# Policy learning for personalized treatment recommendation

Talk PreMeDICaL
30/09/2024

Laura Fuentes Vicente

# Lab presentation

Montpellier Antenna,
INRIA Côte d'Azur

IDESP (Institut Desbrest of epidemiology and Public Health)

## PreMeDICaL (Precision Medicine by Data Integration and Causal Learning)

Julie JOSSE

Antoine CHAMBAZ

Paris Cité University

# Plan

# Plan

# I. Context

# Medical motivations

Given patient's characteristics, what is the **optimal treatment** to give to **maximize** each **patient's outcome**

→ Causal inference, policy learning

Example:

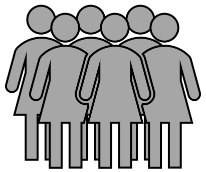
Find the **optimal hormone dose** to **maximize** the **number of oocyte** produced (under no-hyperstimulation constraint)

Gonadotrophin dose classes (treatment)

- Dose 1 … Dose K

# I.1-Mathematical framework

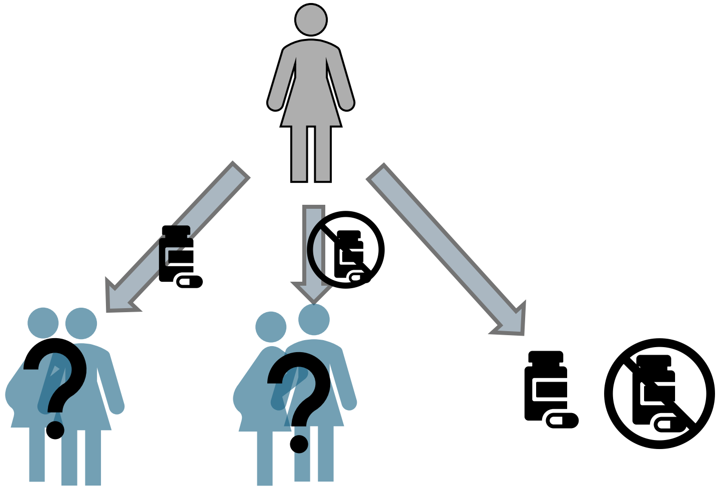
Set of independent and identically distributed subjects



- Covariates: $X_i \in \mathcal{X}$

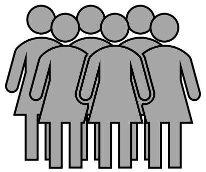
- Binary action: $W_i \in \mathcal{W} = \{0,1\}$

- Potential outcomes: $Y_i(w) \in \mathcal{Y},\ w \in \{0,1\}$
  $Y_i(0)$ *outcome in a world where* $w = 0$
  $Y_i(1)$ *outcome in a world where* $w = 1$

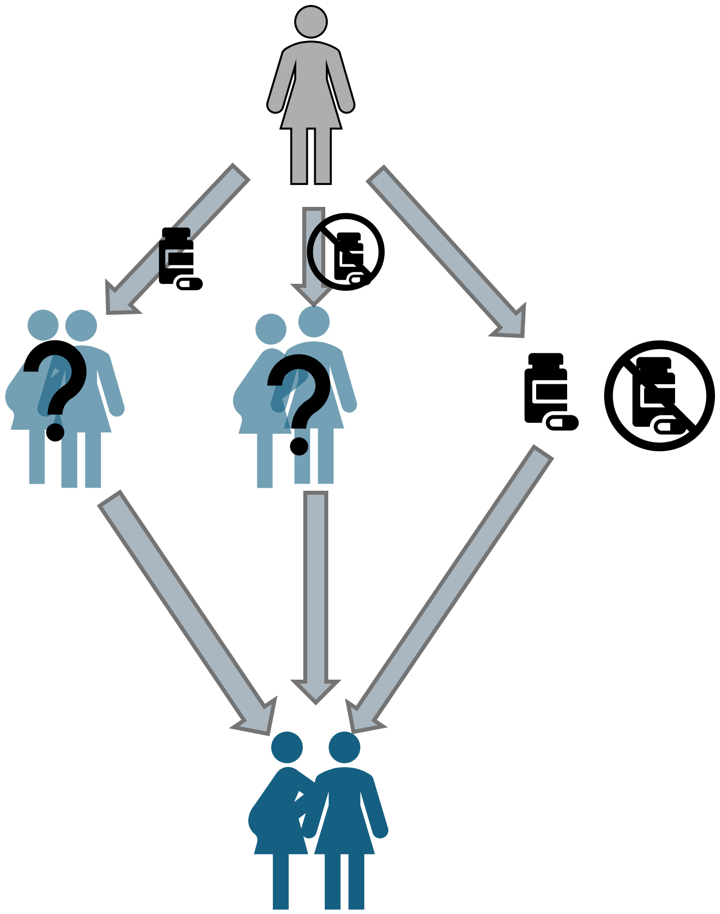


# I.1-Mathematical framework

Set of independent and identically distributed subjects

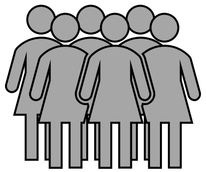


- Covariates: $X_i \in \mathcal{X}$
- Binary action: $W_i \in \mathcal{W} = \{0,1\}$
- Potential outcomes: $Y_i(w) \in \mathcal{Y},\ w \in \{0,1\}$
  $Y_i(0)$ *outcome in a world where* $w = 0$
  $Y_i(1)$ *outcome in a world where* $w = 1$
- Observed outcome: $Y_i = Y_i(W_i) \in \mathcal{Y}$

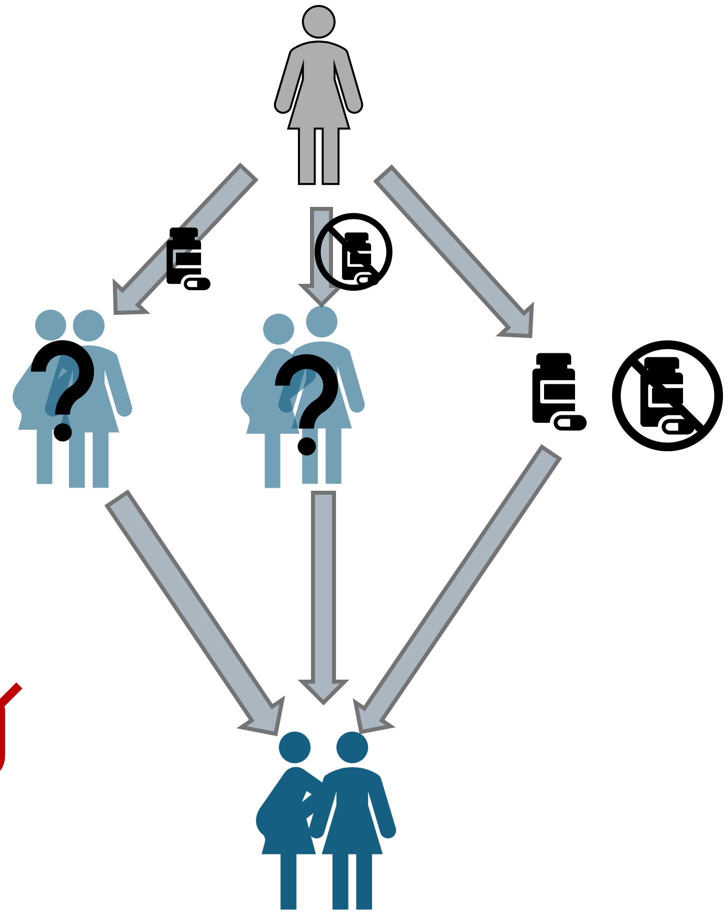


# I.1-Mathematical framework

Set of independent and identically distributed subjects



- Covariates: $X_i \in \mathcal{X}$
- Binary action: $W_i \in \mathcal{W} = \{0,1\}$
- Potential outcomes: $Y_i(w) \in \mathcal{Y},\ w \in \{0,1\}$
  $Y_i(0)$ *outcome in a world where* $w = 0$
  $Y_i(1)$ *outcome in a world where* $w = 1$
- Observed outcome: $Y_i = Y_i(W_i) \in \mathcal{Y}$
- Complete data-structure: $\mathbb{O}_i = (X_i, Y_i(1), Y_i(0), W_i, Y_i) \sim \mathbb{P}_0$
- Observation: $\mathcal{O}_i = (X_i, W_i, Y_i) \sim P_0$



# II. Methods

# II.1- Measures of causal effect

# II.1.1- Average treatment effect

# I.2.1-Average treatment effect

Represents the mean effect of treatment over a population

**Individual treatment effect:** $\Delta_i = Y_i(1) - Y_i(0)$

vs.

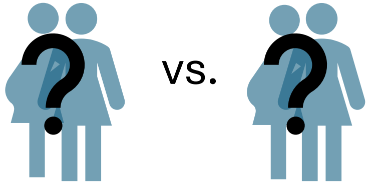
**Average treatment effect:**

$$\theta_{\mathbb{P}_0} = \mathbb{E}_{\mathbb{P}_0}[\Delta] = \mathbb{E}_{\mathbb{P}_0}[Y(1) - Y(0)]$$

# I.2.1-Average treatment effect

Represents the mean effect of treatment over a population

**Individual treatment effect:** $\Delta_i = Y_i(1) - Y_i(0)$




**Average treatment effect:**

$$\theta_{\mathbb{P}_0} = \mathbb{E}_{\mathbb{P}_0}[\Delta] = \mathbb{E}_{\mathbb{P}_0}[Y(1) - Y(0)]$$

Observational data: $O_i = (X_i, W_i, Y_i) \sim P_0$

| Covariates | | | Treatment | Outcome | Potential outcomes | |
|---|---|---|---|---|---|---|
| $X_1$ | $X_2$ | $X_3$ | W | Y | Y(0) | Y(1) |
| 1.1 | 20 | A | 1 | 200 | ? | 200 |
| -6 | 45 | B | 0 | 10 | 10 | ? |
| 0 | 15 | B | 1 | 150 | ? | 150 |
| | ... | | ... | ... | ... | ... |
| -2 | 52 | A | 0 | 100 | 100 | ? |

Assumptions:

1. SUTVA: $Y_i = Y_i(W_i)$
2. Overlap: $\eta < P_0(W = 1|X) < 1 - \eta$, for $\eta$>0
3. Unconfoundedness: $Y(w) \perp W|X,\ w \in \{0,1\}$

**Average treatment effect estimation:**

$$\theta_{\mathbb{P}_0} = \mathbb{E}_{\mathbb{P}_0}[Y(1) - Y(0)] = \mathbb{E}_{\mathbb{P}_0}[Y(1)] - \mathbb{E}_{\mathbb{P}_0}[Y(0)]$$

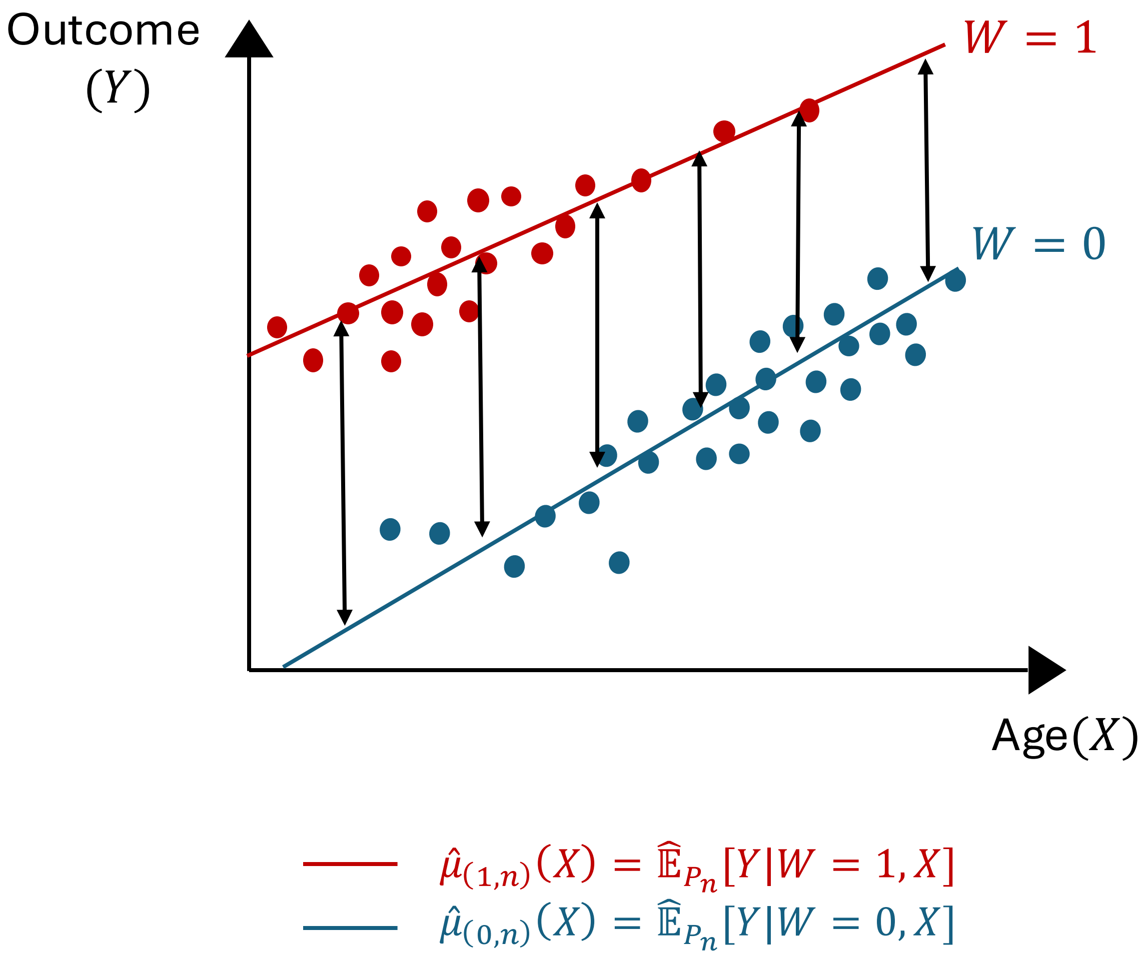
$$\triangleq \mathbb{E}_{P_0}[Y|W = 1] - \mathbb{E}_{P_0}[Y|W = 0] = \mathbb{E}_{P_0}[\,\mathbb{E}_{P_0}[Y|W = 1, X] - \mathbb{E}_{P_0}[Y|W = 0, X]\,]$$

# I.2.1-Average treatment effect estimators

**Average treatment effect:**

$$\psi_{G-comp,n} = \mathbb{E}_{P_n}\left[\hat{\mu}_{(1,n)}(X) - \hat{\mu}_{(0,n)}(X)\right] = \frac{1}{n}\sum_{i=1}^{n} \hat{\mu}_{(1,n)}(X_i) - \hat{\mu}_{(0,n)}(X_i)$$

$$\psi_{IPW,n} = \mathbb{E}_{P_n}\left[\frac{(2W-1)Y}{\widehat{P}_n(W=w|X=x)}\right] = \frac{1}{n}\sum_{i=1}^{n} \frac{2W_i - 1}{\widehat{P}_n(W=W_i|X=X_i)} Y_i$$

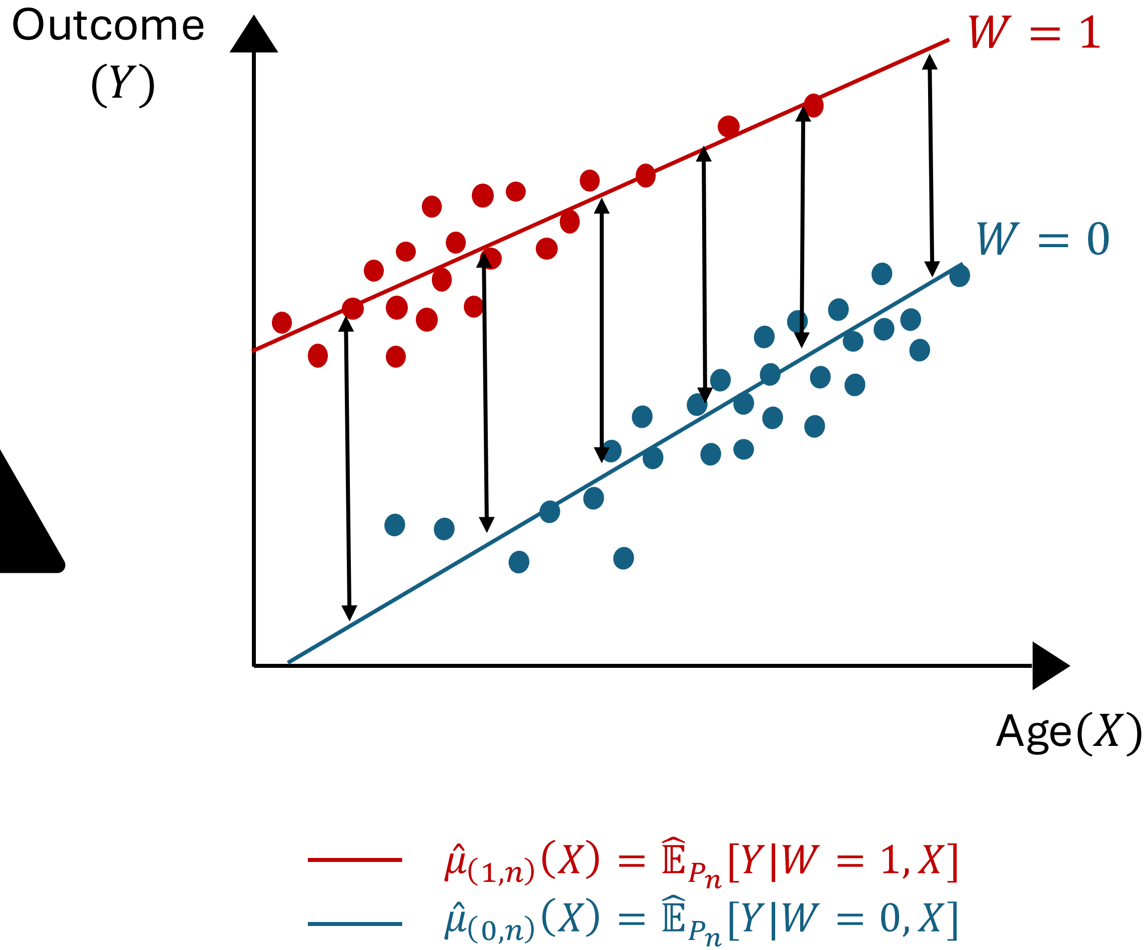


# I.2.1-Average treatment effect estimators

**Average Treatment effect:**

$$\psi_{G-comp,n} = \mathbb{E}_{P_n}\left[\hat{\mu}_{(1,n)}(X) - \hat{\mu}_{(0,n)}(X)\right] = \frac{1}{n}\sum_{i=1}^{n} \hat{\mu}_{(1,n)}(X_i) - \hat{\mu}_{(0,n)}(X_i)$$

$$\psi_{IPW,n} = \mathbb{E}_{P_n}\left[\frac{(2W-1)Y}{\hat{P}_n(W=w|X=x)}\right] = \frac{1}{n}\sum_{i=1}^{n} \frac{2W_i - 1}{\hat{P}_n(W=W_i|X=X_i)} Y_i$$



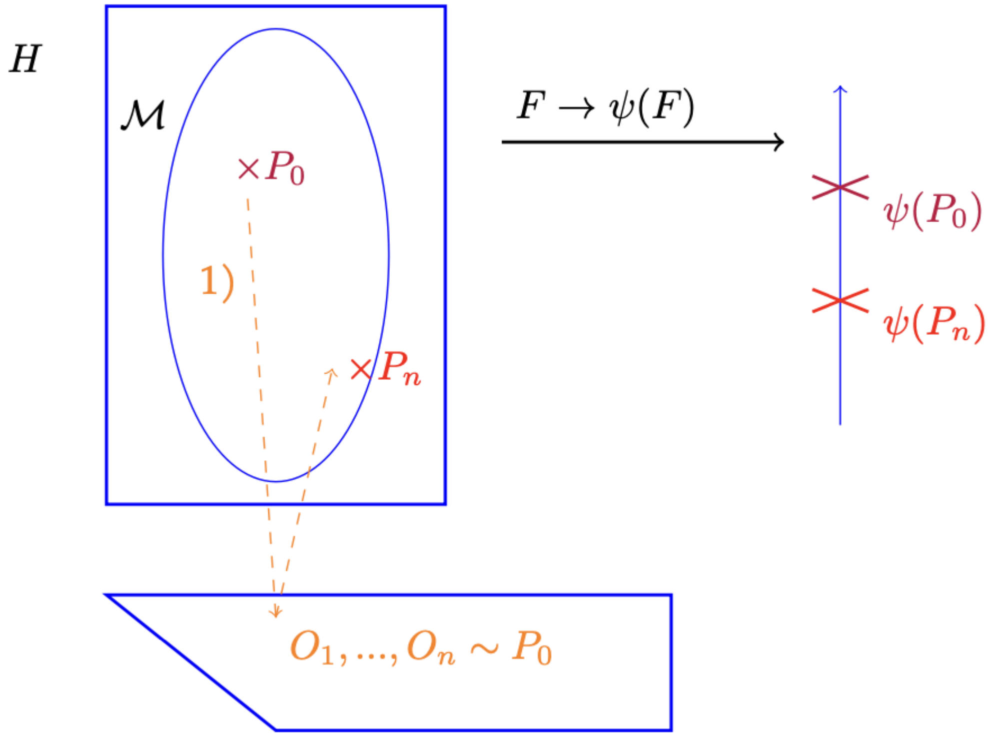

Double robust estimators

→ Augmented IPW / One-step correction estimator

→ Targeted MLE

# II.1.1-ATE double robust estimators

Normal framework:




**Representation of the relationship between:**

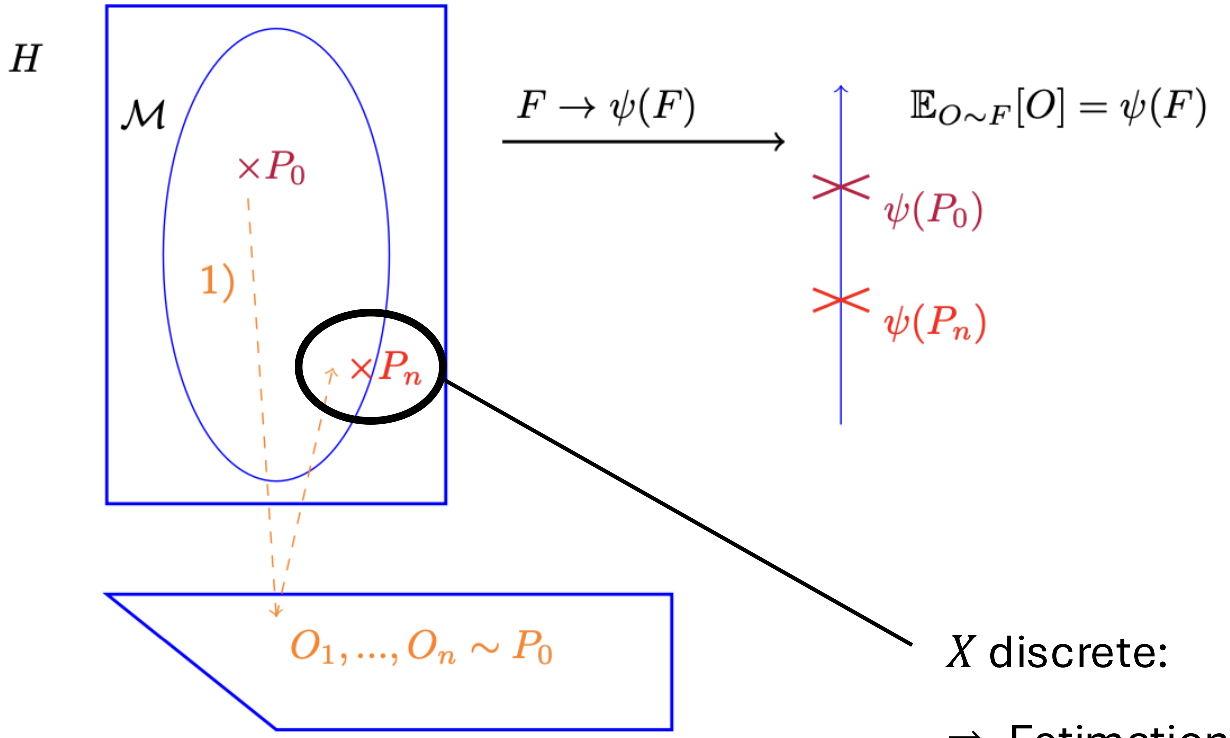
- Observed data
- the empirical distribution ($P_n$)
- the parameter of interest $\psi(P_0)$

where:

- $\mathcal{M}$: set of distributions st. $\psi(P_0)$ well defined
- $\psi(P_0) = \mathbb{E}_{P_0}[\,\mathbb{E}_{P_0}[Y|W=1,X] - \mathbb{E}_{P_0}[Y|W=0,X]\,]$
- $\psi(P_n) = \mathbb{E}_{P_n}[\,\hat{\mathbb{E}}_{P_n}[Y|W=1,X] - \hat{\mathbb{E}}_{P_n}[Y|W=0,X]\,]$

# II.1.1- ATE double robust estimators

Normal framework:




**Representation of the relationship between:**

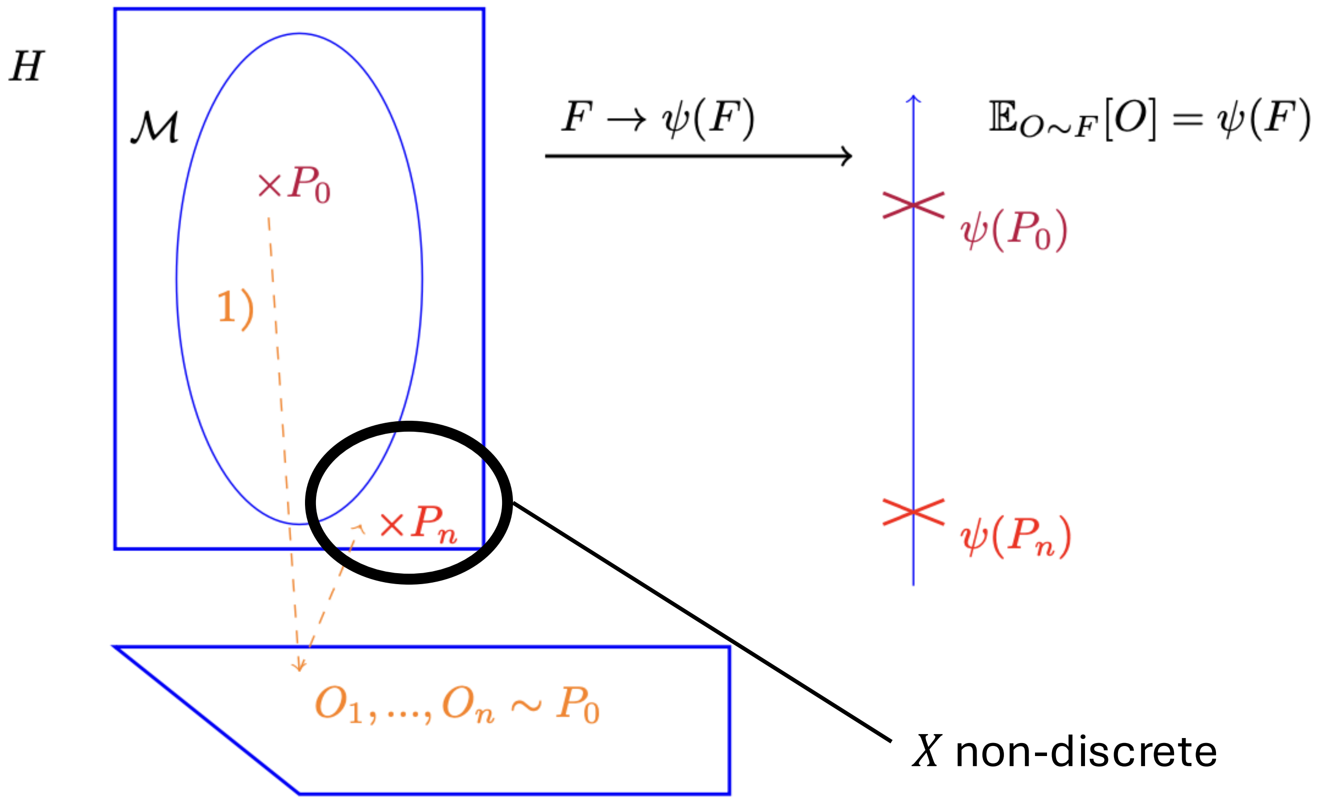
- Observed data
- the empirical distribution ($P_n$)
- the parameter of interest $\psi(P_0)$

$X$ discrete:

$\Rightarrow$ Estimation of $\widehat{\mathbb{E}}_{P_n}[Y|W,X]$ : direct ✓

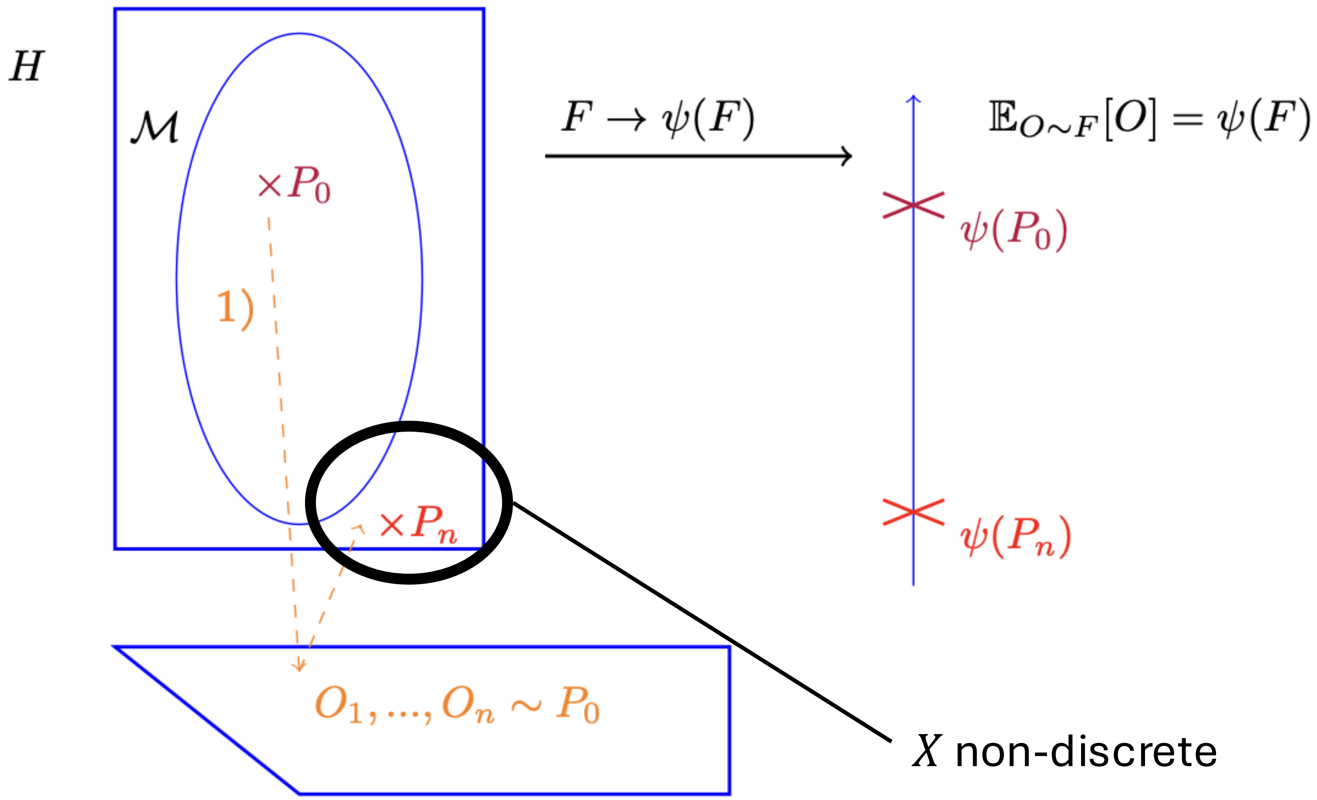
# II.1.1- ATE double robust estimators

Case: $P_n \notin \mathcal{M}$



# II.1.1- ATE double robust estimators

Case: $P_n \notin \mathcal{M}$




**Consequences**:

→ Restrinctive assumptions on model class

or

→ Using algorithms to estimate: $\widehat{\mathbb{E}}_{P_n}[Y|W,X]$

→ slower convergence rate

→ bias

# II.1.1- ATE double robust estimators

$$\psi(P) = \psi(P_0) - \mathbb{E}_{P_0}[\varphi(O;P)] + Rem_{P_0}(P)$$

Influence function (IF)

$o_p(\frac{1}{\sqrt{n}})$

# II.1.1- ATE double robust estimators

$$\psi(P) = \psi(P_0) - \mathbb{E}_{P_0}[\varphi(O;P)] + Rem_{P_0}(P)$$

Influence function (IF) $\qquad o_p(\frac{1}{\sqrt{n}})$

**Assumptions:**

1- $\varphi(O;P) \in \mathcal{L}_0^2(P) = \{\varphi(O;P): \mathbb{E}_P[\varphi(O;P)] = 0 \ \& \mathbb{E}_P[\varphi(O;P)^2] < \infty\}$

2- $\exists P_\infty \in \mathcal{M}$ such that $\|\varphi(O;P) - \varphi(O;P_\infty)\|_{2,p} \xrightarrow[n\to\infty]{} 0$

$$\psi(P) - \psi(P_0) = a + b - c + o_p(\frac{1}{\sqrt{n}})$$

$a)$ $\mathbb{E}_{P_n}[\varphi(O;P_\infty)] - \mathbb{E}_{P_0}[\varphi(O;P_\infty)] \Rightarrow \sqrt{n}(\mathbb{E}_{P_n}[\varphi(O;P_\infty)] - \mathbb{E}_{P_0}[\varphi(O;P_\infty)]) \to \mathcal{N}(0, Var_P(O;P_\infty)(O))$

$b)$ $(\mathbb{E}_{P_n}[\varphi(O;P)] - \mathbb{E}_{P_n}[\varphi(O;P_\infty)]) - (\mathbb{E}_{P_0}[\varphi(O;P)] - \mathbb{E}_{P_0}[\varphi(O;P_\infty)]) = o_p(\frac{1}{\sqrt{n}})$

$c)$ $\mathbb{E}_{P_n}[\varphi(O;P)]$: random term!

# II.1.1- ATE double robust estimators

**Augmented IPW**

$$\psi(P) = \psi(P_0) - \mathbb{E}_{P_0}[\varphi(O;P)] + Rem_{P_0}(P) \qquad \Rightarrow \psi(P) - \psi(P_0) = a + b - c + o_p(\frac{1}{\sqrt{n}})$$

**Solution:**

$$\psi_{AIPW}(P) = \psi(P) + c = \psi(P) + \mathbb{E}_{P_n}[\varphi(O;P)] = \psi(P) + \frac{1}{n}\sum_{i=1}^{n} \varphi(O_i;P)$$

**Objective: compute $\varphi(O;P)$!**

# II.1.1- ATE double robust estimators

**Augmented IPW**

$$\psi(P) = \psi(P_0) - \mathbb{E}_{P_0}[\varphi(O;P)] + Rem_{P_0}(P) \qquad \Rightarrow \psi(P) - \psi(P_0) = a + b - c + o_p(\frac{1}{\sqrt{n}})$$

**Solution:**

$$\psi_{AIPW}(P) = \psi(P) + c = \psi(P) + \mathbb{E}_{P_n}[\varphi(O;P)] = \psi(P) + \frac{1}{n}\sum_{i=1}^{n}\varphi(O_i;P)$$

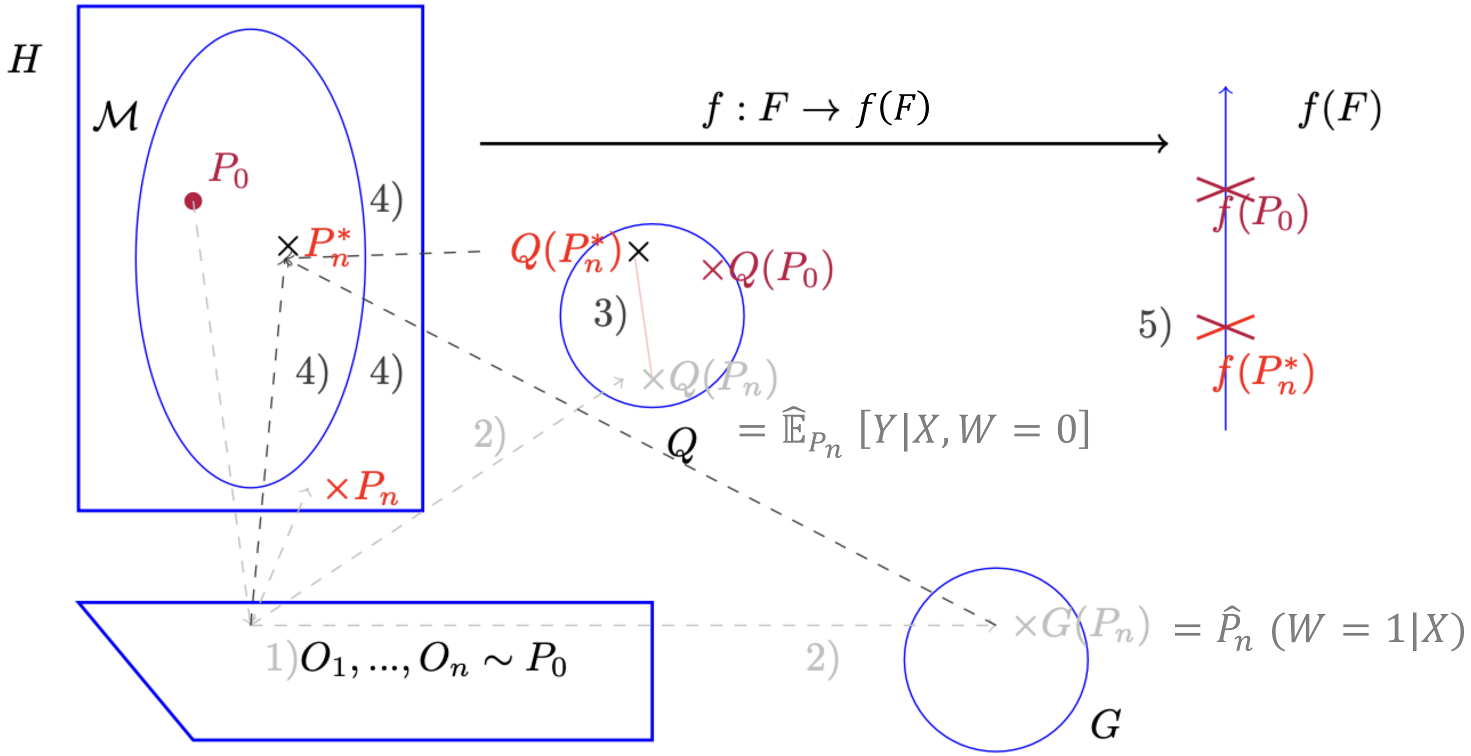
**Objective: compute $\boldsymbol{\varphi(O;P)}$!**

$$\psi_{AIPW}(P) = \mathbb{E}_P[\hat{\mathbb{E}}_P[Y|X,W=1] - \hat{\mathbb{E}}_P[Y|X,W=0]$$
$$+ \frac{1_{W=1}}{P(W=1|X=x)}(Y - \hat{\mathbb{E}}_P[Y|X,W=1]) - \frac{1-1_{W=1}}{1-P(W=1|X=x)}(Y - \hat{\mathbb{E}}_P[Y|X,W=0])]$$

# II.1.1- ATE double robust estimators

**Targeted Maximum Likelihood Estimator (TMLE)**

Performs the bias correction in the regression space $Q$




- **Build initial estimators**
  - Regression: $Q(P) \in Q$
  - Propensity score: $G(P) \in G$

- **Build our fluctuation:**

  Correct initial regression $Q(P_n)$, s.t. :

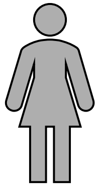
  $c = \mathbb{E}_{P_n}[\varphi(O; P)] = 0$

- **Estimate $\psi(P_0)$ with corrected regression!**

# II.1.2- Conditional average treatment effect

# I.2.2-Conditional average treatment effect

Expected difference in outcome between receiving and not receiving treatment within a specific population defined by covariates $X = x$

Example:  $X = x$

**Conditional average treatment effect:**

$$CATE_{\mathbb{P}_0}(x) = \mathbb{E}_{\mathbb{P}_0}[\,\Delta\,|\mathrm{X} = x] = \mathbb{E}_{\mathbb{P}_0}[Y(1) - Y(0)|X = x]$$

Same assumptions

**Conditional average treatment effect estimation:**

$$CATE_{\mathbb{P}_0}(x) = \tau_{P_0}(x), \quad \tau_P\colon \mathcal{X} \to \mathbb{R}, \quad \forall P \in \mathcal{M}$$
$$= \mathbb{E}_{P_0}[Y|W = 1, X = x] - \mathbb{E}_{P_0}[Y|W = 0, X = x]$$

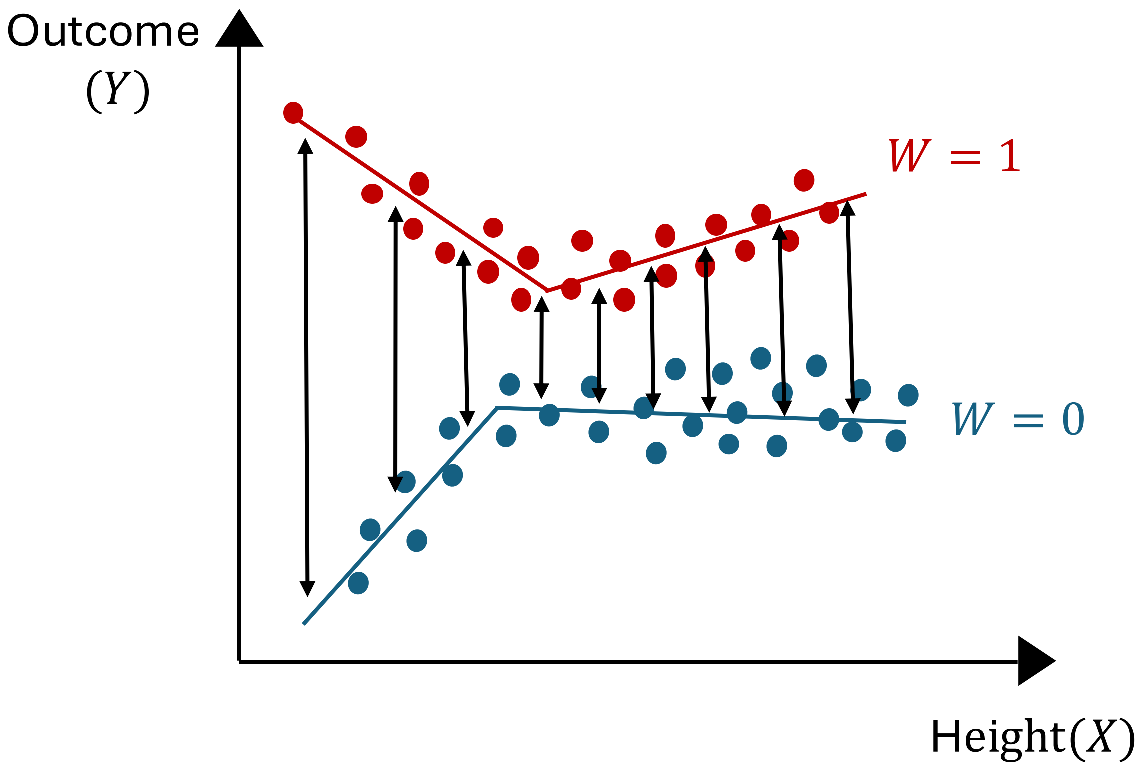
# I.2.2-Conditional treatment effect estimators

$\hat{\mu}_{(1,n)}(x) = \widehat{\mathbb{E}}_{P_n}[Y|W = 1, X = x]$
$\hat{\mu}_{(0,n)}(x) = \widehat{\mathbb{E}}_{P_n}[Y|W = 0, X = x]$

**Conditional average treatment effect:**

$$\tau_{G-comp,n}(x) = \hat{\mu}_{(1,n)}(x) - \hat{\mu}_{(0,n)}(x)$$

*i.e. X-learner, R-learner, DR-learner, MACF, etc.*



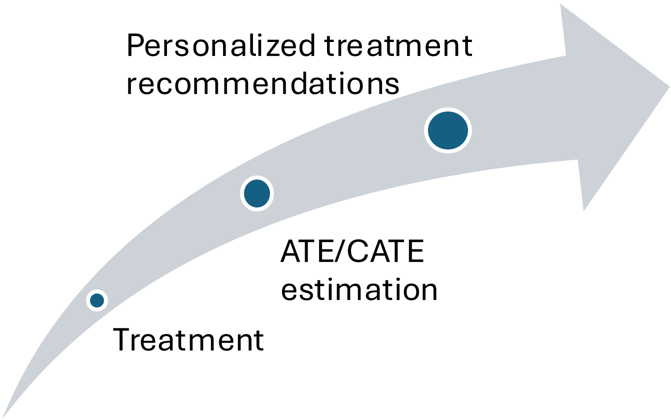

<u>Double robust estimators</u>

→ Augmented IPW / One-step correction estimator

→ Targeted MLE

# II.2- Policy learning

# II.2-Policy learning framework




<u>Mathematical framework:</u>

Let’s consider a decision maker:

- Patient characteristics: $X_i \in \mathcal{X}$ (covariates)
- Actions: $W_i \in \mathcal{W}$ (here action = treatment)
- Observed outcome: $Y_i \in \mathcal{Y}$ (for chosen action)

- **Policy:** $d \in \mathcal{D}$ decision maker’s support
  $d\colon \mathcal{X} \to \mathcal{W}$

# II.2-Policy learning framework: policy value

The value of a policy reflects the mean outcome expected following the given policy $(\boldsymbol{d})$

$$V_d(\mathbb{P}_0) = \mathbb{E}_{\mathbb{P}_0}[Y(d(X))] = \mathbb{E}_{\mathbb{P}_0}[d(X)Y(1) + (1 - d(X))Y(0)]$$

$$\triangleq \mathbb{E}_{P_0}\left[\mathbb{E}_{P_0}[Y|X, W = d(X)]\right] = V_d(P_0)$$

- Assess performance of a policy
- Compare policies
- ...

**<u>Two possible goals:</u>**

1. **<u>Optimization:</u>** Find the best treatment policy (maximizing the total expected value)

$$d^* \in \underset{d \in \mathcal{D}}{\operatorname{argmax}} V_d(P_0)$$

2. **Evaluation:** Estimating the expected value of a given policy:

$$V_d(\mathbb{P}_0) \triangleq V_d(P_0) \rightsquigarrow \hat{V}_d(P_n)$$

# II.2.1- Policy optimization

# II.2.1- Outcome modeling approaches

**1- Outcome modeling-based methods:**

Model Y(1) and Y(0)

→ Estimate $CATE_{\mathbb{P}_0}(x)$ :

$\triangleq \tau_{P_0}(x) = \mathbb{E}_{P_0}[Y|W=1, X=x] - \mathbb{E}_{P_0}[Y|W=0, X=x]$

$$d^*(X) = 1_{sign(\tau_{P_0}(X))>0}$$

| Covariates | | | Treatment | Estimated potential outcomes | | Treatment rule |
|---|---|---|---|---|---|---|
| $X_1$ | $X_2$ | $X_3$ | $W$ | $\hat{\mu}_0(X)$ | $\hat{\mu}_1(X)$ | $d(X)$ |
| 1.1 | 20 | F | 1 | 100 | 200 | 1 |
| -6 | 45 | F | 0 | 10 | 9 | 0 |
| 0 | 15 | M | 1 | 180 | 150 | 0 |
| | ... | | ... | ... | | ... |
| -2 | 52 | M | 0 | 70 | 170 | 1 |

$$\hat{\tau}_n(x) = \hat{\mu}_{1,n}(x) - \hat{\mu}_{0,n}(x)$$

T/S-learner    IPW/AIPW    TMLE    MACF

$\hat{\mu}_{(1,n)}(X) = \hat{\mathbb{E}}_{P_n}[Y|W=1, X]$

$\hat{\mu}_{(0,n)}(X) = \hat{\mathbb{E}}_{P_n}[Y|W=0, X]$

$\hat{\mu}_{(n)}(X, W) = \hat{\mathbb{E}}_{P_n}[Y|W, X]$

# II.2.1- Outcome modeling approaches

**1- Outcome modeling-based methods:**

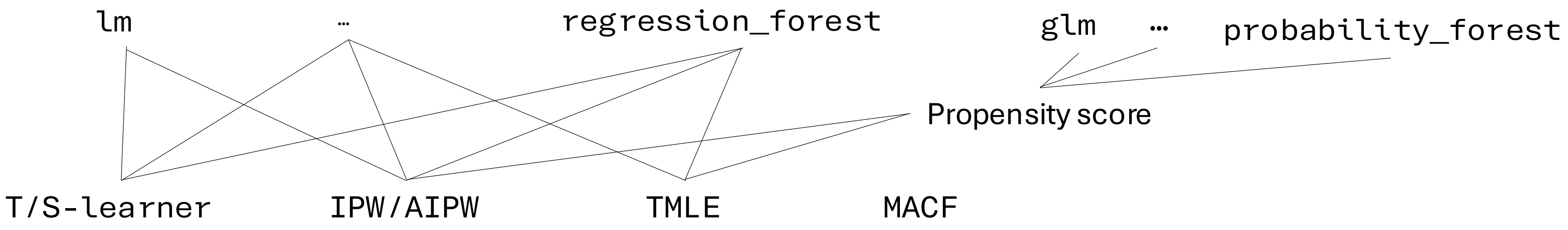
Model Y(1) and Y(0)

→ Estimate $CATE_{\mathbb{P}_0}(x)$ :

$$\triangleq \tau_{P_0}(x) = \mathbb{E}_{P_0}[Y|W=1, X=x] - \mathbb{E}_{P_0}[Y|W=0, X=x]$$

$$d^*(X) = 1_{sign\left(\tau_{P_0}(X)\right)>0}$$

| Covariates | | | Treatment | Estimated potential outcomes | | Treatment rule |
|---|---|---|---|---|---|---|
| $X_1$ | $X_2$ | $X_3$ | $W$ | $\hat{\mu}_0(X)$ | $\hat{\mu}_1(X)$ | $d(X)$ |
| 1.1 | 20 | F | 1 | 100 | 200 | 1 |
| -6 | 45 | F | 0 | 10 | 9 | 0 |
| 0 | 15 | M | 1 | 180 | 150 | 0 |
| … | | | … | … | | … |
| -2 | 52 | M | 0 | 70 | 170 | 1 |



# II.2.1- Outcome modeling approaches

**1- Outcome modeling-based methods:**

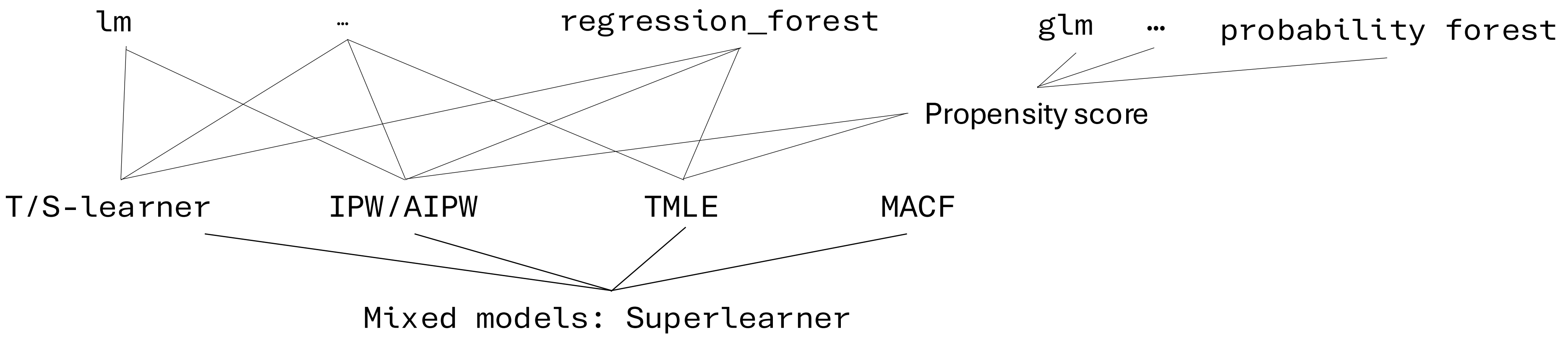
Model Y(1) and Y(0)

→ Estimate $CATE_{\mathbb{P}_0}(x)$ :

$$\triangleq \tau_{P_0}(x) = \mathbb{E}_{P_0}[Y|W=1, X=x] - \mathbb{E}_{P_0}[Y|W=0, X=x]$$

$$d^*(X) = 1_{sign\left(\tau_{P_0}(X)\right)>0}$$

| Covariates | | | Treatment | Estimated potential outcomes | | Treatment rule |
|---|---|---|---|---|---|---|
| $X_1$ | $X_2$ | $X_3$ | $W$ | $\hat{\mu}_0(X)$ | $\hat{\mu}_1(X)$ | $d(X)$ |
| 1.1 | 20 | F | 1 | 100 | 200 | 1 |
| -6 | 45 | F | 0 | 10 | 9 | 0 |
| 0 | 15 | M | 1 | 180 | 150 | 0 |
| | … | | … | … | | … |
| -2 | 52 | M | 0 | 70 | 170 | 1 |



# II.2.2- Direct estimation techniques

**2- Direct estimation techniques:**

$$d_0^* \in \underset{d \in \mathcal{D}}{\operatorname{argmax}} V_d(P_0) = \mathbb{E}_{P_0}\left[\mathbb{E}_{P_0}[Y|X, d(X)]\right]$$

(1)

vs.

(0)

$\Longrightarrow$ Classification task

# II.2.2- Direct estimation techniques

**2- Direct estimation techniques:**

1. **Single stage outcome weighted learning**

$$d_0^* \in \underset{d \in \mathcal{D}}{\operatorname{argmax}} V_d(P_0),\ V_d(\mathbb{P}_0) = \mathbb{E}_{\mathbb{P}_0}[\mathbb{E}_{\mathbb{P}_0}[Y(d(X))]] = \mathbb{E}_{\mathbb{P}_0}[\mathbb{E}_{\mathbb{P}_0}[\frac{1_{W=d(X)}Y(d(X))}{\mathbb{P}_0(W|X)}|X]] \triangleq \mathbb{E}_{P_0}\left[\frac{Y}{P_0(W|X)}1_{W=d(X)}\right]$$

$$= \underset{d \in \mathcal{D}}{\operatorname{argmin}} \mathbb{E}_{P_0}\left[\frac{Y}{P_0(W|X)}1_{W \neq d(X)}\right] \quad \Rightarrow \underset{f \in \mathcal{F}}{\operatorname{argmin}}\ \mathbb{E}_{P_0}[\alpha_i \underbrace{1_{(2W-1)f(X)}}_{\Phi(1-(2W-1)f(X))} + \lambda Pen(d(X))]$$

Find $f_0^*$ whose sign defines the OTR: $d_0^* = \frac{sign(f_0^*)+1}{2}$

loss: Hinge, logistic, etc.
penalization: Lasso, Ridge, ElasticNet, None
weight: IPW ($\frac{Y}{P_0(W=w|X)}$), AIPW ($\frac{Y-\hat{\mu}_W(X)}{P_0(W=w|X)}$)

# II.2.2- Direct estimation techniques

**2- Direct estimation techniques:**

1. **Single stage outcome weighted learning**

$$d_0^* \in \underset{d\in\mathcal{D}}{\text{argmax}}\, V_d(P_0),\ V_d(\mathbb{P}_0) = \mathbb{E}_{\mathbb{P}_0}[\mathbb{E}_{\mathbb{P}_0}[Y(d(X))]] = \mathbb{E}_{\mathbb{P}_0}\left[\mathbb{E}_{\mathbb{P}_0}\left[\frac{1_{W=d(X)}Y(d(X))}{\mathbb{P}_0(W|X)}|X\right]\right] \triangleq \mathbb{E}_{P_0}\left[\frac{Y}{P_0(W|X)}1_{W=d(X)}\right]$$

$$= \underset{d\in\mathcal{D}}{\text{argmin}}\,\mathbb{E}_{P_0}\left[\frac{Y}{P_0(W|X)}1_{W\neq d(X)}\right] \qquad \Rightarrow \underset{f\in\mathcal{F}}{\text{argmin}}\ \mathbb{E}_{P_0}\left[\alpha_i\, \underbrace{1_{(2W-1)f(X)}}_{\Phi(1-(2W-1)f(X))} + \lambda Pen(d(X))\right]$$

Find $f_0^*$ whose sign defines the OTR:

$$d_0^* = \frac{sign(f_0^*)+1}{2}$$

**Estimate $f_0^*(X)$ with $\mathrm{P_n}$:**

$$f_n^* \in \underset{d\in\mathcal{D}}{\text{argmin}} \frac{1}{n}\sum_{i=1}^{n} \frac{Y_i}{\hat{P}_n(W_i|X=X_i)}\Phi(1-(2W_i-1)f(X_i)) + \lambda Pen(d(X_i))$$

$$d_n^* = \frac{sign(f_n^*)+1}{2}$$

loss: Hinge, logistic, etc.
penalization: Lasso, Ridge, ElasticNet, None
weight: IPW ($\frac{Y}{P_0(W=w|X)}$), AIPW ($\frac{Y-\hat{\mu}_W(X)}{P_0(W=w|X)}$)

# II.2.2- Direct estimation techniques

**2- Direct estimation techniques:**

1. Single stage outcome weighted learning
2. **Weighted classification**

$$V_d(\mathbb{P}_0) = \mathbb{E}_{\mathbb{P}_0}[Y(d(X))] = \mathbb{E}_{\mathbb{P}_0}\big[d(X)Y(1) + \big(1 - d(X)\big)Y(0)\big]$$

$$= \underbrace{\mathbb{E}_{\mathbb{P}_0}[Y(0)]}_{\text{baseline effect}} + \underbrace{\mathbb{E}_{\mathbb{P}_0}\big[\big(Y(1) - Y(0)\big)d(X)\big]}_{\text{ATE dependence}}$$

# II.2.2- Direct estimation techniques

**2- Direct estimation techniques:**

$$V_d(\mathbb{P}_0) = \underbrace{\mathbb{E}_{\mathbb{P}_0}[Y(0)]}_{\text{baseline effect}} + \underbrace{\mathbb{E}_{\mathbb{P}_0}[\big(Y(1) - Y(0)\big)d(X)]}_{\text{ATE dependence}}$$

1. Single stage outcome weighted learning

2. **Weighted classification**

$$\Longrightarrow A(d) = 2(V_d(\mathbb{P}_0) - \mathbb{E}_{\mathbb{P}_0}[Y(0)]) - \mathbb{E}_{\mathbb{P}_0}[Y(1) - Y(0)]$$

$$\triangleq \mathbb{E}_{P_0}[\boldsymbol{\tau(X)}\,(2d(X) - 1)] = \mathbb{E}_{P_0}[\underbrace{|\tau(X)|}_{\alpha}\underbrace{sign(\tau(X))(2d(X) - 1)}_{\in\{-1,1\}}]$$

$$d_0^* \in \underset{d}{\operatorname{argmax}}\, A(d), \qquad A(d) = \mathbb{E}_{P_0}[\alpha\, sign(\tau(X))(2d(X) - 1)]$$

Options for $\alpha$ :

IPW: $\alpha^{IPW} = |\frac{Y}{P_0(W = w|X)}|$

AIPW: $\alpha^{AIPW} = |\frac{Y - \hat{\mu}_W(X)}{P_0(W=w|X)}|$

$$d_n^* \in \underset{d \in \mathcal{D}}{\operatorname{argmax}} \frac{1}{n} \sum_{i=1}^{n} |\alpha_i| sign(\tau(X_i))(2d(X_i) - 1)$$

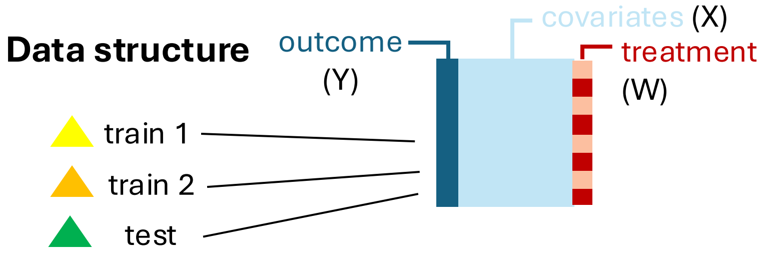
# II.2.2- Policy evaluation

# II.3- Policy evaluation

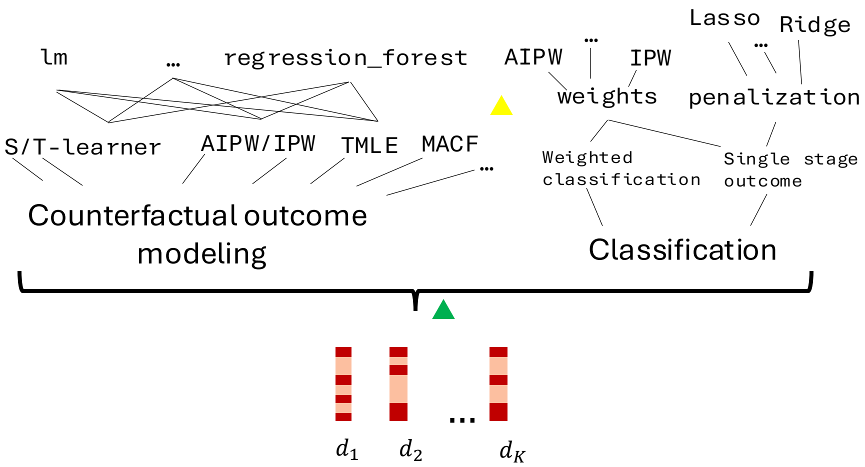
The value of a policy reflects the mean outcome expected following the given policy $(d)$

**Objective:** Compute the value of each policy to compare their performances

$$V_d(P) = \mathbb{E}_P[\mathbb{E}_P[Y|W = d(X), X]]$$




**Step 1: Gather policies $(d)$ to evaluate**

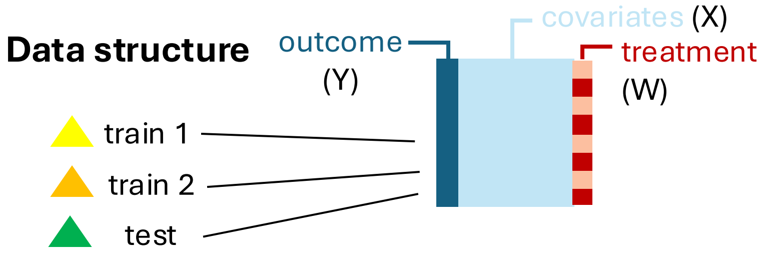


# II.3- Policy evaluation

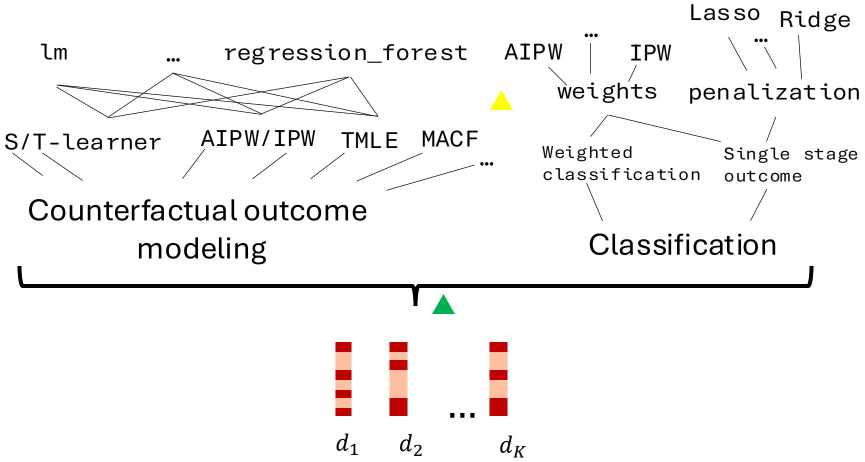
The value of a policy reflects the mean outcome expected following the given policy $(\boldsymbol{d})$

**Objective:** Compute the value of each policy to compare their performances

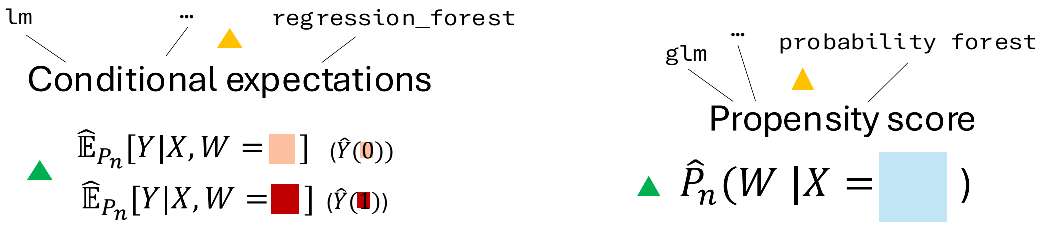
$$V_d(P) = \mathbb{E}_P[\mathbb{E}_P[Y|W = d(X), X]]$$




## Step 1: Gather policies $(\boldsymbol{d})$ to evaluate




## Step 2: Train nuisance parameters



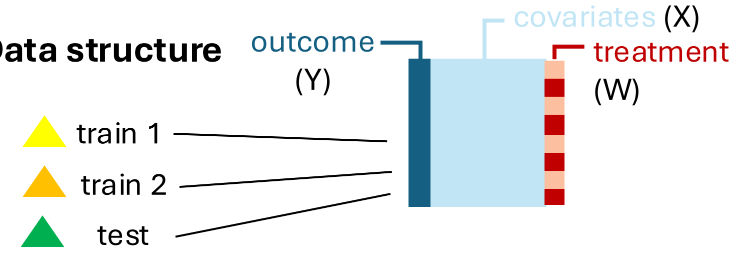
# II.3- Policy evaluation

The value of a policy reflects the mean outcome expected following the given policy $(d)$

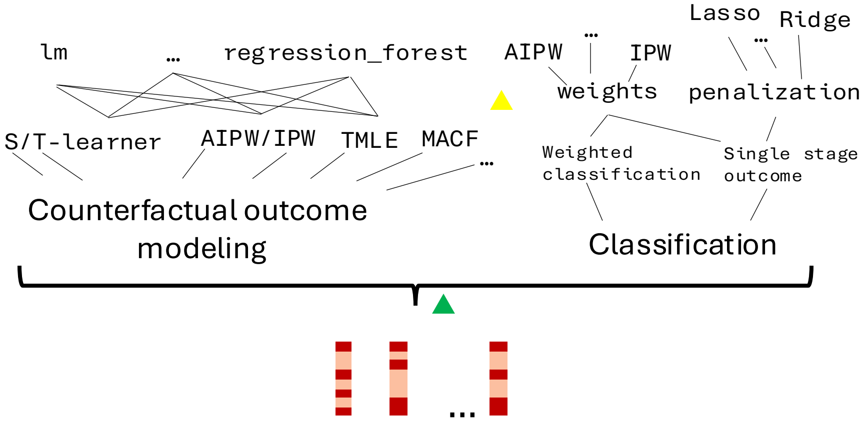
**Objective:** Compute the value of each policy to compare their performances

$$V_d(P) = \mathbb{E}_P[\mathbb{E}_P[Y|W = d(X), X]]$$

**Data structure**

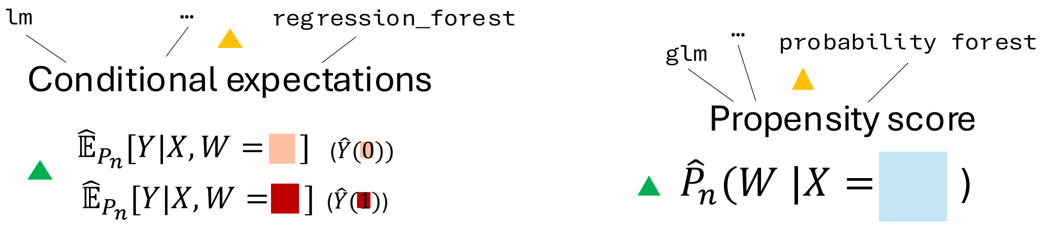



## Step 1: Gather policies $(d)$ to evaluate




$d_1$ $d_2$ ... $d_K$

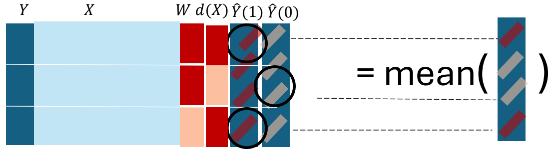
## Step 2: Train nuisance parameters




## Step 3: Compute policy value

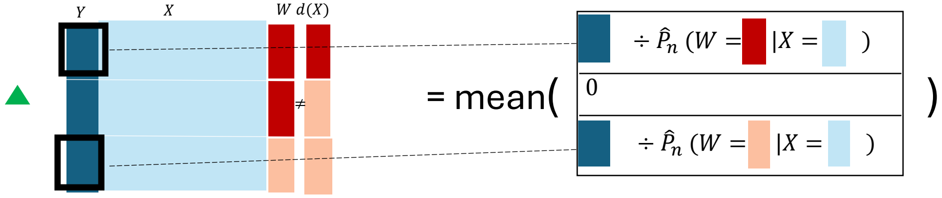
Substitution estimator

$$V_{subs.est,d}(P_n) = \frac{1}{n}\sum_{i=1}^{n} \widehat{\mathbb{E}}_{P_n}[\; |X = \;, W = d(X)\;]$$




IPW estimator

$$V_{IPW,d}(P_n) = \frac{1}{n}\sum_{i=1}^{n} \frac{1_{W = d(X)}}{\widehat{P}_n(W = \; |X = \;)}$$




And other double robust estimators: AIPW, TMLE

# III. Results

# III-Synthetic setting

$$X \sim \text{Unif}(0,1)^p$$

$$W \sim \text{Ber}(p = 0.5)$$

$$Y = u(X) + Wc(X) + Z$$

intercept function

$u\colon \mathbb{R}^p \mapsto \mathbb{R}$

Gaussian noise:

$Z \sim \mathcal{N}(0,1)$

Optimal treatment rule function

$$c(X) = c_0(2d^*(X) - 1)$$

$$d^*\colon \mathbb{R}^p \mapsto \{0,1\}$$

$:X \mapsto d^*(X) = 1_{f(X)<b}$ or $1_{f(X)>b}$

$n = 1000$ (individuals), $p = 5$ (covariates)

# III-Synthetic setting

$$X \sim \text{Unif}(0,1)^p$$

$$W \sim \text{Ber}(p = 0.5)$$

$$Y = u(X) + Wc(X) + Z$$

intercept function
$u\colon \mathbb{R}^p \mapsto \mathbb{R}$

Gaussian noise:
$Z \sim \mathcal{N}(0,1)$

Optimal treatment rule function

$$c(X) = c_0(2d^*(X) - 1)$$

$$d^*\colon \mathbb{R}^p \mapsto \{0,1\}$$

$$: X \mapsto d^*(X) = 1_{f(X)<b} \text{ or } 1_{f(X)>b}$$

**Tree setting:**

$$u(X) = k + \sqrt{\frac{5}{\sum_{i=1}^{p} X_p} \times \sum_{i=i}^{p} X_p}$$

$$k = 10 - \frac{1}{n}\sum_{i=1}^{p} X_p |c(X)| - \sqrt{\frac{5}{\sum_{i=1}^{p} X_p} \times \sum_{i=i}^{p} X_p}$$

$c_0 = \sqrt{5}$ $\quad b_1 = 0.6$ $\quad b_2 = 0.2$

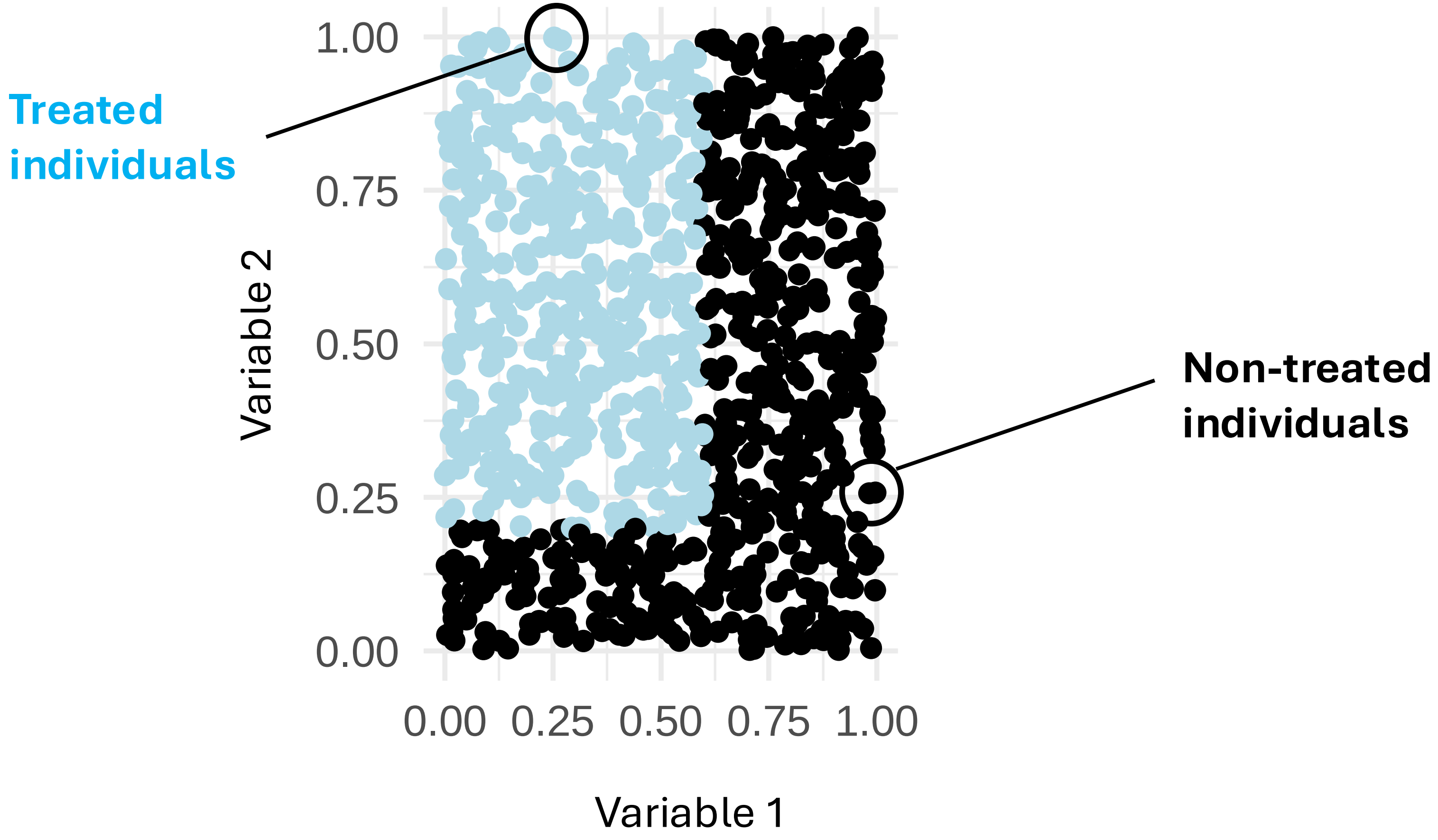
$f_1(X) = X_1$ $\quad f_2(X) = X_2$

$$d^*(X) = 1_{f_1(X)<b_1 \,\wedge\, f_2(X)\geq b_2}$$

$n = 1000$ (individuals), $p = 5$ (covariates)

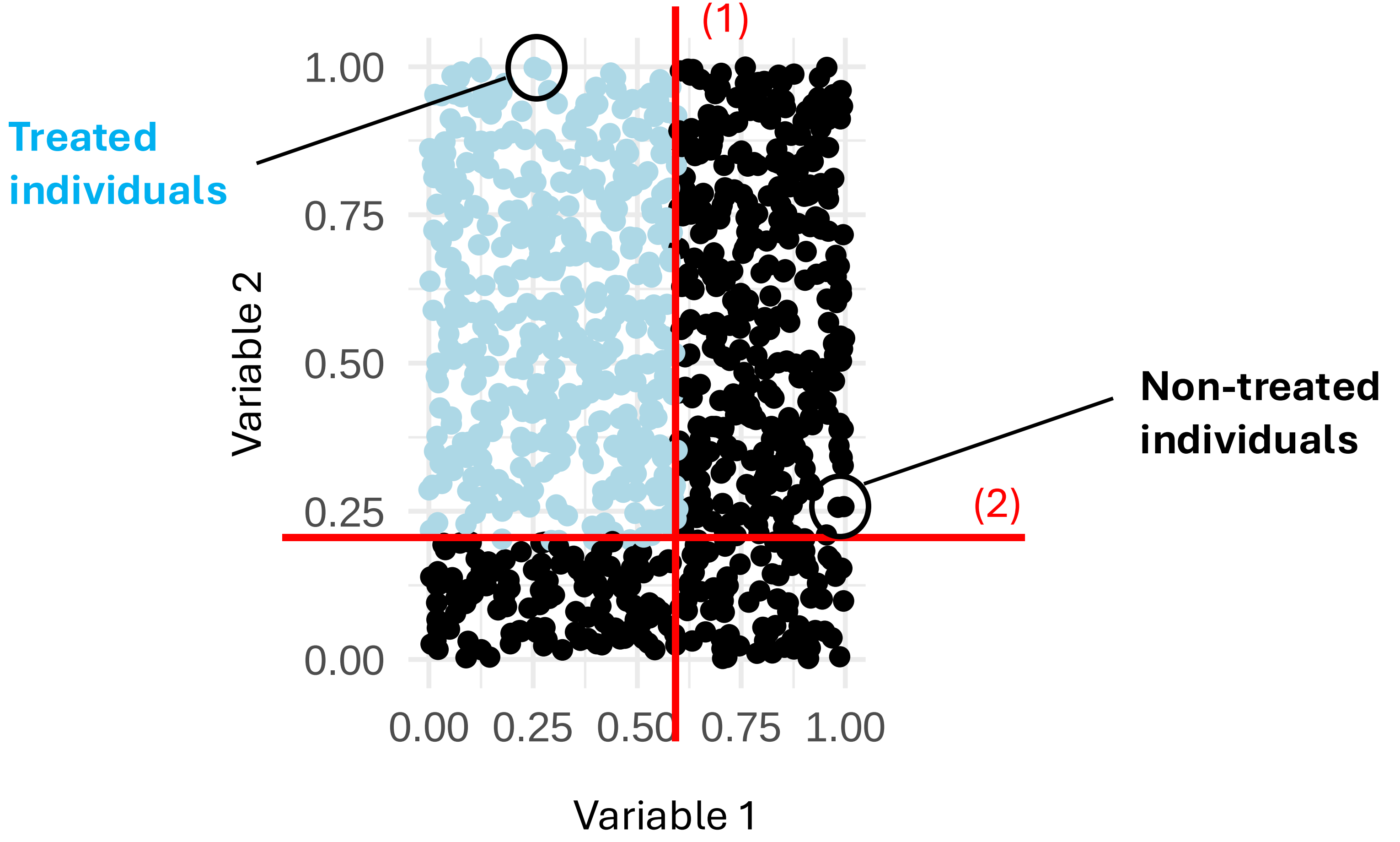
# III-Tree setting

Optimal treatment rule (d_Opt )



# III-Tree setting

## Optimal treatment rule (d_Opt )



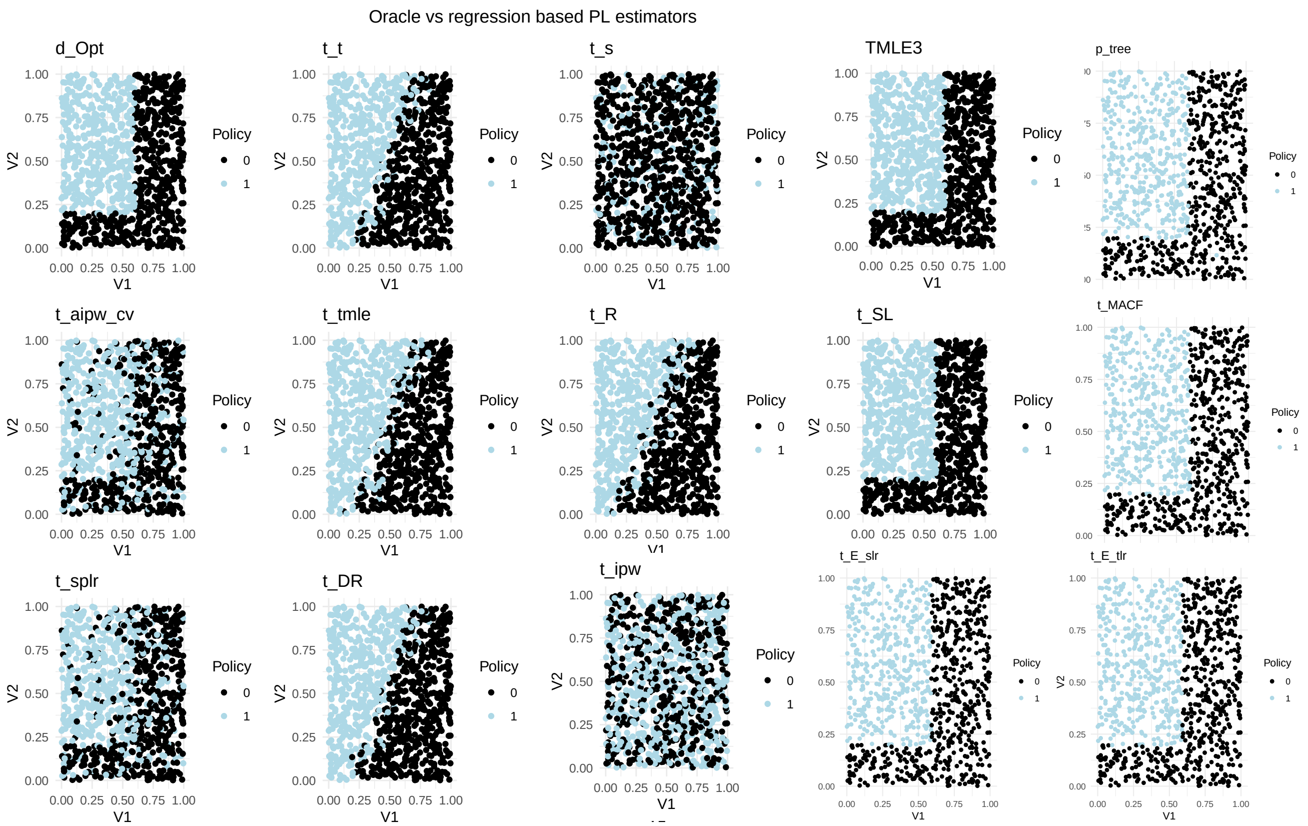

**Treats:**

Variable 1 < 0.6 (1)

**and**

Variable 2 > 0.2 (2)

# III-Tree setting




**Outcome modeling-based approach:**
Estimate CATE:

$$\hat{\tau}_n(x) = \hat{\mu}_{1,n}(x) - \hat{\mu}_{0,n}(x)$$

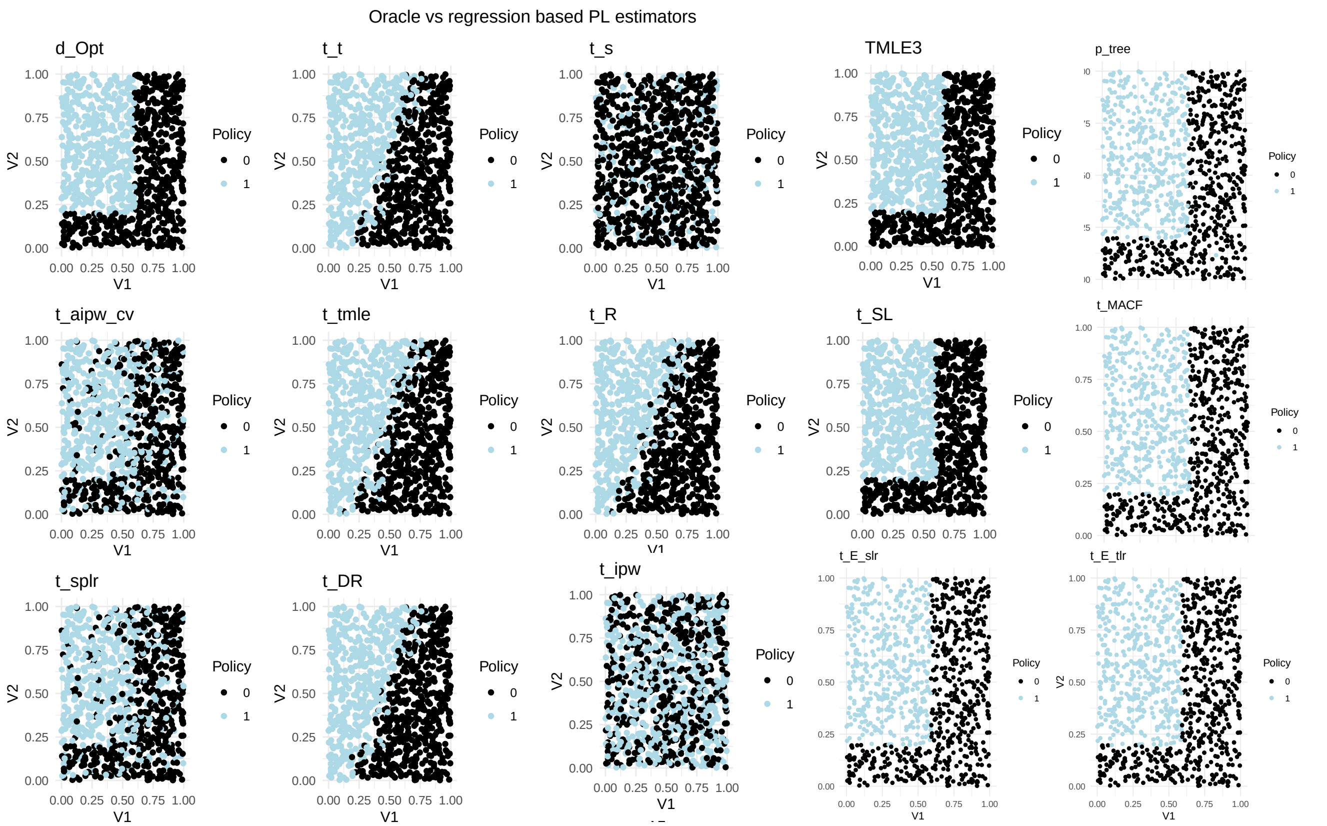
$$\hat{d}_n(x) = 1_{sign(\hat{\tau}_n(x)>0)}$$

**Figure 1: Policy optimization results for regression and tree-based algorithms in a tree setting**
**Left**: Visual representation of treatment rules



$$\hat{\mu}_{w,n}(x) = \widehat{\mathbb{E}}_{P_n}[Y|X = x, W = w]$$

# III-Tree setting




$\hat{\mu}_{w,n}(x)$: lm

$\hat{\mu}_{w,n}(x)$: tree-based techniques

**Outcome modeling-based approach:**
Estimate CATE:

$$\hat{\tau}_n(x) = \hat{\mu}_{1,n}(x) - \hat{\mu}_{0,n}(x)$$

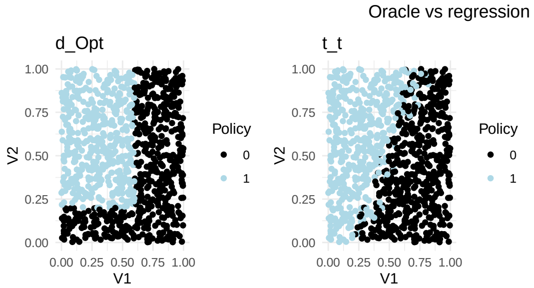
$$\hat{d}_n(x) = 1_{sign(\hat{\tau}_n(x)>0)}$$

**Figure 1: Policy optimization results for regression and tree-based algorithms in a tree setting**
**Left**: Visual representation of treatment rules



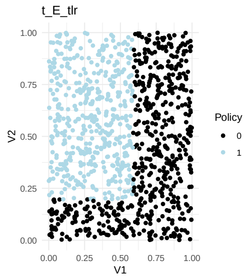
$\hat{\mu}_{w,n}(x) = \hat{\mathbb{E}}_{P_n}[Y|X = x, W = w]$

# III-Tree setting




$$\hat{\tau}_{T-learner\,(lm),\,n}(x) = \hat{\mu}_{1,n}(x) - \hat{\mu}_{0,n}(x)$$

$$\hat{d}_{T-learner\,(lm),\,n}\,(x) = 1_{sign(\hat{\tau}_{T-learner\,(lm),\,n}(x)>0)}$$




$$\hat{\tau}_{T-learner\,(grf),\,n}(x) = \hat{\mu}_{1,n}(x) - \hat{\mu}_{0,n}(x)$$

$$\hat{d}_{T-learner\,(grf),\,n}\,(x) = 1_{sign(\hat{\tau}_{T-learner\,(grf),\,n}(x)>0)}$$

$\hat{\mu}_{w,n}(x)$: lm

$\hat{\mu}_{w,n}(x)$: tree-based techniques

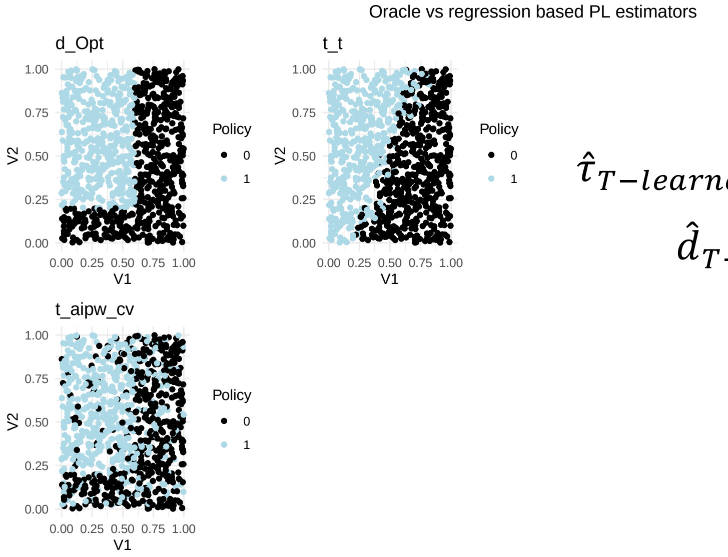
**Figure 1: Policy optimization results for regression and tree-based algorithms in a tree setting**

**Left**: Visual representation of treatment rules



$$\hat{\mu}_{w,n}(x) = \hat{\mathbb{E}}_{P_n}[Y|X = x, W = w]$$

# III-Tree setting



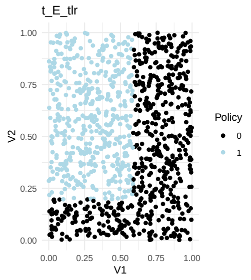

$$\hat{\tau}_{T-learner\,(lm),\,n}(x) = \hat{\mu}_{1,n}(x) - \hat{\mu}_{0,n}(x)$$

$$\hat{d}_{T-learner\,(lm),\,n}\,(x) = 1_{sign(\hat{\tau}_{T-learner\,(lm),\,n}(x)>0)}$$

$$\hat{\tau}_{AIPW-CV\,(lm),\,n}(x) = \hat{\mu}_{1,n}(x) - \hat{\mu}_{0,n}(x) + \frac{1_{W=1}}{\hat{P}(W=1|X=x)}(\mathrm{Y}-\hat{\mu}_{1,n}(x)) - \frac{1-1_{W=1}}{1-\hat{P}(W=1|X=x)}(\mathrm{Y}-\hat{\mu}_{0,n}(x))$$

$$\hat{d}_{AIPW-CV(lm),\,n}(x) = 1_{sign(\hat{\tau}_{AIPW-CV\,(lm),\,n}(x)>0)}$$

$\hat{\mu}_{w,n}(x)$: lm




$\hat{\mu}_{w,n}(x)$: tree-based techniques

$$\hat{\tau}_{T-learner\,(grf),\,n}(x) = \hat{\mu}_{1,n}(x) - \hat{\mu}_{0,n}(x)$$

$$\hat{d}_{T-learner\,(grf),\,n}\,(x) = 1_{sign(\hat{\tau}_{T-learner\,(grf),\,n}(x)>0)}$$

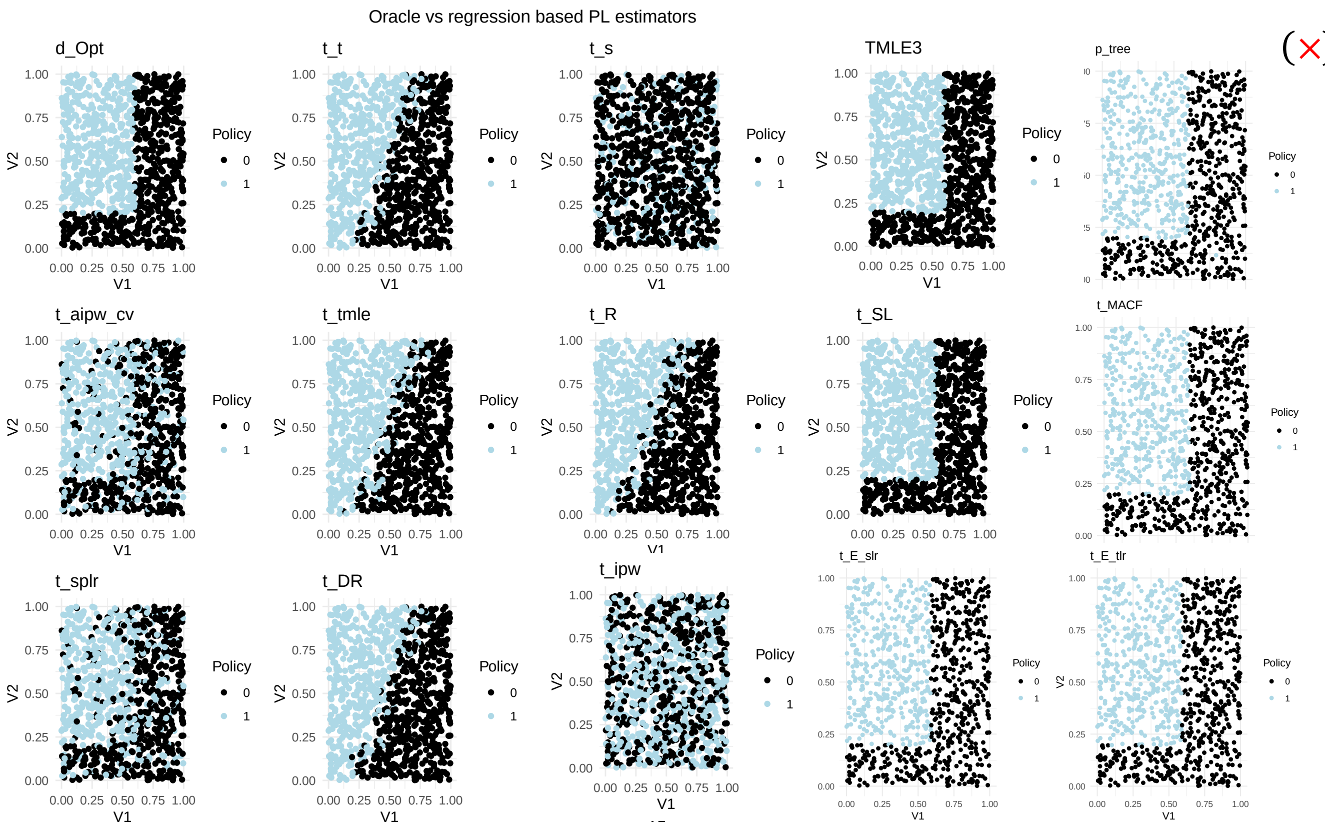
**Figure 1: Policy optimization results for regression and tree-based algorithms in a tree setting**

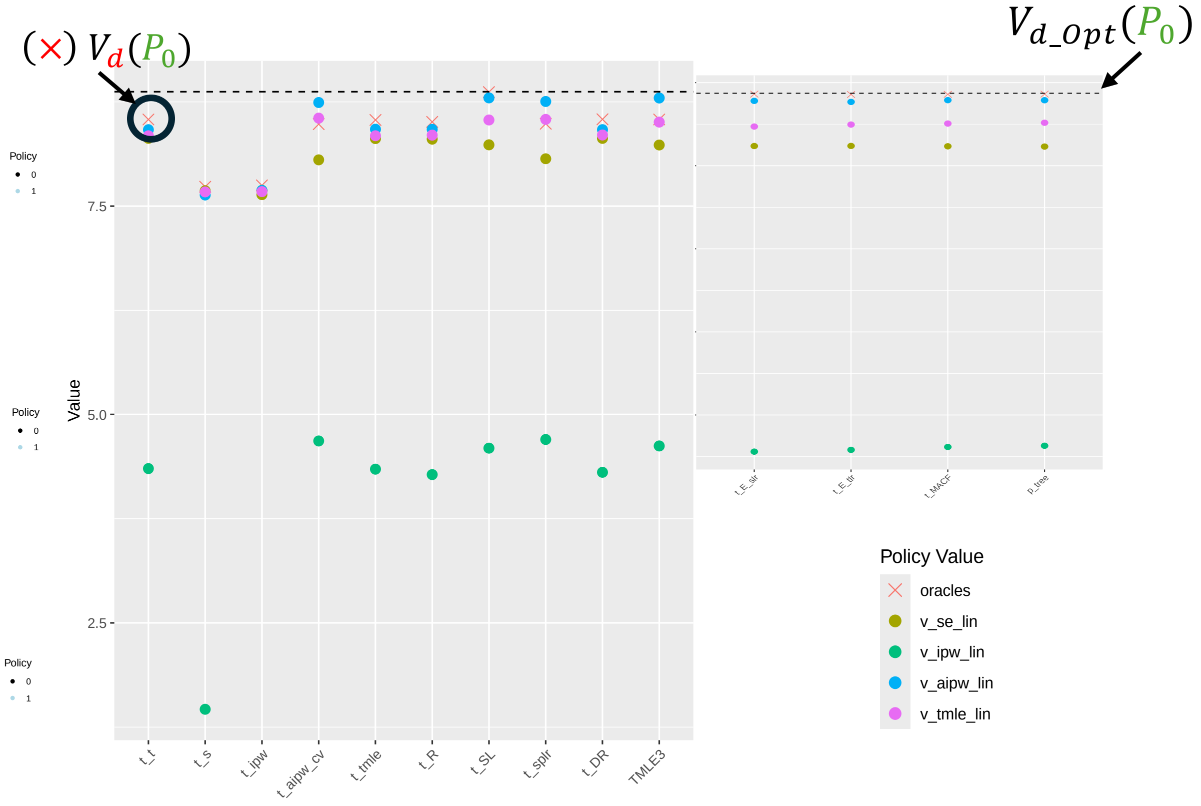
**Left**: Visual representation of treatment rules



$$\hat{\mu}_{w,n}(x) = \hat{\mathbb{E}}_{P_n}[Y|X = x, W = w]$$

# III-Tree setting



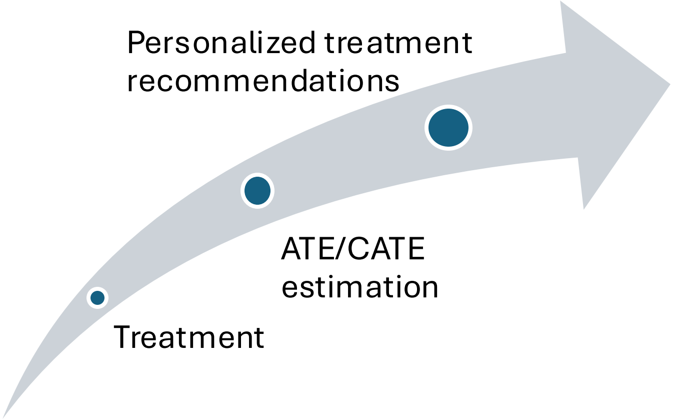

Oracle policy values for $d : \mathbb{E}_{P_0}[Y] = \mathbb{E}_{P_0}[u(X) + d(X)c(X) + Z]$

**Figure 1: Policy optimization results for regression and tree-based algorithms in a tree setting**

**Left**: Visual representation of treatment rules

# IV. Discussion

# Discussion




- Policy optimization and evaluation techniques
- Synthetic simulation for binary treatment
  - To improve:
    - Test multiple n sizes (boxplots)
    - Test non RCT scenario
- Other contributions:
  - Multi-treatment extensions:
    - Policy optimization algorithm
    - Policy evaluation technique
  - Application to IVF data

**Perspectives:** adding constraints to the policy optimization problem (Ph.D thesis)

- Explainability
- Fairness, No-harm criteria ...

# Thank you!

# Annexes

# II.1.1- ATE double robust estimators

**Targeted Maximum Likelihood Estimator (TMLE)**

Build a fluctuation: **Find** a regression $\mathbf{Q}(\boldsymbol{P_n^*}) = \mathbb{E}_{P_n^*}[Y|X=x, W=w]$ **closest** to $Q(P_0) = \mathbb{E}_{P_0}[Y|X=x, W=w]$

$$\mathrm{Q}_{n,\epsilon} = \{(w,x) \to expit\big(logit\big(\widehat{\mathbb{E}}_{P_n}[Y|X=x, W=w]\big) + \epsilon H_n(x,w)\big)\}$$

If $Y \in \{0,1\}$ or $[0,1]$

$$\underset{\epsilon}{\operatorname{argmin}}\, \mathbb{E}_{P_n}[R_n(\epsilon)] = \sum_{i=1}^{n} -Y_i \log(Q_{n,\epsilon}(X_i, W_i)) - (1-Y_i)\log\big(1 - Q_{n,\epsilon}(X_i, W_i)\big)$$

$$H_n(x,w) = \frac{2w - 1}{wP_n(W=1|X=x) + (1-w)P_n(W=0|X=x)}$$

# II.1.1- ATE double robust estimators

**Targeted Maximum Likelihood Estimator (TMLE)**

Build a fluctuation: **Find** a regression $\mathbf{Q}(\boldsymbol{P_n^*}) = \mathbb{E}_{P_n^*}[Y|X = x, W = w]$ **closest** to $Q(P_0) = \mathbb{E}_{P_0}[Y|X = x, W = w]$

If $Y \in \{0,1\}$ or $[0,1]$

$$\mathrm{Q}_{n,\epsilon} = \{(w, x) \to expit\left(logit\left(\widehat{\mathbb{E}}_{P_n}[Y|X = x, W = w]\right) + \epsilon H_n(x, w)\right)\}$$

$$\underset{\epsilon}{\operatorname{argmin}}\, \mathbb{E}_{P_n}[R_n(\epsilon)] = \sum_{i=1}^{n} -Y_i \log(Q_{n,\epsilon}(X_i, W_i)) - (1 - Y_i) \log\left(1 - Q_{n,\epsilon}(X_i, W_i)\right)$$

If $Y \in [a, b]$, $a < b$

$$\mathrm{Q}_{n,\epsilon} = \{(w, x) \to \left(\widehat{\mathbb{E}}_{P_n}[Y|X = x, W = w] + \epsilon H_n(x, w)\right)\}$$

$$\underset{\epsilon}{\operatorname{argmin}}\, \mathbb{E}_{P_n}[R_n(\epsilon)] = \sum_{i=1}^{n}(Y_i - Q_{n,\epsilon}(X_i, W_i))^2$$

$$H_n(x, w) = \frac{2w - 1}{wP_n(W = 1|X = x) + (1 - w)P_n(W = 0|X = x)}$$

# II.1.1- ATE double robust estimators

**Targeted Maximum Likelihood Estimator (TMLE)**

$$\mathbf{Q}(\boldsymbol{P_n^*}) = \mathbb{E}_{P_n^*}[Y|X = x, W = w]$$

$$= \widehat{\mathbb{E}}_{P_n}[Y|X = x, W = w] + \boldsymbol{\epsilon_n}\, H_n(x, w)$$

$$\mathbb{E}_{P_n^*}[Y|X, W = 1] = \widehat{\mathbb{E}}_{P_n}[Y|X, W = 1] + \boldsymbol{\epsilon_n}\, H_n(x, 1)$$

$$\mathbb{E}_{P_n^*}[Y|X, W = 0] = \widehat{\mathbb{E}}_{P_n}[Y|X, W = 0] + \boldsymbol{\epsilon_n}\, H_n(x, 0)$$

$$\psi(P_n^*) = \mathbb{E}_{P_n^*}[\mathbb{E}_{P_n^*}[Y|X, W = 1] - \mathbb{E}_{P_n^*}[Y|X, W = 0]]$$

$$H_n(x, w) = \frac{2w - 1}{wP_n(W = 1|X = x) + (1 - w)P_n(W = 0|X = x)}$$

# A VISUAL GUIDE TO POLICY EVALUATION