considering standardized mean differences of the confounding variables between treatment and control groups in Figure 3 it appears indeed that certain features such as the hemoglobin level differ considerably between the two groups. As detailed in Section **??**, a possible solution to deal with this potential situation is the use of overlap weights instead of the inverse propensity weights (**?**). However, in our case, when using the corresponding modified estimands and estimators, i.e. the average treatment effect on the overlap population, the results reported in

Figure 4 are very similar to those from the normal average treatment effect estimation on the entire population (Figure 9) and lead to the same conclusion about the treatment effect.

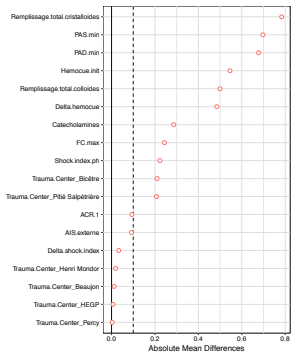


Fig 3: Unadjusted standardized mean differences of the confounding variables.

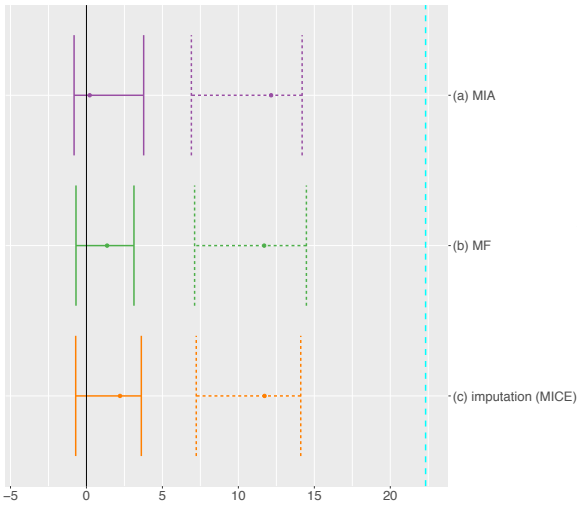


Fig 4: ATE estimations on overlap population on Traumabase data (solid: doubly robust estimates; dotted: IPW estimates; dashed vertical line: without adjustment; x-axis: $\hat{\tau}$ and bootstrap confidence intervals[1]).

[1] Values on the x-axis are multiplied by 100 for better readability.

I. Mayer
Centre d'Analyse et de Mathématique Sociales
EHESS
75006 Paris, France
E-mail: imke.mayer@ehess.fr

J. Josse
Centre de Mathématiques Appliquées
École Polytechnique
91128 Palaiseau Cedex, France
E-mail: julie.josse@polytechnique.edu

S. Wager
Graduate School of Business
Stanford University
CA 94305, USA
E-mail: swager@stanford.edu

T. Gauss and J.-D. Moyer
Department of anesthesia and intensive care
Beaujon hospital, AP–HP
92110 Clichy, France
E-mail: tgauss@protonmail.com
E-mail: jean-denis.moyer@aphp.fr