Proximal Causal Inference for Modified Treatment Policies

Andrea Rotnitzky
Professor of Biostatistics, University of Washington

Based on https://doi.org/10.48550/arXiv.2512.12038

April 15, 2026

Joint work with

Antonio Olivas-Martinez
University of Washington

Dr. Peter Gilbert
Fred Hutch Cancer Center,
University of Washington

Outline

- Biomarkers in vaccine research
- Negative control outcomes and treatments
- Modified treatment policies (MTPs): Identification in the absence of unmeasured confounding.
- Modified treatment policies (MTPs): Proximal Causal Inference for MTP means with unmeasured confounding:
  - **Identification**
  - **Estimation**
- Simulation experiments
- Application to the analysis of a vaccine trial

Outline

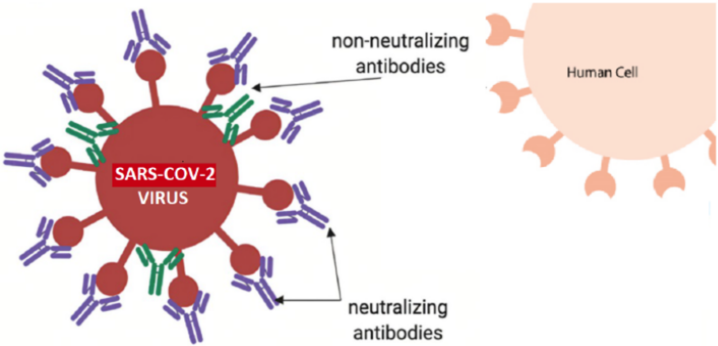
- **Biomarkers in vaccine research**
- Negative control outcomes and treatments
- Modified treatment policies (MTPs): Identification in the absence of unmeasured confounding.
- Modified treatment policies (MTPs): Proximal Causal Inference for MTP means with unmeasured confounding
  - **Identification**
  - **Estimation**
- Simulation experiments
- Application to the analysis of a vaccine trial

Biomarkers in vaccine research

- Large Phase 3 vaccine trials require prolonged follow-up because infection may take time to occur
- This generates need for identifying early immune markers that can serve as surrogate outcomes—measured sooner and used in place of infection endpoints—to aid in the evaluation of vaccines and thereby accelerate vaccine evaluation and approval
- Neutralizing antibody (NAbs) titers measured at a fixed time after vaccination are widely regarded as valuable candidate surrogate markers.
- Establishing their validity as surrogate markers, however, requires several distinct stages of scientific investigation to evaluate the causal effect of NAb on infection:
  - Animal studies
  - Randomized studies of monoclonal antibody infusion
  - Analysis of earlier Phase 3 vaccine trials that recorded both NAb titers and infection occurrence

Neutralizing and binding antibodies

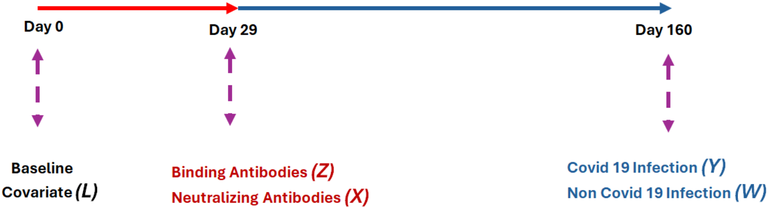
- **Candidate biomarker:** Concentration of neutralizing antibodies in the bloodstream against the current virus strain.



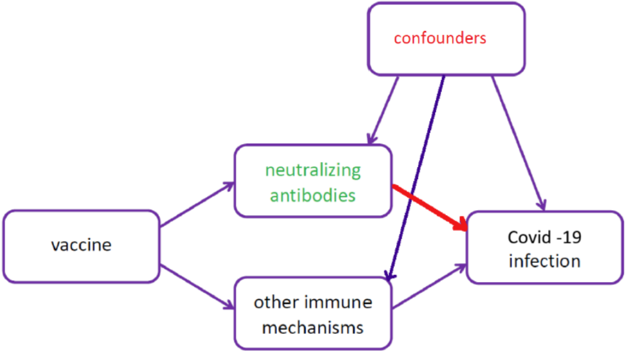
The Ensemble COVID-19 Vaccine Trial

- Multinational, placebo-controlled trial of efficacy of a single dose of Ad26.Cov2.S in preventing COVID-19 infection

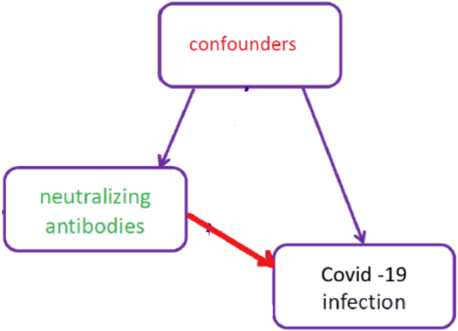
- Data in vaccine arm:



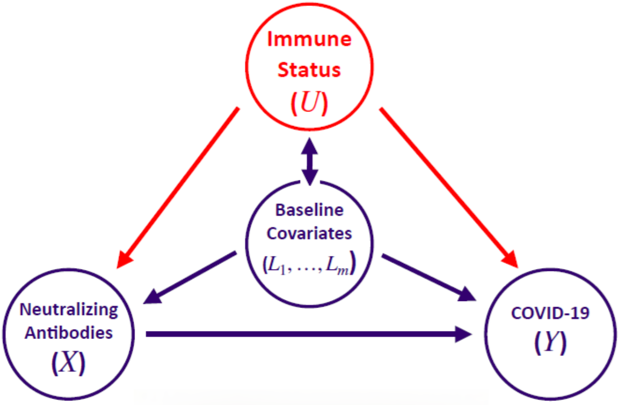
Confounding



Confounding



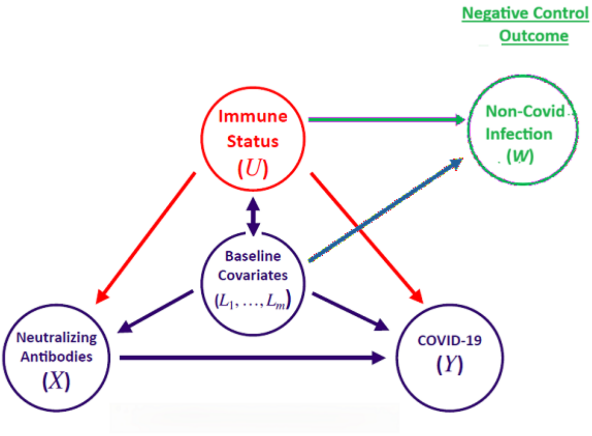
Unmeasured confounding



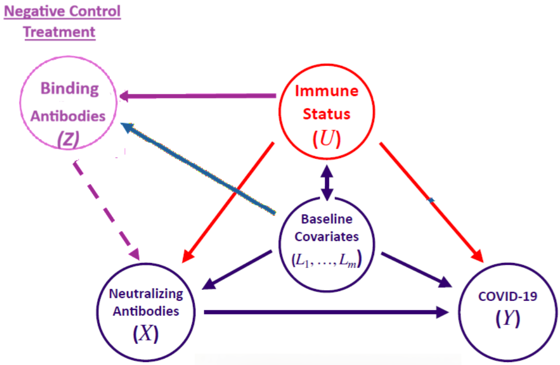
Outline

- Biomarkers in vaccine research
- **Negative control outcomes and treatments**
- Modified treatment policies (MTPs): Identification in the absence of unmeasured confounding.
- Modified treatment policies (MTPs): Proximal Causal Inference for MTP means with unmeasured confounding
  - **Identification**
  - **Estimation**
- Simulation experiments
- Application to the analysis of a vaccine trial

Negative control outcomes

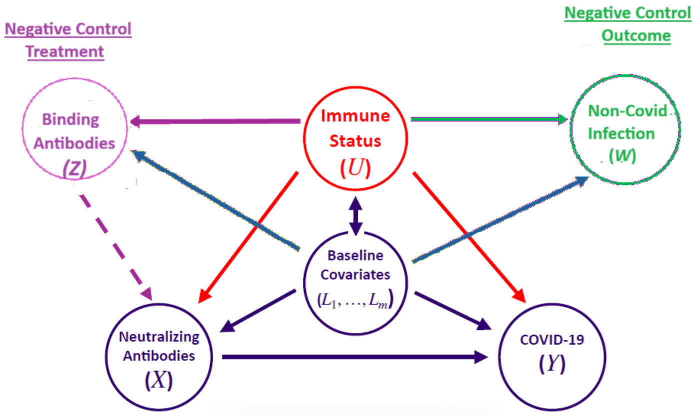


Negative control treatments



Proximal Inference

- **Tchetgen-Tchetgent, Miao and colleagues (**Miao et al., 2018 and Tchetgen Tchetgen et al., 2020 and more...): causal inference with unmeasured confounding when both negative control treatments and outcomes are available.



Outline

- Biomarkers in vaccine research
- Negative control outcomes and treatments
- **Modified treatment policies (MTPs): Identification in the absence of unmeasured confounding.**
- Modified treatment policies (MTPs): Proximal Causal Inference for MTP means with unmeasured confounding
  - **Identification**
  - **Estimation**
- Simulation experiments
- Application to the analysis of a vaccine trial

Dose response curve

- **Data**: n i.i.d copies of

$$\left(\underbrace{L}_{\substack{\text{baseline}\\\text{covariates}}}, \underbrace{X}_{\substack{\text{NAb}\\\text{concentration}}}, \underbrace{Y}_{\text{infection}} \right)$$

- **Counterfactuals for subject** i:

$$\{Y_i(x) : x \in \mathcal{X}\}$$

- **Dose-response curve:**

$$x \mapsto \mu(x) \equiv E[Y(x)]$$

Dose response curve

- **Data**: n i.i.d copies of

$$\left(\underbrace{L}_{\substack{\text{baseline}\\\text{covariates}}}, \underbrace{X}_{\substack{\text{NAb}\\\text{concentration}}}, \underbrace{Y}_{\text{infection}} \right)$$

- **Counterfactuals for subject** i:

$$\{Y_i(x) : x \in \mathcal{X}\}$$

- **Dose-response curve:**

$$x \mapsto \mu(x) \equiv E[Y(x)]$$

- **If X had been fully randomized**, then $Y(x)$ would have been independent of X and therefore

$$x \mapsto \mu(x) = E[Y|X = x]$$

would be the **conditional mean curve.**

Dose response curve

- Suppose we were prepared to assume our study emulated a **conditionally randomized trial given observed baseline covariates** L.

Dose response curve

- Suppose we were prepared to assume our study emulated a **conditionally randomized trial given observed baseline covariates** L. Then, the following exchangeability condition would hold

$$E\left[Y(x) \mid L = l\right] = E\left[Y \mid X = x, L = l\right]$$

Dose response curve

- Suppose we were prepared to assume our study emulated a **conditionally randomized trial given observed baseline covariates** L. Then, the following exchangeability condition would hold

$$E\left[Y\left(x\right)|L=l\right]=E\left[Y|X=x,L=l\right]$$

- provided **positivity holds,** that is

$$f\left(x|l\right)>0 \text{ for each possible value } l \text{ of } L$$

Dose response curve

- Suppose we were prepared to assume our study emulated a **conditionally randomized trial given observed baseline covariates** L. Then, the following exchangeability condition would hold

$$E\left[Y\left(x\right)|L=l\right]=E\left[Y|X=x,L=l\right]$$

- provided **positivity holds,** that is

$$f\left(x|l\right)>0 \text{ for each possible value } l \text{ of } L$$

  Then

$$\mu\left(x\right)\equiv E\left\{e_0\left(x,L\right)\right\}$$

  where

$$e_0\left(x,l\right)\equiv E\left[Y|X=x,L=l\right]$$

Dose response curve

- Suppose we were prepared to assume our study emulated a **conditionally randomized trial given observed baseline covariates** L. Then, the following exchangeability condition would hold

$$E\left[Y(x)\,|L=l\right]=E\left[Y|X=x,L=l\right]$$

- provided **positivity holds,** that is

$$f\left(x|l\right)>0 \text{ for each possible value } l \text{ of } L$$

  Then

$$\mu\left(x\right)\equiv E\left\{e_0\left(x,L\right)\right\}$$

  where

$$e_0\left(x,l\right)\equiv E\left[Y|X=x,L=l\right]$$

- **Issues:** Some NAb levels are implausible for some type of subjects $\Longrightarrow$
  - Hypothetical worlds in which those subjects will be exposed to such levels are unrealistic and uninteresting.

Dose response curve

- Suppose we were prepared to assume our study emulated a **conditionally randomized trial given observed baseline covariates** L. Then, the following exchangeability condition would hold

$$E\left[Y(x)\,|L=l\right]=E\left[Y|X=x,L=l\right]$$

- provided **positivity holds,** that is

$$f(x|l)>0 \text{ for each possible value } l \text{ of } L$$

  Then

$$\mu(x)\equiv E\left\{e_0(x,L)\right\}$$

  where

$$e_0(x,l)\equiv E\left[Y|X=x,L=l\right]$$

- **Issues:** Some NAb levels are implausible for some type of subjects $\Longrightarrow$
  - Hypothetical worlds in which those subjects will be exposed to such levels are unrealistic and uninteresting.
  - failure of positivity.

Modified treatment policy (MTP) curve

- The curve

$$\delta \to \psi(\delta) \equiv E\left[Y\left(X+\delta\right)\right],\ \delta > 0$$

quantifies how infection risk changes as everyone's NAb concentration is δ units greater than their naturally occurring concentration.

Modified treatment policy (MTP) curve

- The curve
$$\delta \to \psi(\delta) \equiv E\left[Y\left(X+\delta\right)\right],\ \delta > 0$$
quantifies how infection risk changes as everyone's NAb concentration is δ units greater than their naturally occurring concentration.

- Suppose we were prepared to assume the following weaker exchangeability condition held
$$\begin{aligned} E\left[Y\left(X+\delta\right)|X=x, L=l\right] &= E\left[Y|X=x+\delta, L=l\right] \\ &\equiv e_0\left(x+\delta, l\right) \end{aligned}$$

- In words: within levels of L, the infection risk among those with natural NAb level x , had they boosted their level to $x+\delta$ is equal to the infection risk of those whose natural NAb level was $x+\delta$.

Modified treatment policy (MTP) curve

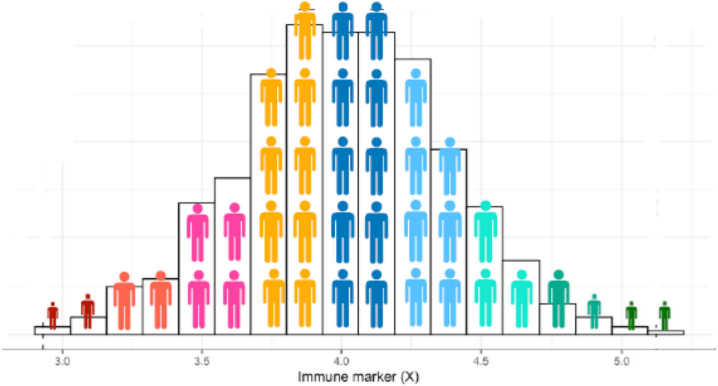
- The curve
$$\delta \to \psi(\delta) \equiv E\left[Y\left(X+\delta\right)\right],\ \delta > 0$$
quantifies how infection risk changes as everyone's NAb concentration is δ units greater than their naturally occurring concentration.

- Suppose we were prepared to assume the following weaker exchangeability condition held
$$\begin{aligned} E\left[Y\left(X+\delta\right)|X=x, L=l\right] &= E\left[Y|X=x+\delta, L=l\right] \\ &\equiv e_0\left(x+\delta, l\right) \end{aligned}$$

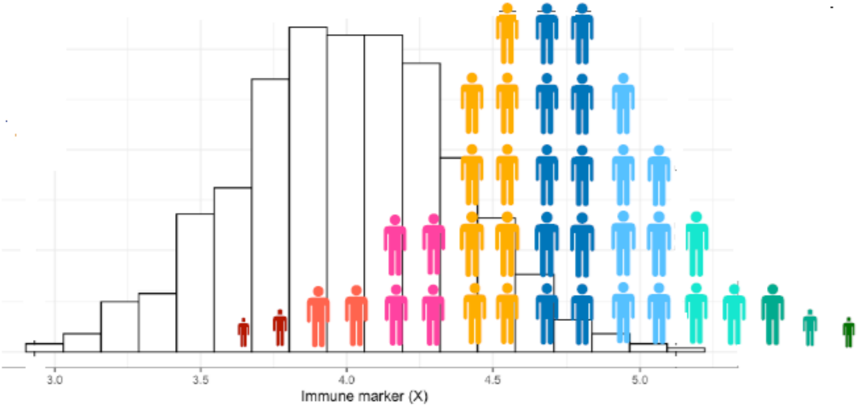
- In words: within levels of L, the infection risk among those with natural NAb level x , had they boosted their level to $x+\delta$ is equal to the infection risk of those whose natural NAb level was $x+\delta$.

- Under such assumption,
$$\psi(\delta) \equiv E\left[e_0\left(X+\delta, L\right)\right]$$

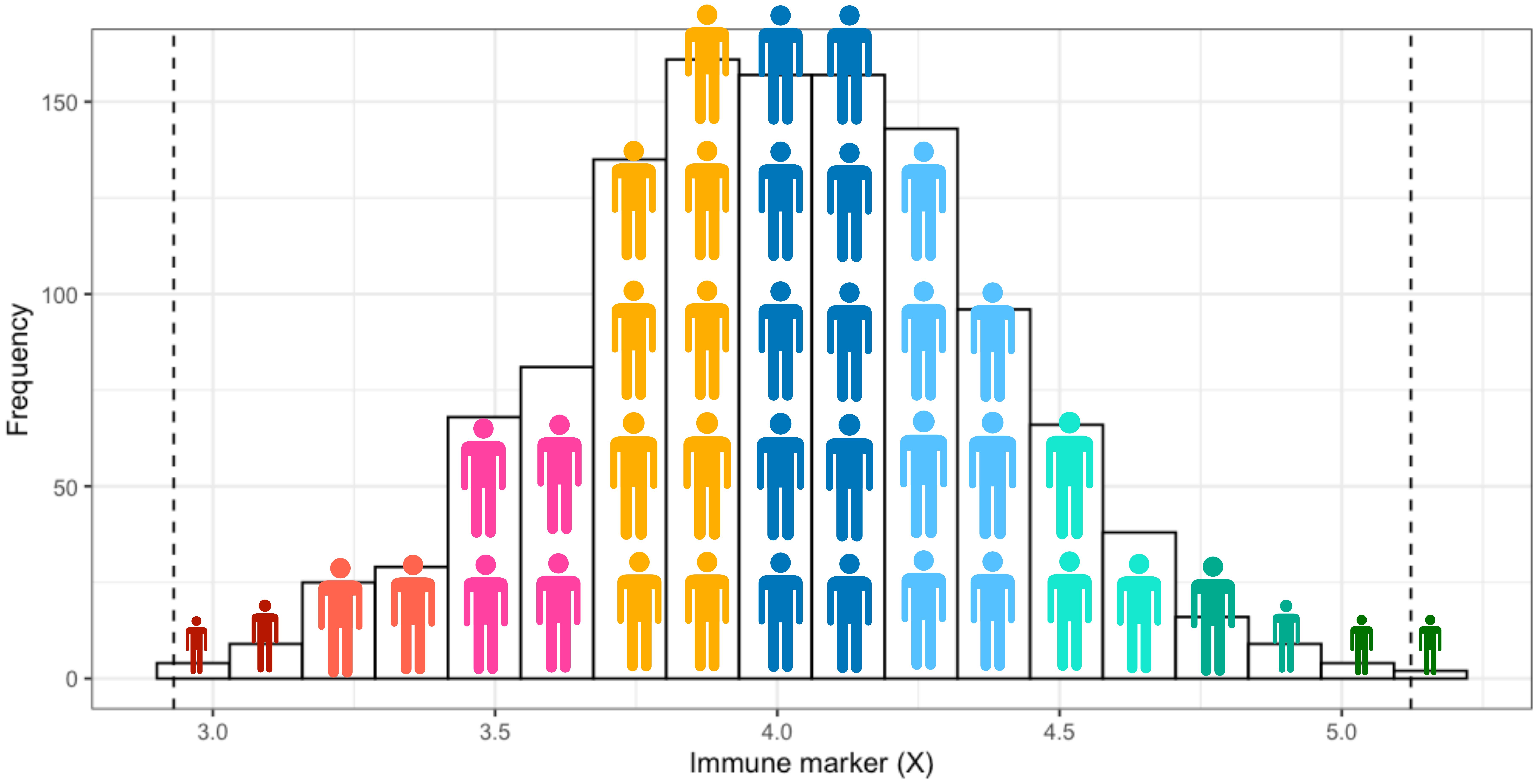
Challenge: lack of positivity



Challenge: lack of positivity

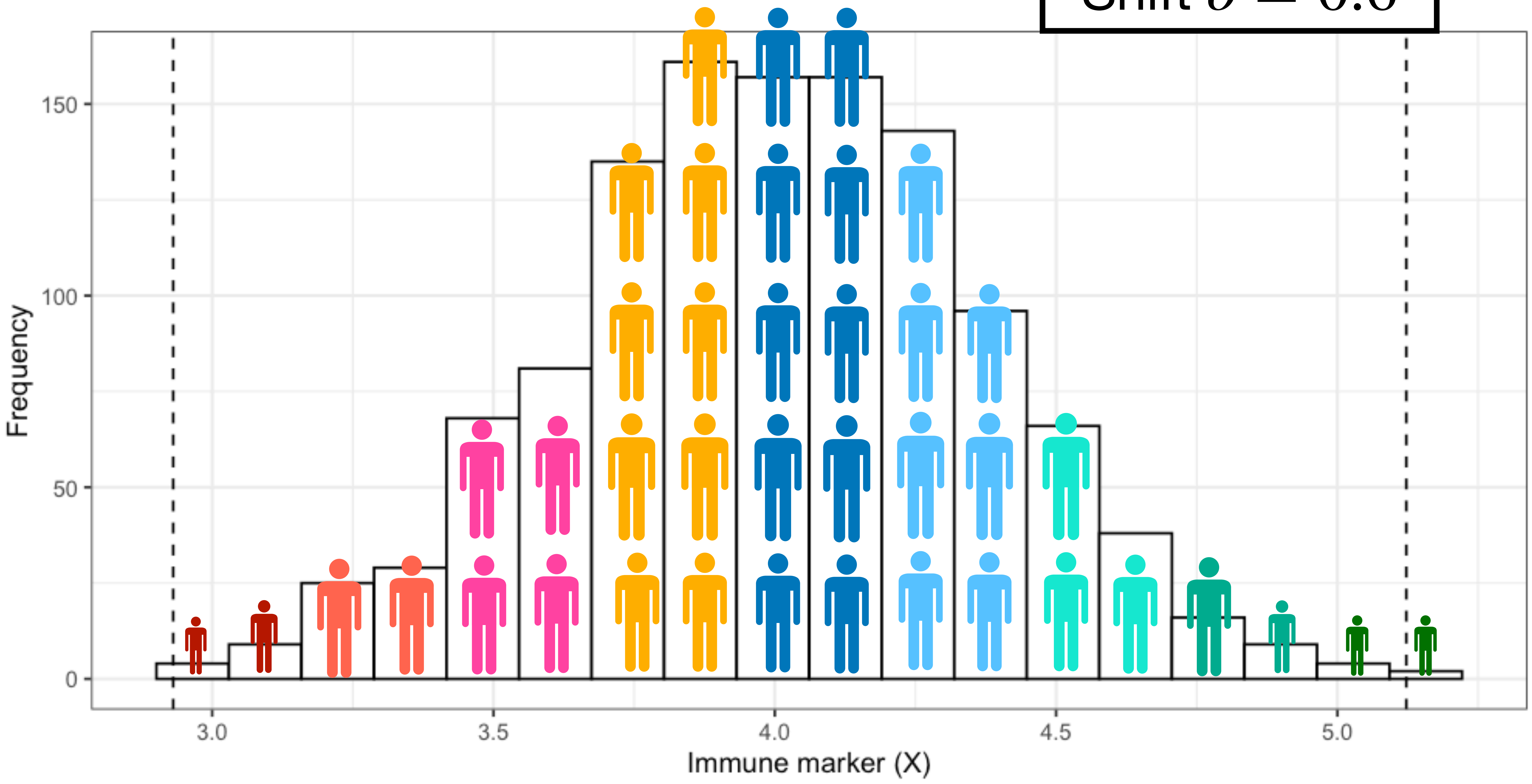


Strategy 1: Restrict the Population

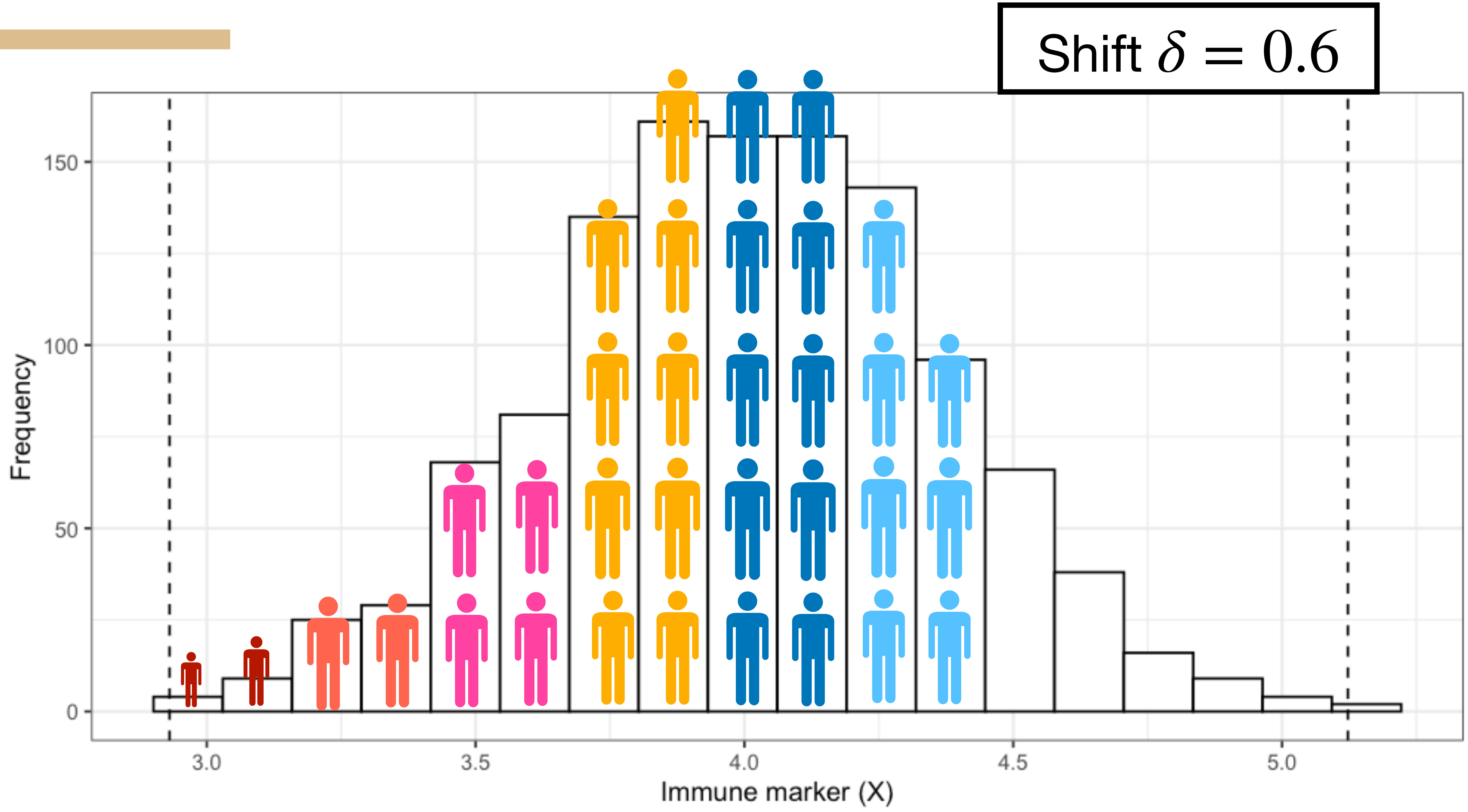


Strategy 1: Restrict the Population

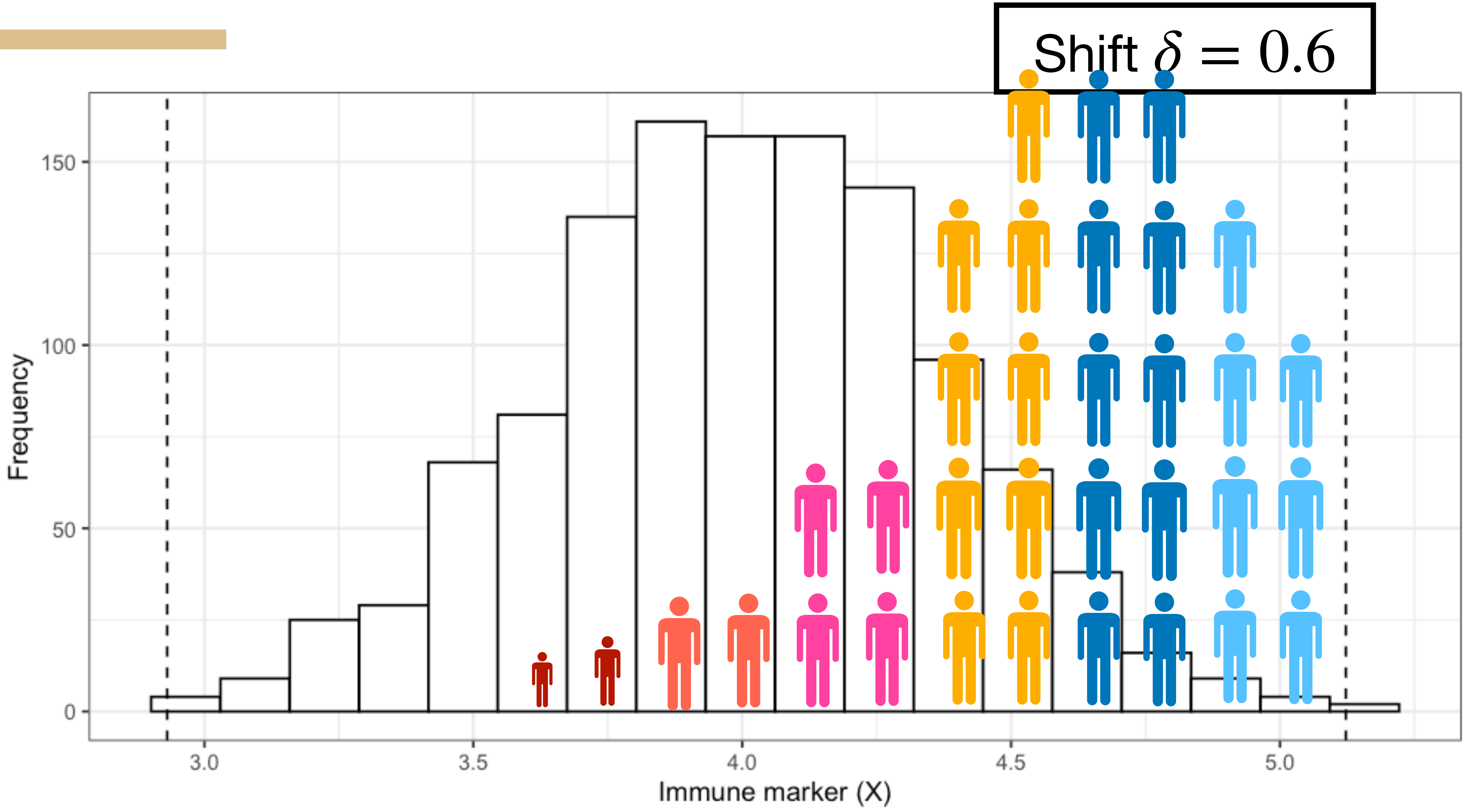
Shift $\delta = 0.6$



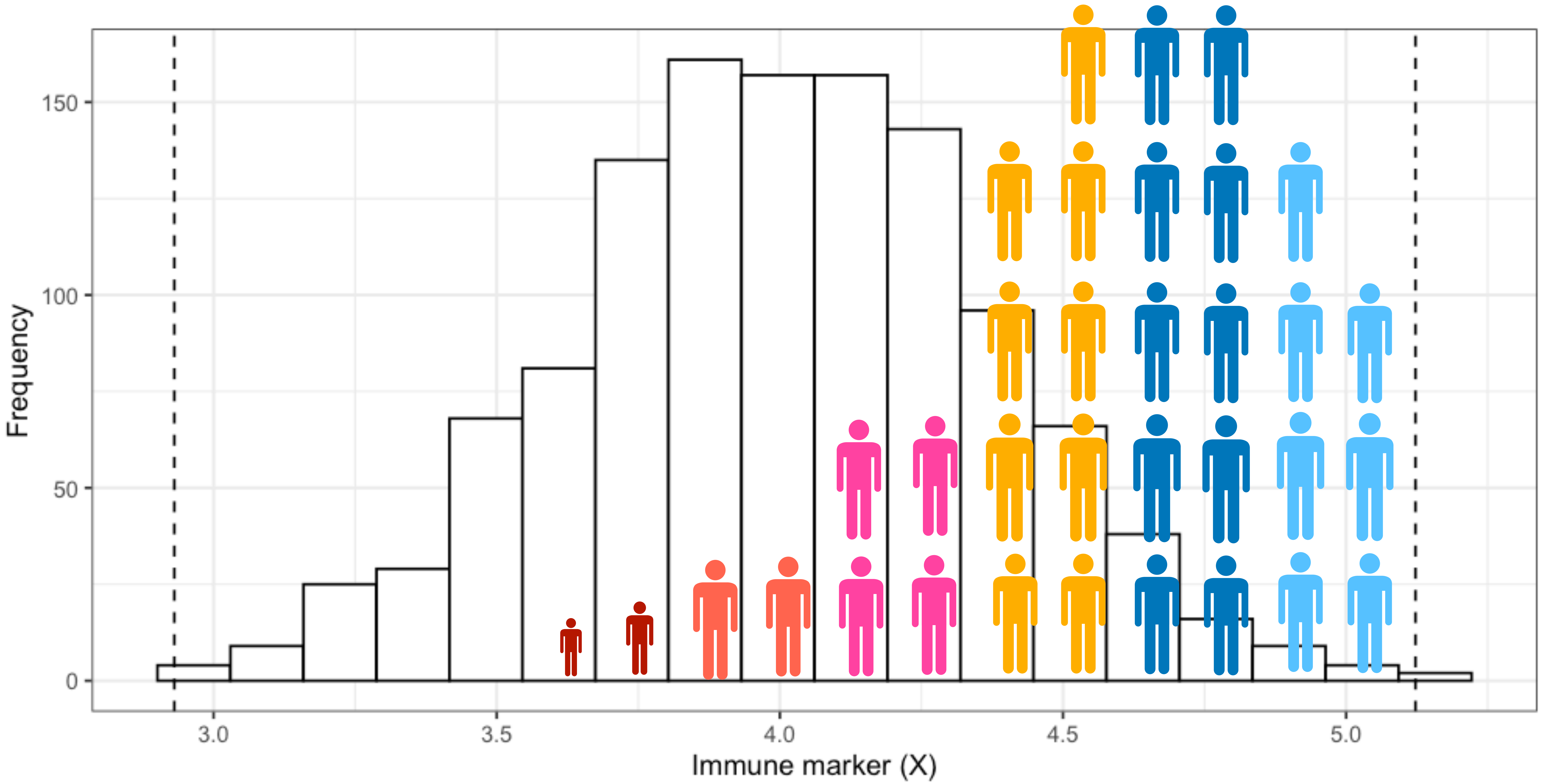
Strategy 1: Restrict the Population



Strategy 1: Restrict the Population



Strategy 1: Restrict the Population



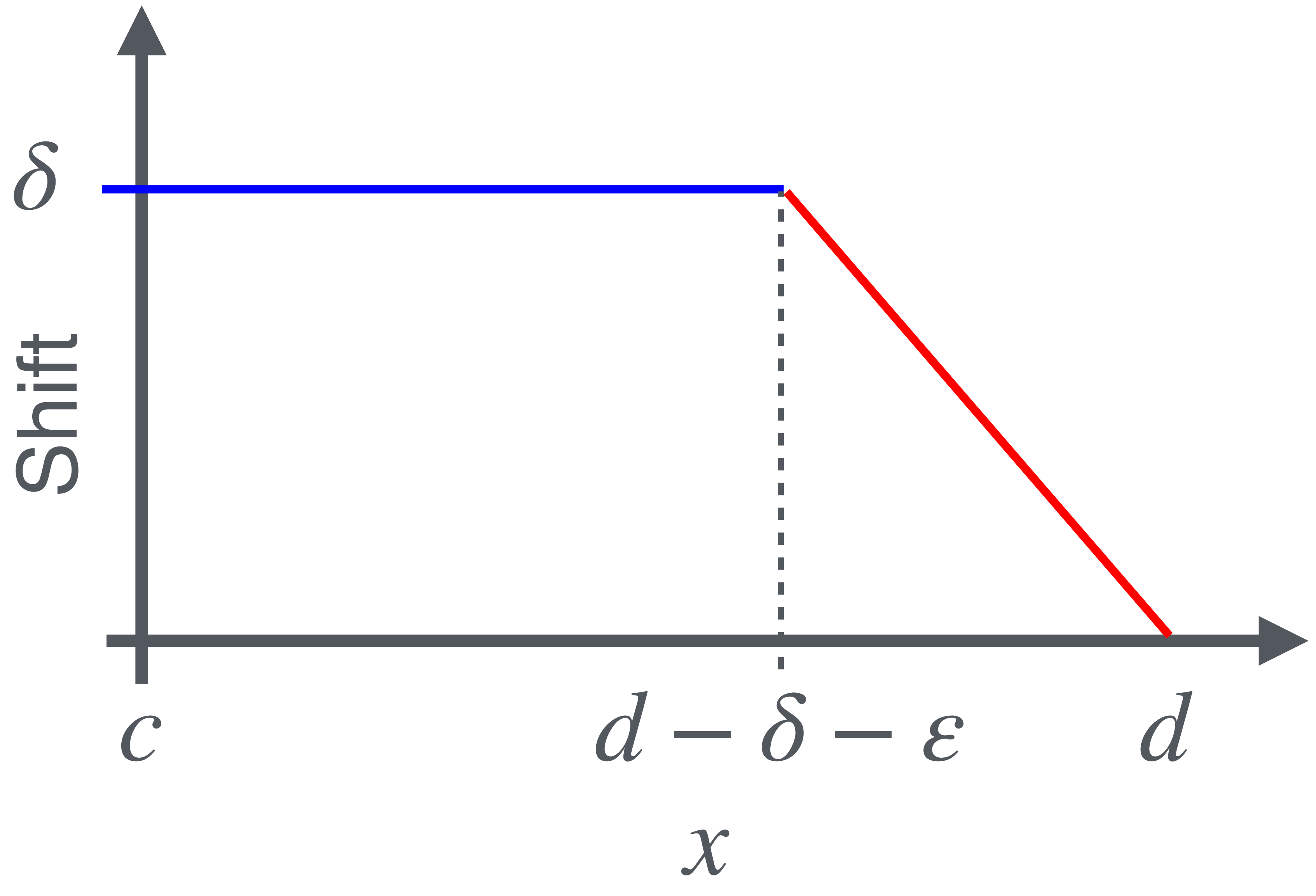
Target parameter: $\psi(\delta) = \mathbb{E}\left\{Y(X+\delta) \,|\, X \in [3\,,\; 5.1-\delta]\right\}$

Strategy 2: Change the policy

Choose $\varepsilon > 0$ such that $c + \delta < d - \varepsilon$.

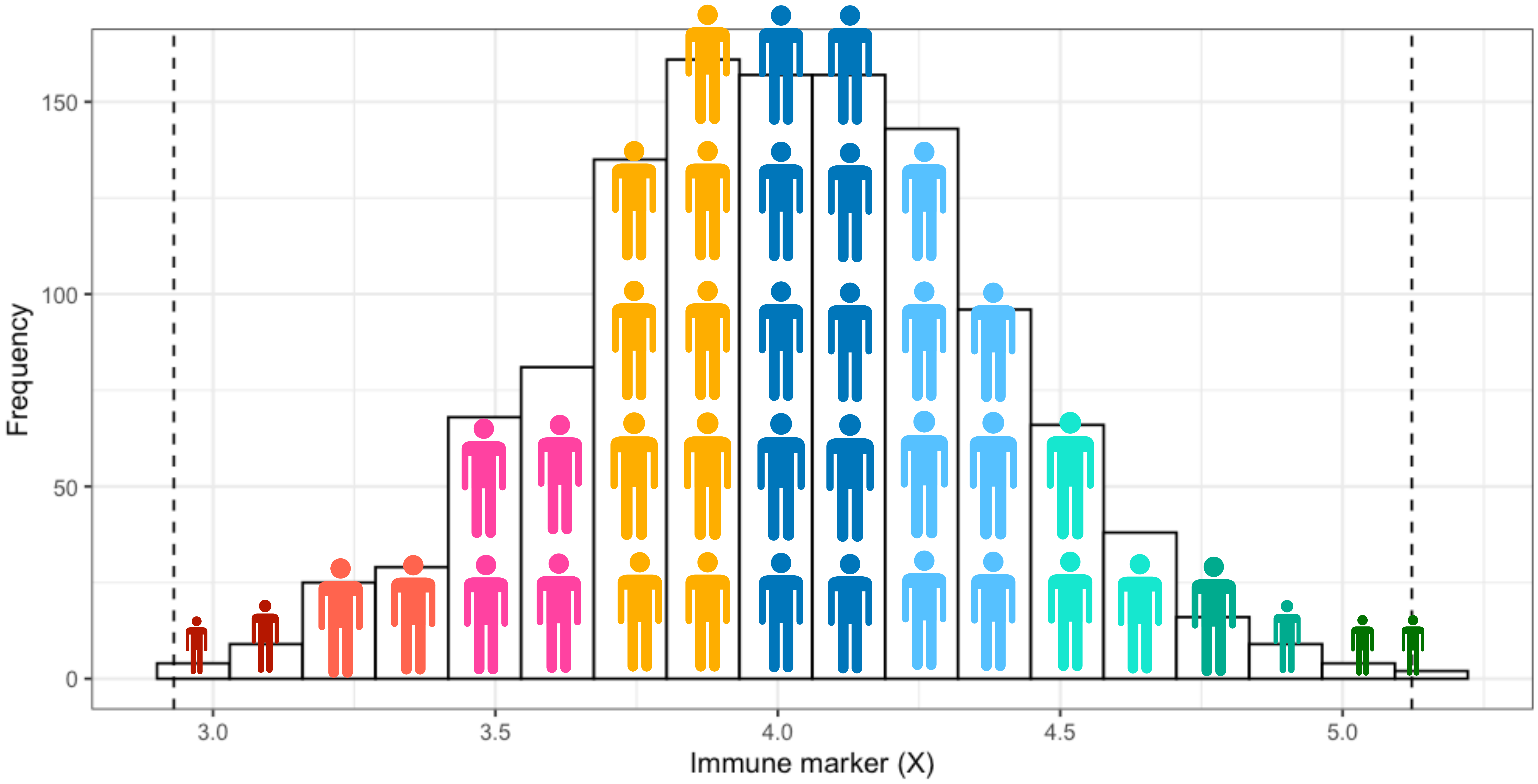
The following policy does not suffer lack of positivity:

$$q(x) = \begin{cases} x + \delta, & \boxed{\text{if } x \in [c, d - \delta - \varepsilon],} \\ x + \frac{\delta}{\delta + \varepsilon}[d - x], & \boxed{\text{if } x \in (d - \delta - \varepsilon, d]} \end{cases}$$



Strategy 2. Change the Policy

Shift $\delta = 0.6$

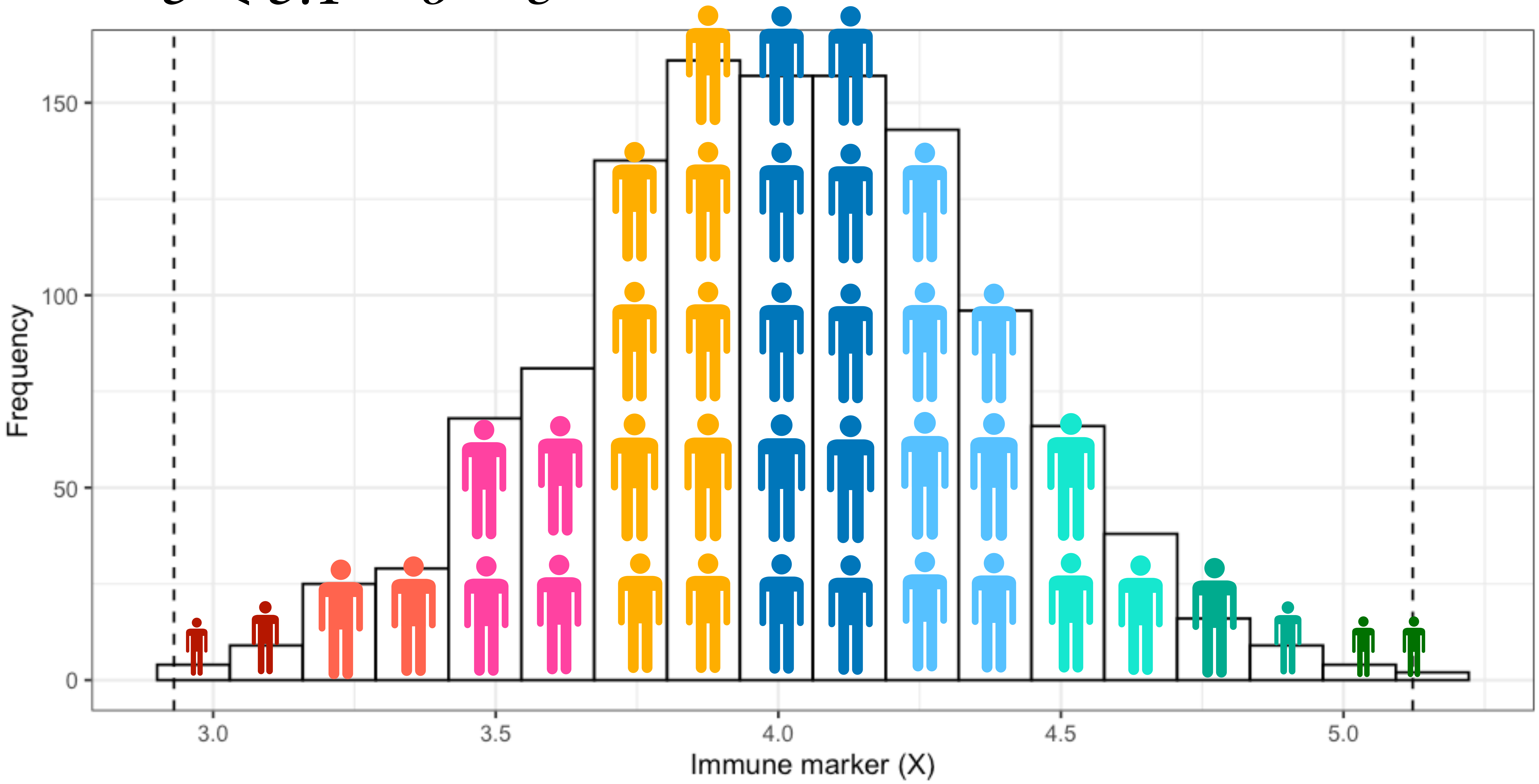


Strategy 2. Change the Policy

Choose $\varepsilon > 0$ such that

$$3 < 5.1 - \delta - \varepsilon$$

Shift $\delta = 0.6$

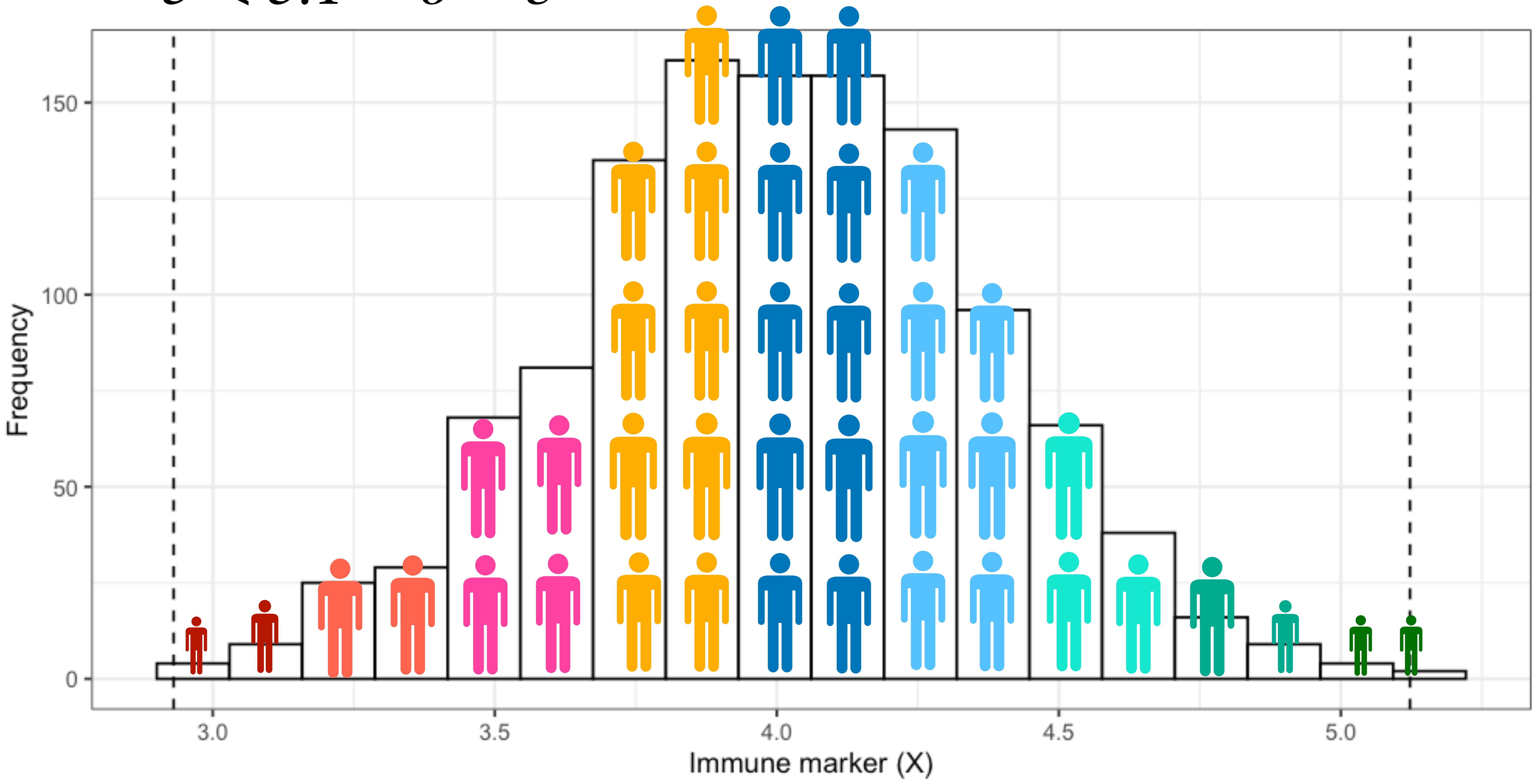


Strategy 2. Change the Policy

Choose $\varepsilon > 0$ such that

$$3 < 5.1 - \delta - \varepsilon$$

Shift $\delta = 0.6$

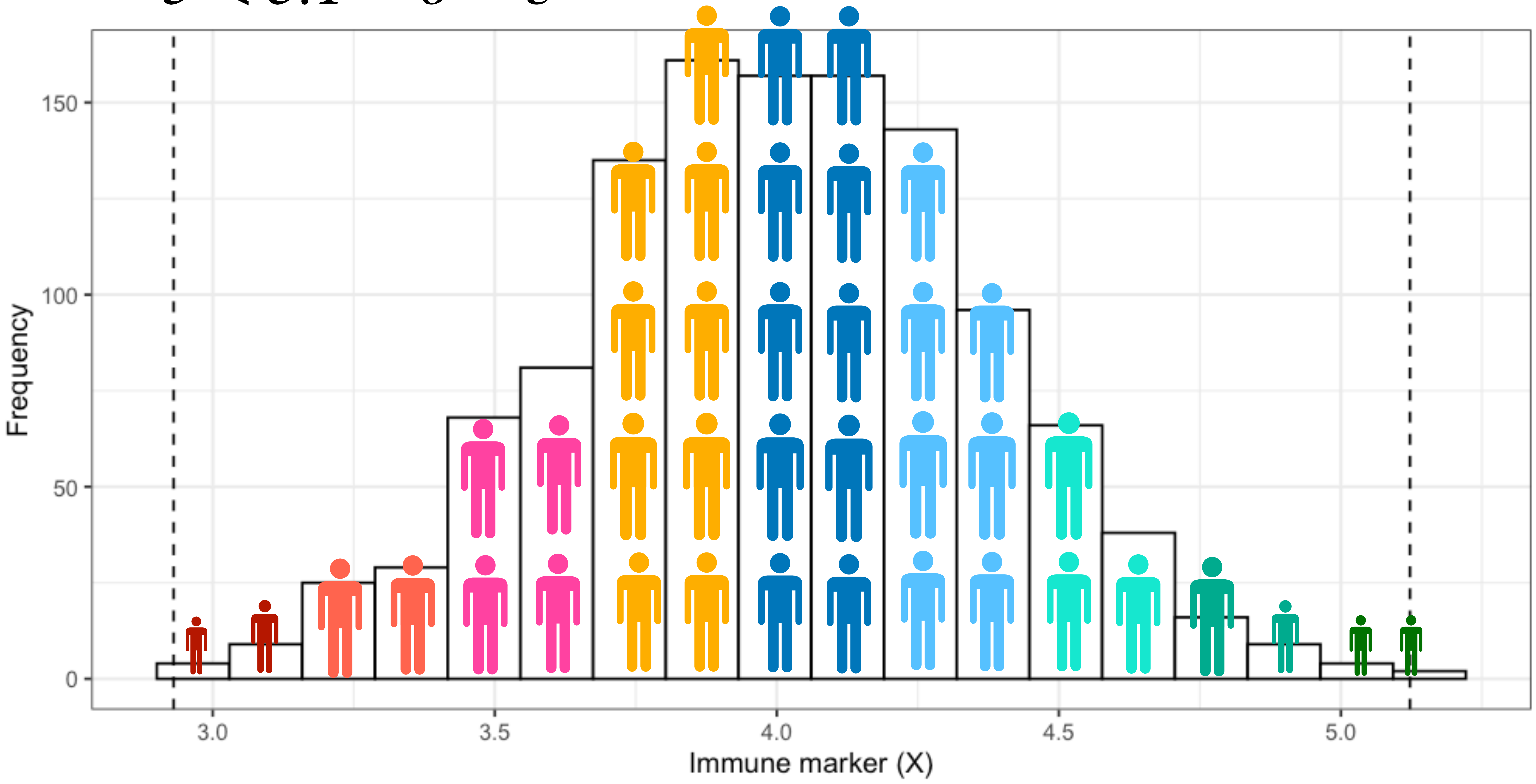


Strategy 2. Change the Policy

Choose $\varepsilon > 0$ such that

$$3 < 5.1 - \delta - \varepsilon$$

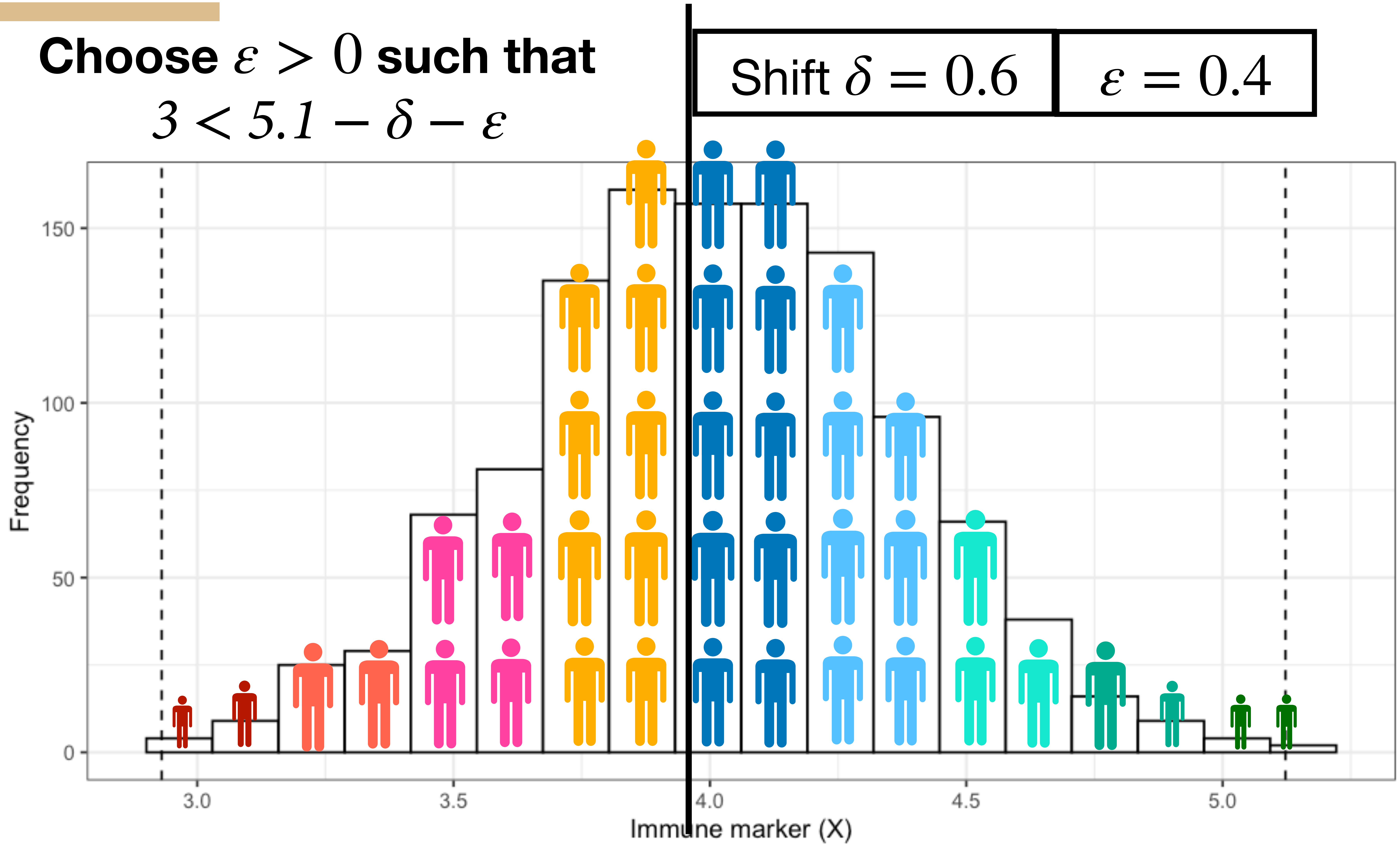
Shift $\delta = 0.6$	$\varepsilon = 0.4$



Strategy 2. Change the Policy

Choose $\varepsilon > 0$ such that

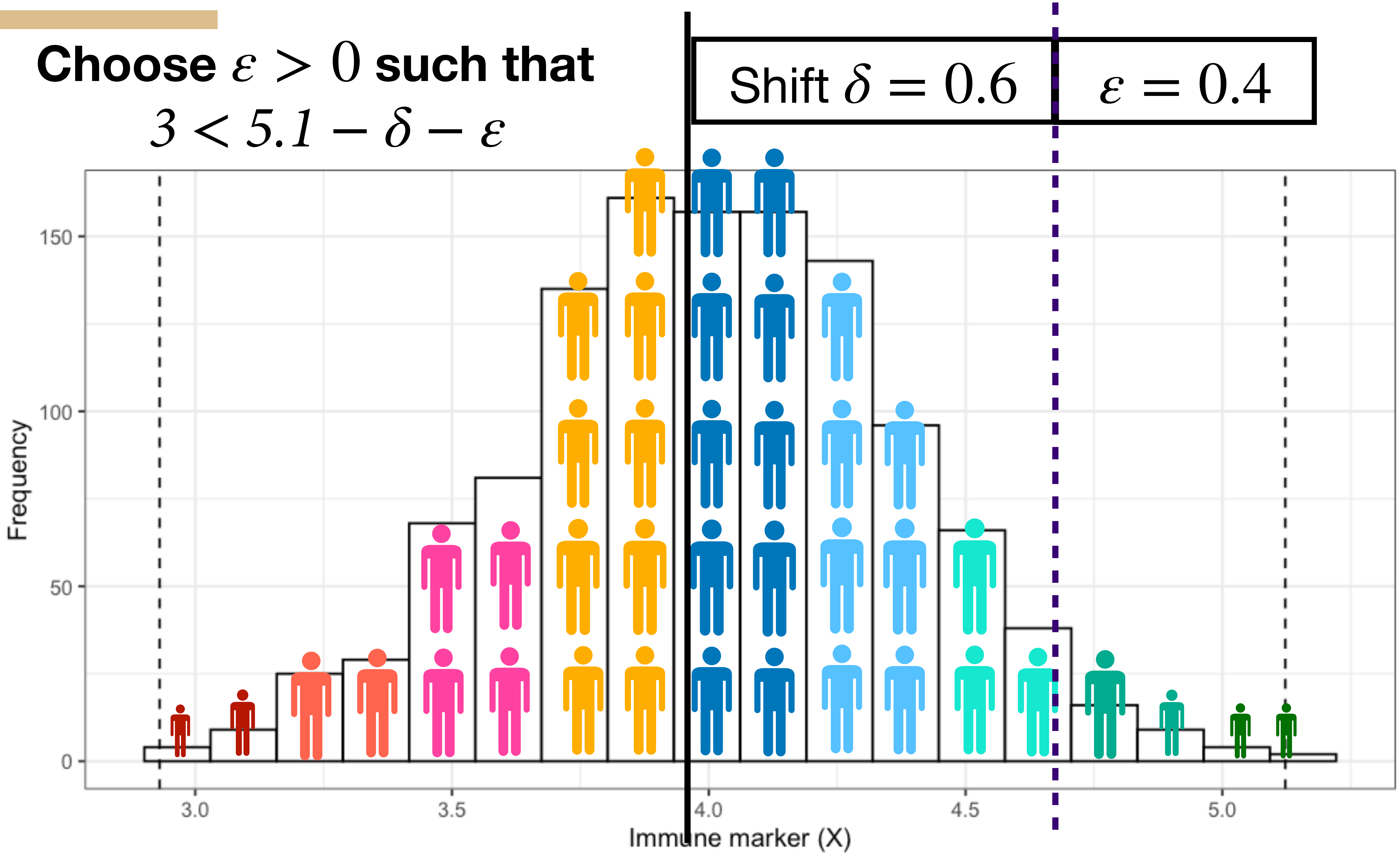
$$3 < 5.1 - \delta - \varepsilon$$



Strategy 2. Change the Policy

Choose $\varepsilon > 0$ such that

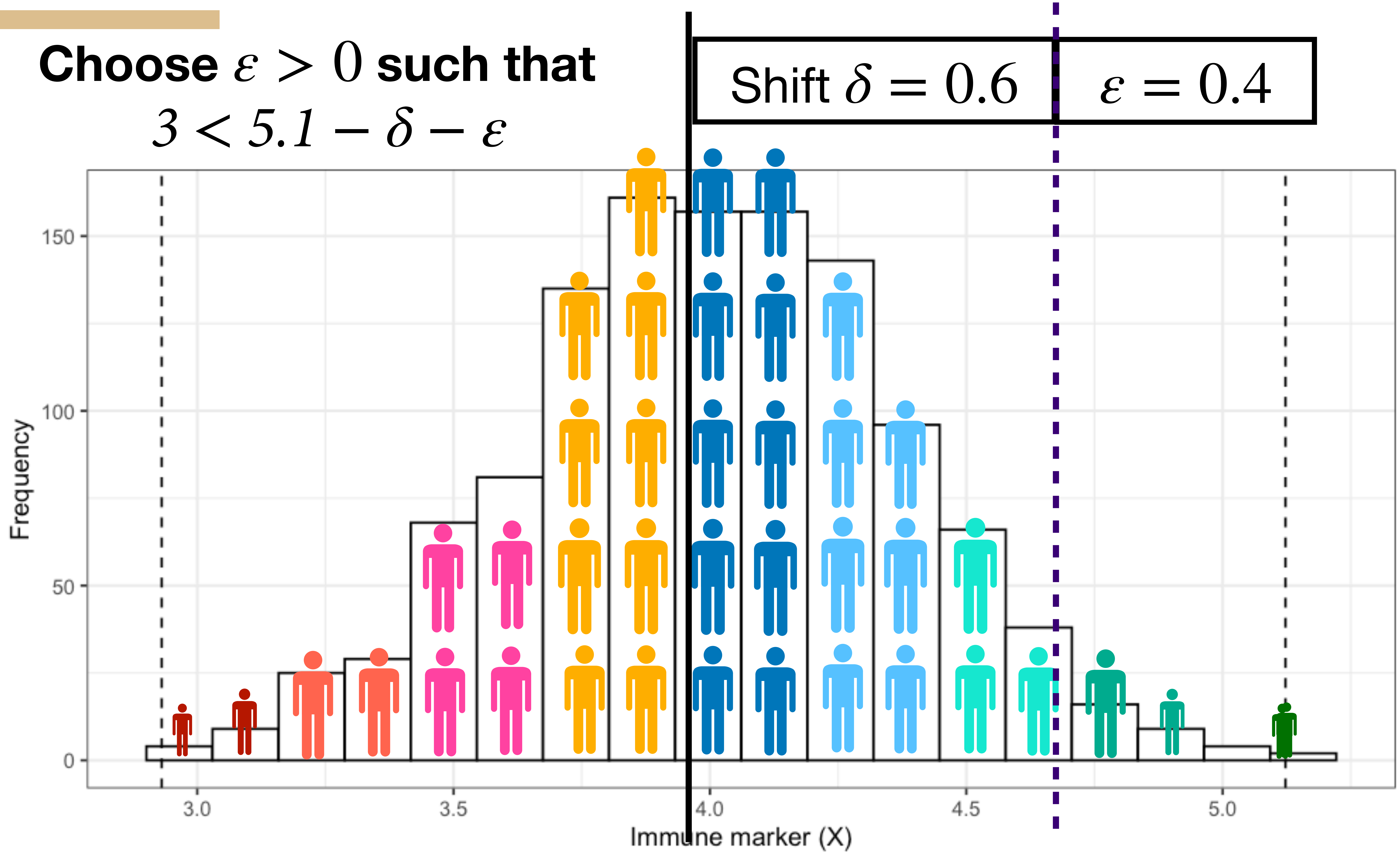
$$3 < 5.1 - \delta - \varepsilon$$



Strategy 2. Change the Policy

Choose $\varepsilon > 0$ such that

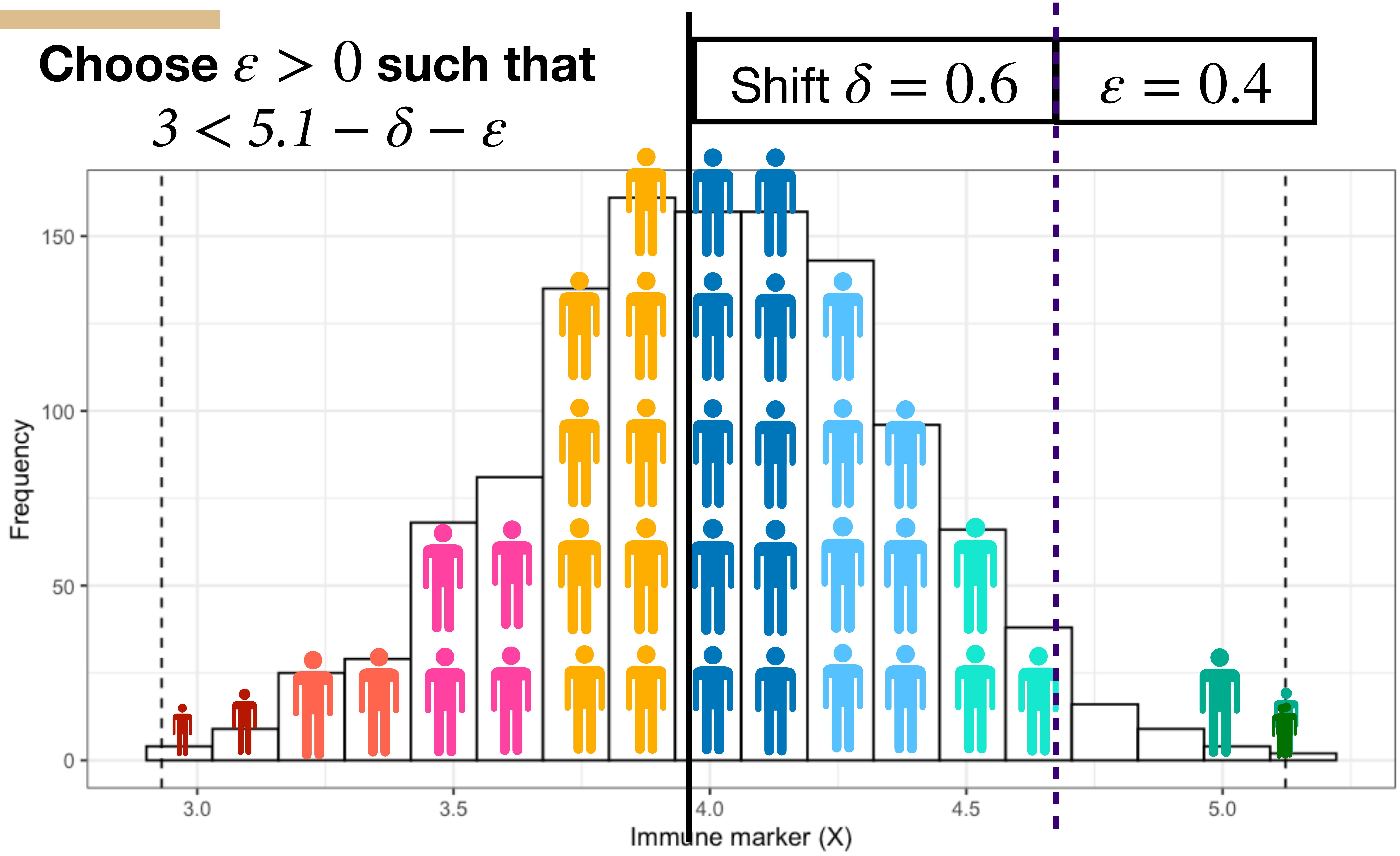
$$3 < 5.1 - \delta - \varepsilon$$



Strategy 2. Change the Policy

Choose $\varepsilon > 0$ such that

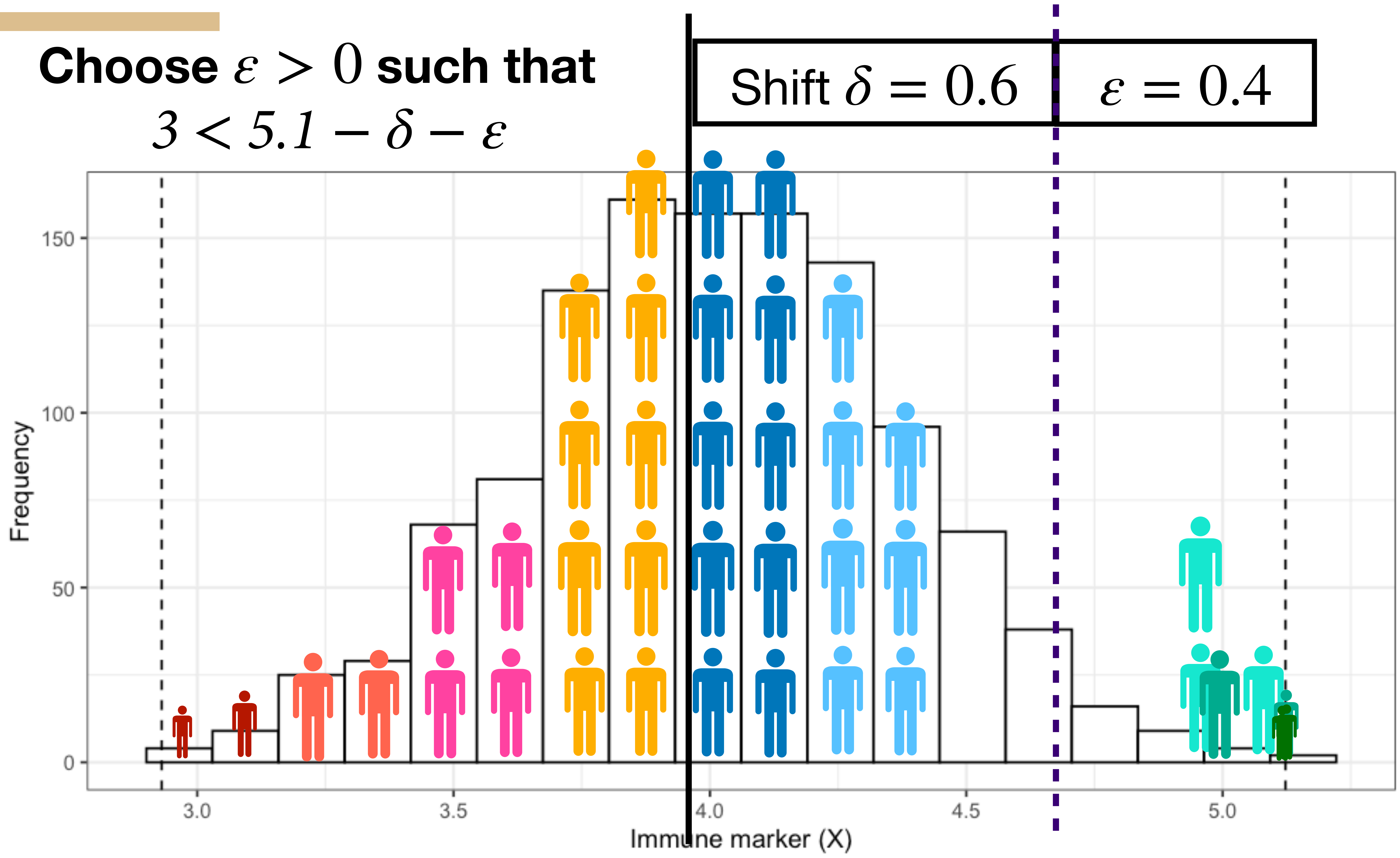
$$3 < 5.1 - \delta - \varepsilon$$



Strategy 2. Change the Policy

Choose $\varepsilon > 0$ such that

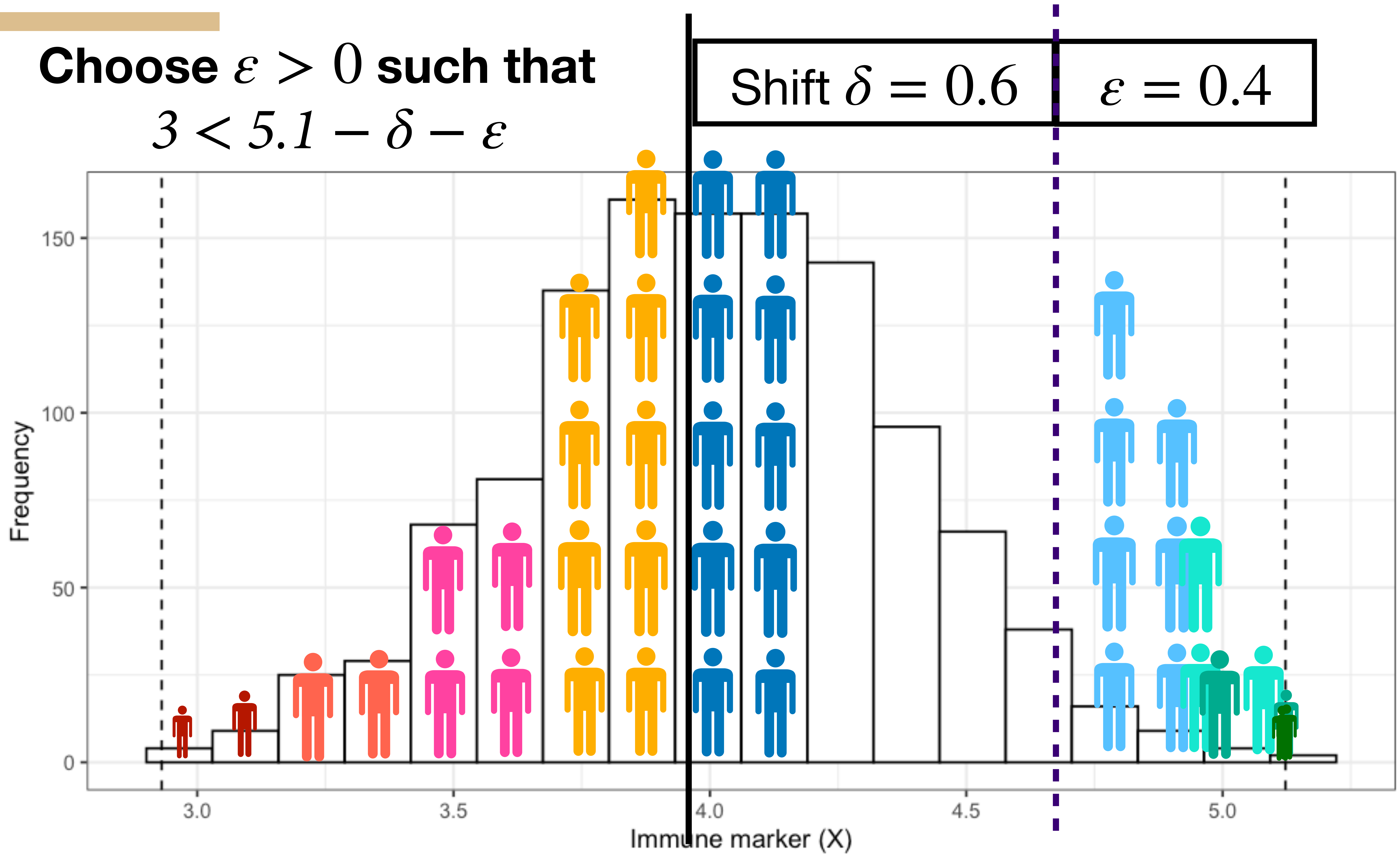
$$3 < 5.1 - \delta - \varepsilon$$



Strategy 2. Change the Policy

Choose $\varepsilon > 0$ such that

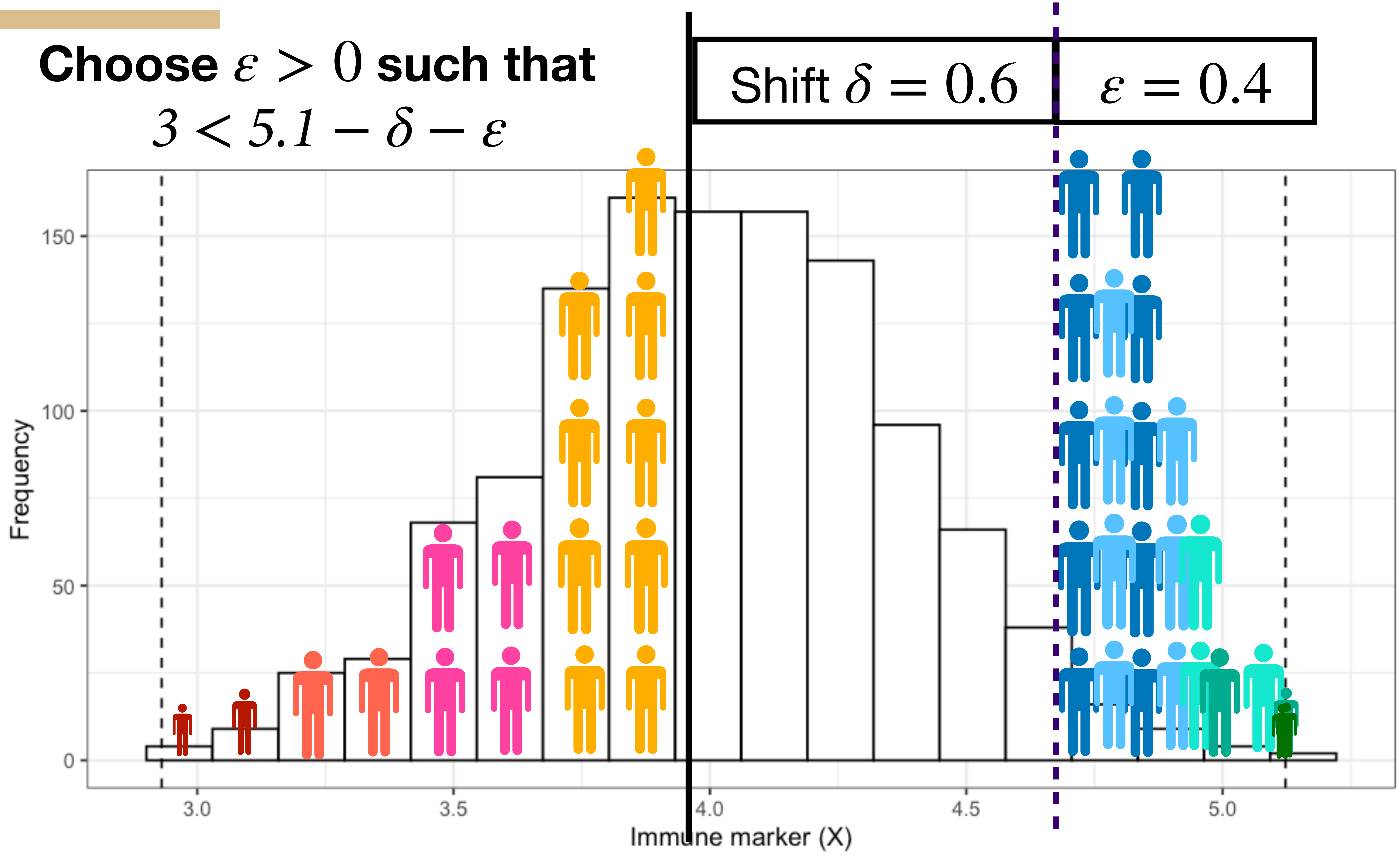
$$3 < 5.1 - \delta - \varepsilon$$



Strategy 2. Change the Policy

Choose $\varepsilon > 0$ such that

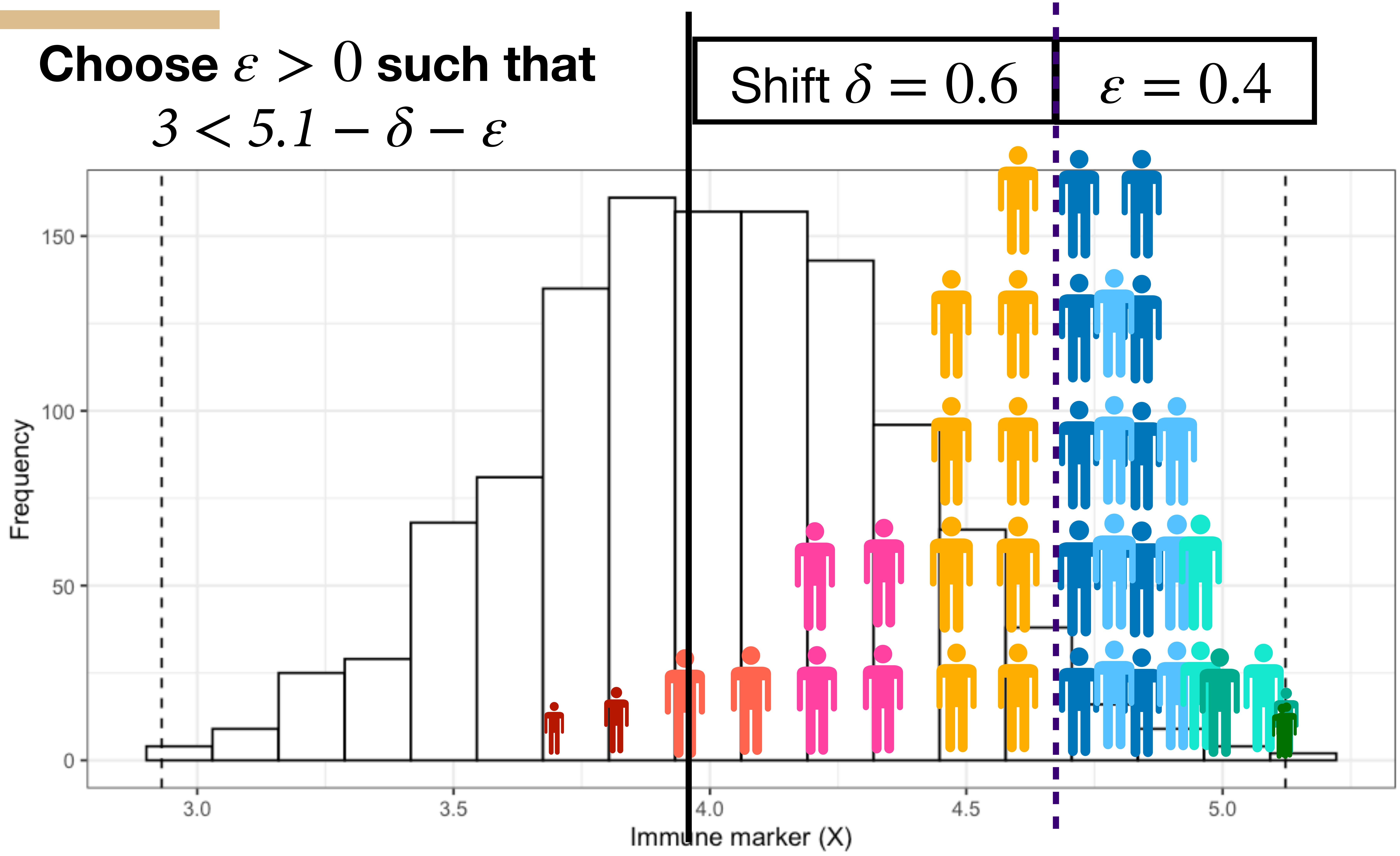
$$3 < 5.1 - \delta - \varepsilon$$



Strategy 2. Change the Policy

Choose $\varepsilon > 0$ such that

$$3 < 5.1 - \delta - \varepsilon$$



Modified treatment policy (MTP) mean

- Given a function $x \to q(x)$ we define the counterfactual $Y_i(q)$ to be the outcome in the i^{th} subject had exposure (NAb) level been $q(X_i)$ instead of X_i
- The MTP counterfactual mean is

$$\psi_0 \equiv E[Y(q)]$$

Modified treatment policy (MTP) mean

- Given a function $x \rightarrow q(x)$ we define the counterfactual $Y_i(q)$ to be the outcome in the i^{th} subject had exposure (NAb) level been $q(X_i)$ instead of X_i
- The MTP counterfactual mean is

$$\psi_0 \equiv E[Y(q)]$$

- Suppose we were prepared to assume the following weaker exchangeability condition held

$$\begin{aligned} E[Y(q) | X = x, L = l] &= E[Y | X = q(x), L = l] \\ &\equiv e_0(q(x), l) \end{aligned}$$

Modified treatment policy (MTP) mean

- Given a function $x \rightarrow q(x)$ we define the counterfactual $Y_i(q)$ to be the outcome in the i^{th} subject had exposure (NAb) level been $q(X_i)$ instead of X_i
- The MTP counterfactual mean is

$$\psi_0 \equiv E[Y(q)]$$

- Suppose we were prepared to assume the following weaker exchangeability condition held

$$\begin{aligned} E[Y(q)|X = x, L = l] &= E[Y|X = q(x), L = l] \\ &\equiv e_0(q(x), l) \end{aligned}$$

- Under this assumption,

$$\psi_0 \equiv E[e_0(q(X), L)]$$

Modified treatment policy (MTP) mean

- Given a function $x \to q(x)$ we define the counterfactual $Y_i(q)$ to be the outcome in the i^{th} subject had exposure (NAb) level been $q(X_i)$ instead of X_i
- The MTP counterfactual mean is

$$\psi_0 \equiv E[Y(q)]$$

- Suppose we were prepared to assume the following weaker exchangeability condition held

$$\begin{aligned} E[Y(q)|X = x, L = l] &= E[Y|X = q(x), L = l] \\ &\equiv e_0(q(x), l) \end{aligned}$$

- Under this assumption,

$$\psi_0 \equiv E[e_0(q(X), L)]$$

- **Challenge**: in the presence of unmeasured confounding by U,

$$\psi_0 \equiv E[e_0(q(X), L, U)]$$

but we do not measure U

Outline

- Biomarkers in vaccine research
- Negative control outcomes and treatments
- Modified treatment policies (MTPs): Identification in the absence of unmeasured confounding.
- **Modified treatment policies (MTPs): Proximal Causal Inference for MTP means with unmeasured confounding**
  - **Identification**
  - Estimation
- Simulation experiments
- Application to the analysis of a vaccine trial

Modified treatment policy (MTP) mean

- **Mathematical result:** Assume q is continuously differentiable (a.e.) and strictly monotone. Then,

$$
\begin{aligned}
E\left[e_0\left(q\left(X\right), L, U\right)\right] &= E\left[e_0\left(X, L, U\right)\ \alpha_0\left(X, L, U\right)\right] \\
&= E\left[Y\ \alpha_0\left(X, L, U\right)\right]
\end{aligned}
$$

Modified treatment policy (MTP) mean

- **Mathematical result:** Assume q is continuously differentiable (a.e.) and strictly monotone. Then,

$$
\begin{aligned}
E\left[e_0\left(q\left(X\right), L, U\right)\right] &= E\left[e_0\left(X, L, U\right)\ \alpha_0\left(X, L, U\right)\right] \\
&= E\left[Y\ \alpha_0\left(X, L, U\right)\right]
\end{aligned}
$$

- where

$$
\alpha_0\left(x, l, u\right) = I\left[x \in \mathrm{Image}(q)\right]\ \frac{dq^{-1}\left(x\right)}{dx}\ \frac{p_{X|L,U}\left(q^{-1}\left(x\right)|l, u\right)}{p_{X|L,U}\left(x|l, u\right)}
$$

Modified treatment policy (MTP) mean

- **Mathematical result:** Assume q is continuously differentiable (a.e.) and strictly monotone. Then,

$$\begin{aligned} E\left[e_0\left(q\left(X\right), L, U\right)\right] &= E\left[e_0\left(X, L, U\right)\ \alpha_0\left(X, L, U\right)\right] \\ &= E\left[Y\ \alpha_0\left(X, L, U\right)\right] \end{aligned}$$

- where

$$\alpha_0\left(x, l, u\right) = I\left[x \in \mathrm{Image}(q)\right]\ \frac{dq^{-1}\left(x\right)}{dx}\ \frac{p_{X|L,U}\left(q^{-1}\left(x\right)|l, u\right)}{p_{X|L,U}\left(x|l, u\right)}$$

- **Remark:** This holds because the map

$$e \rightarrow E\left[e\left(q\left(X\right), L, U\right)\right]$$

is bounded and linear and α_0 is its **Riesz representer**, i.e.

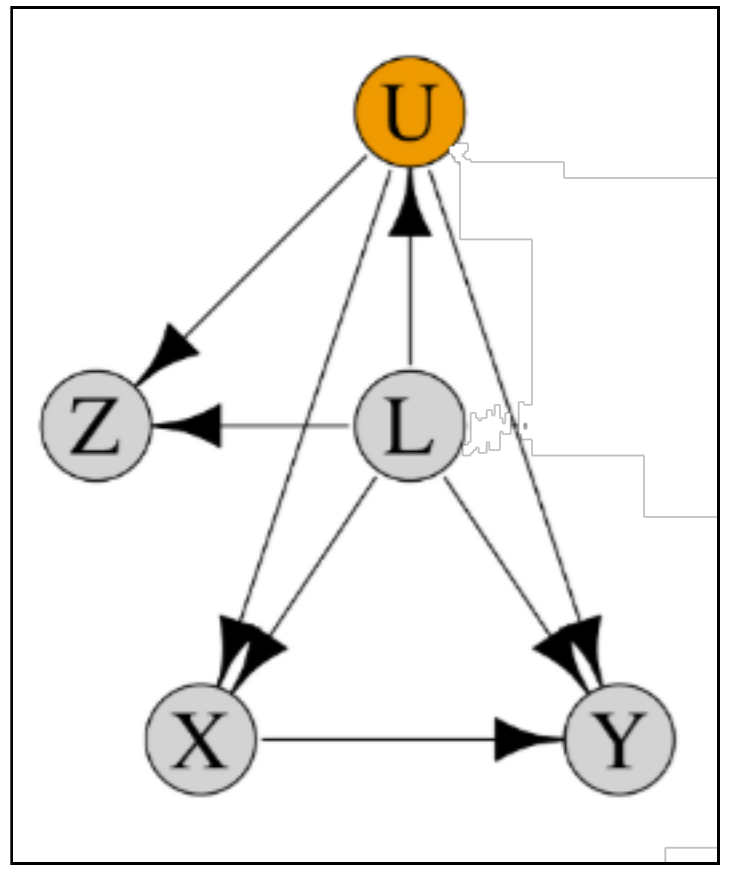
$$E\left[e\left(q\left(X\right), L, U\right)\right] = E\left[e\left(X, L, U\right)\ \alpha_0\left(X, L, U\right)\right] \text{ for all } e$$

Negative control trx and latent trx bridge

1. $Z \perp Y \mid X, L, U$ true if **Z is a Negative Control Trx**

2. **Latent trx bridge:** any g_0 that satisfies

$$\alpha_0(X, L, U) = \mathbb{E}\left[g_0(X, L, Z) \mid X, L, U\right]$$

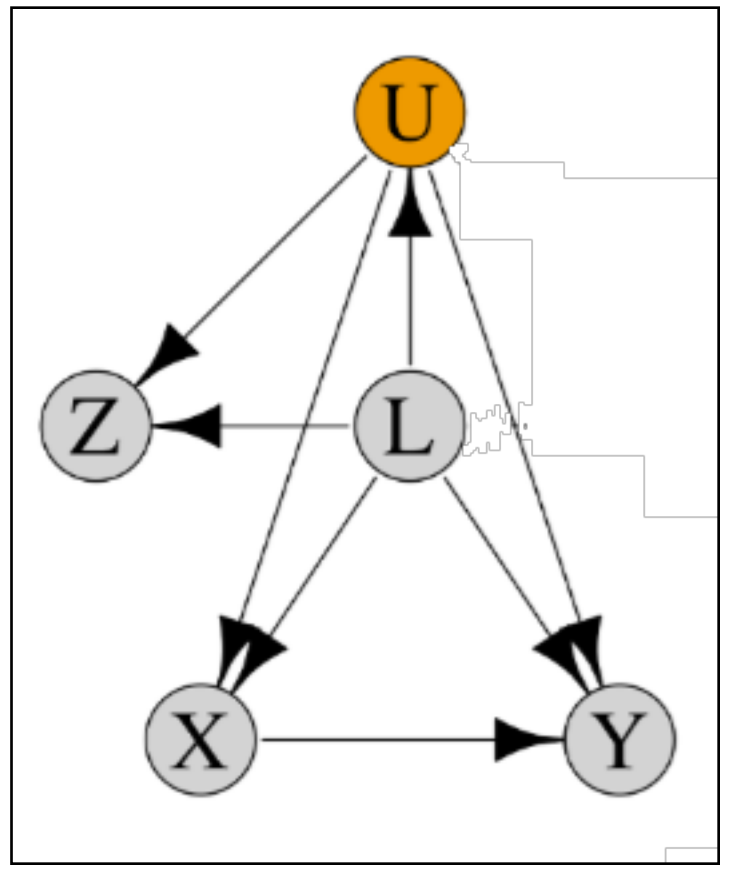


Negative control trx and latent trx bridge

1. $Z \perp Y \mid X, L, U$ true if **Z is a Negative Control Trx**

2. **Latent trx bridge function :** any g_0 that satisfies

$$\alpha_0(X, L, U) = \mathbb{E}\left[g_0(X, L, Z) \mid X, L, U\right]$$

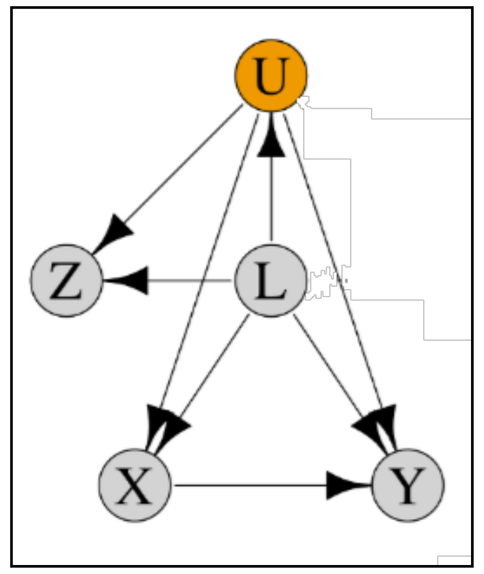


The existence of g_0 is essentially tantamount to the requirement that Z is a strong proxy of U

For instance: For g_0 to exist the following completeness condition suffices under regularity conditions:

$$\mathbb{E}\left[\eta(X, L, U) \mid X, L, Z\right] = \text{c} \quad \Rightarrow \quad \eta(X, L, U) = \text{c a.e.}$$

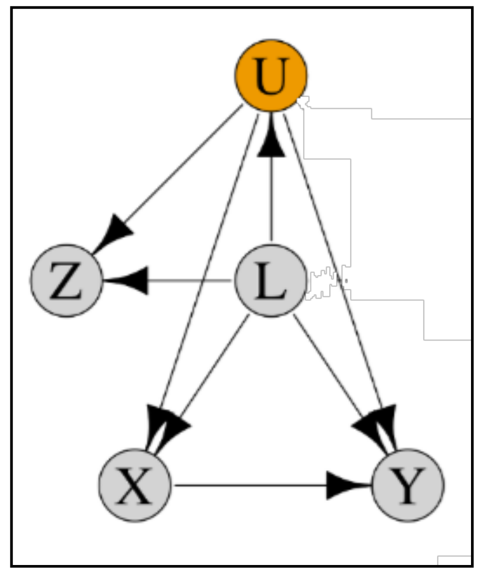
Negative control trx and latent trx bridge



If Z is a negative control trx and g_0 is a latent trx bridge function then

$$
\begin{aligned}
\mathbb{E}\left[Y\, \alpha_0(X, L, U)\right] &= \mathbb{E}\left[\mathbb{E}\left[Y \,|\, X, L, U\right] \mathbb{E}\left[g_0(X, L, Z) \,|\, X, L, U\right]\right] && \text{(by } g_0 \text{ trx br fcn)} \\
&= \mathbb{E}\left[\mathbb{E}\left[Y\, g_0(X, L, Z) \,|\, X, L, U\right]\right] && \text{(by } Z \text{ neg. trx)} \\
&= \mathbb{E}\left[Y\, g_0(X, L, Z)\right]
\end{aligned}
$$

Negative control trx and latent trx bridge



If Z is a negative control trx and g_0 is a latent trx bridge function then

$$\mathbb{E}\left[Y\,\alpha_0(X,L,U)\right] = \mathbb{E}\left[\mathbb{E}\left[Y \mid X,L,U\right]\mathbb{E}\left[g_0(X,L,Z) \mid X,L,U\right]\right] \quad \text{(by } g_0 \text{ trx br fcn)}$$

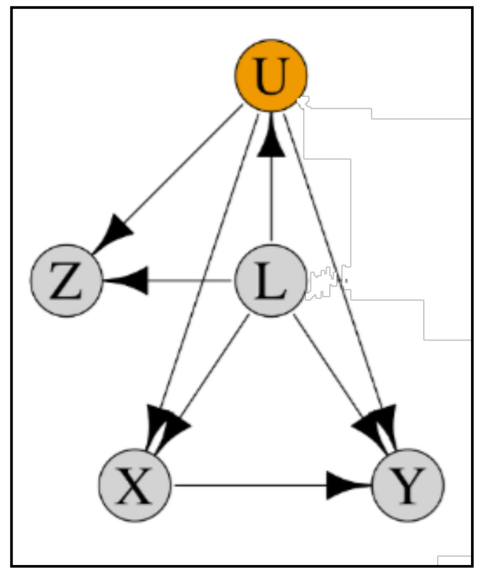
$$= \mathbb{E}\left[\mathbb{E}\left[Y\,g_0(X,L,Z) \mid X,L,U\right]\right] \quad \text{(by } Z \text{ neg. trx)}$$

$$= \mathbb{E}\left[Y\,g_0(X,L,Z)\right]$$

So under assumptions 1 and 2

$$\psi_0 = \mathbb{E}\left[Y\,g_0(X,L,Z)\right]$$

Negative control trx and latent trx bridge



If Z is a negative control trx and g_0 is a latent trx bridge function then

$$\mathbb{E}\left[Y\,\alpha_0(X,L,U)\right] = \mathbb{E}\left[\mathbb{E}\left[Y \mid X,L,U\right]\mathbb{E}\left[g_0(X,L,Z) \mid X,L,U\right]\right] \quad \text{(by } g_0 \text{ trx br fcn)}$$

$$= \mathbb{E}\left[\mathbb{E}\left[Y\,g_0(X,L,Z) \mid X,L,U\right]\right] \quad \text{(by } Z \text{ neg. trx)}$$

$$= \mathbb{E}\left[Y\,g_0(X,L,Z)\right]$$

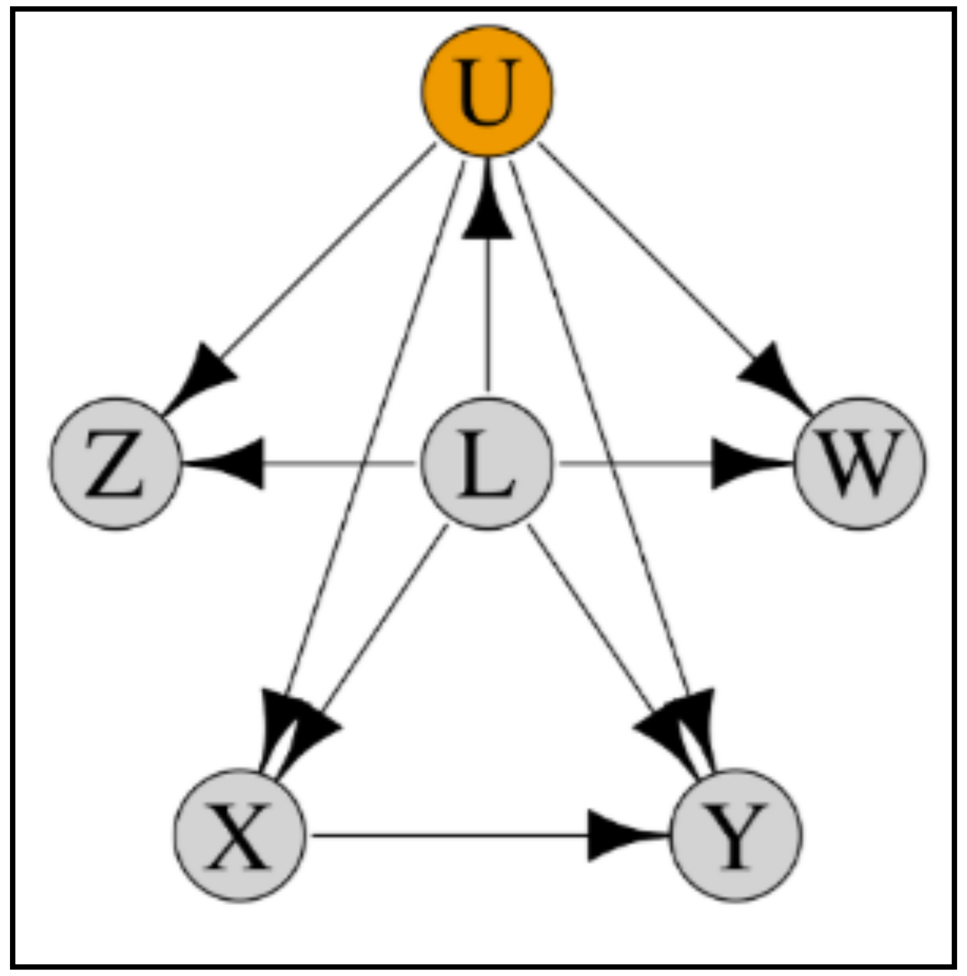
So under assumptions 1 and 2

$$\psi_0 = \mathbb{E}\left[Y\,g_0(X,L,Z)\right]$$

However… g_0 is not identified because it solves an integral equation that depends on the unknown law of $Z \mid X, L, U$

Negative control outcome and latent outcome bridge

3. $(X, Z) \perp W \mid L, U$ true if **W is a Negative Control Outcome**



4. **Latent outcome bridge fcn**: any h_0 that satisfies

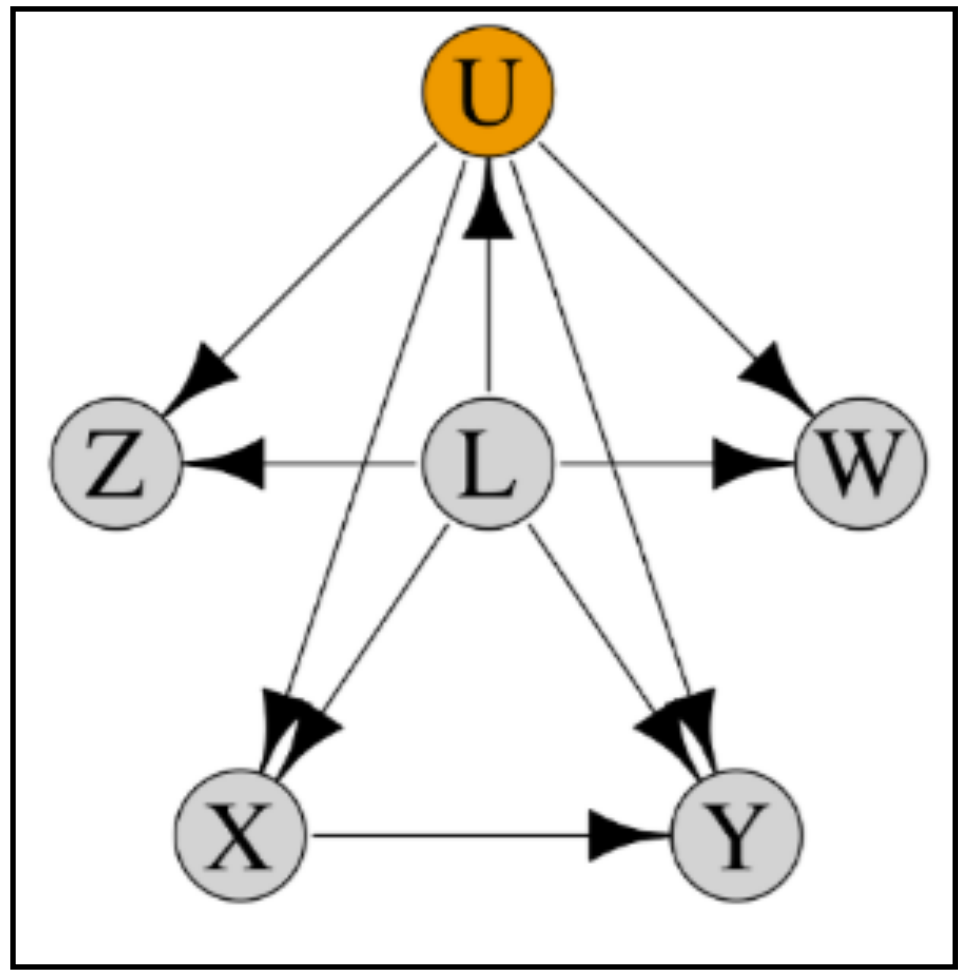
$$\mathbb{E}\left[Y \mid X, L, U\right] = \mathbb{E}\left[h_0(X, L, W) \mid X, L, U\right]$$

The existence of h0 is essentially tantamount to the requirement that W is a strong proxy of U

Negative control outcome and latent outcome bridge

Result: If W is a negative control outcome and h_0 is a latent outcome bridge function it can be shown that

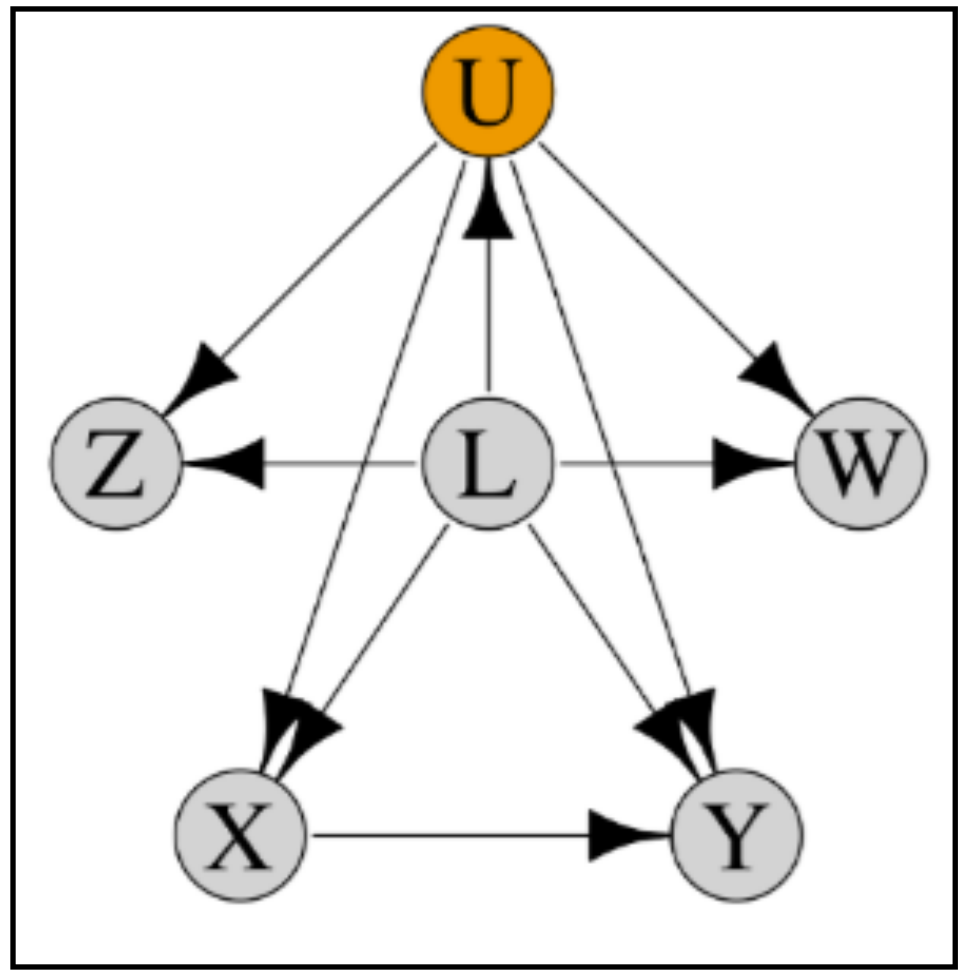
$$\mathbb{E}\left[e_0(q(X), L, U)\right] = \mathbb{E}\left[h_0(q(X), L, W)\right]$$



Negative control outcome and latent outcome bridge

Result: If W is a negative control outcome and h_0 is a latent outcome bridge function it can be shown that

$$\mathbb{E}\left[e_0(q(X), L, U)\right] = \mathbb{E}\left[h_0(q(X), L, W)\right]$$



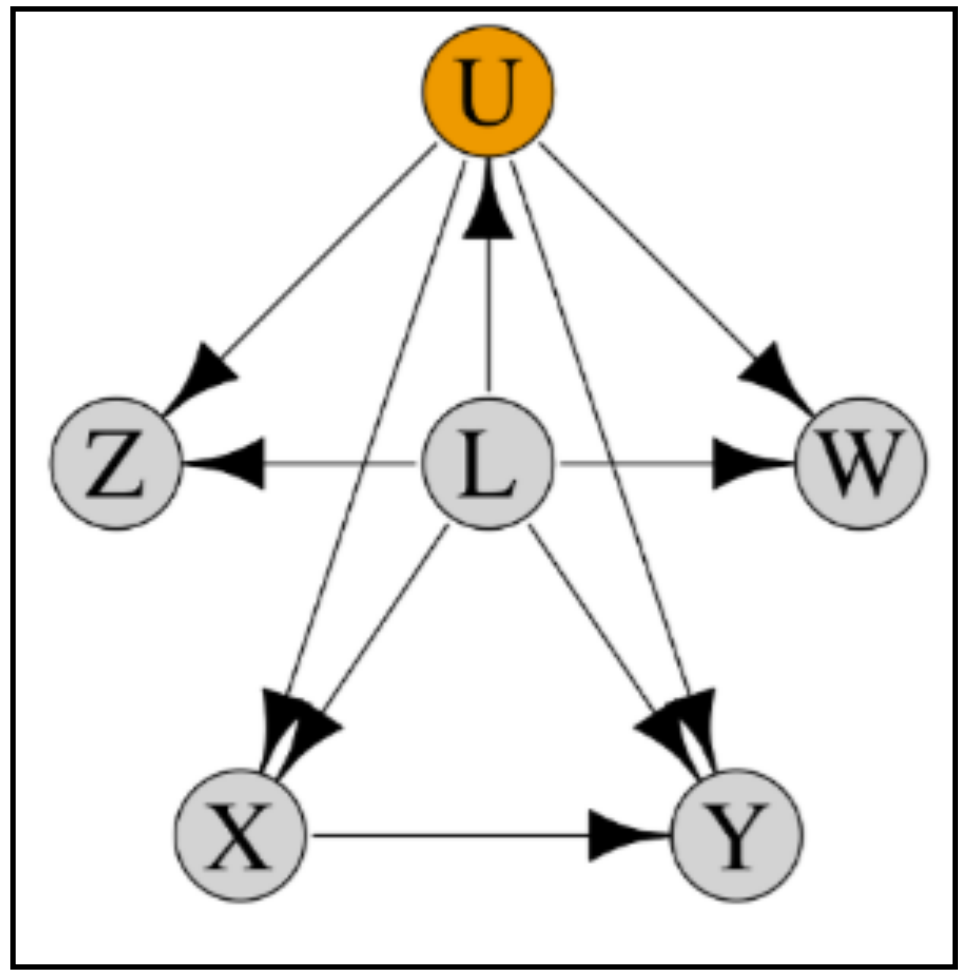
So under assumptions $3 - 4$

$$\psi_0 = \mathbb{E}\left[h_0(q(X), L, W)\right]$$

Negative control outcome and latent outcome bridge

Result: If W is a negative control outcome and h_0 is a latent outcome bridge function it can be shown that

$$\mathbb{E}\left[e_0(q(X), L, U)\right] = \mathbb{E}\left[h_0(q(X), L, W)\right]$$



So under assumptions $3-4$

$$\psi_0 = \mathbb{E}\left[h_0(q(X), L, W)\right]$$

However… h_0 is not identified because it solves an integral equation that depends on the unknown law of $W \mid X, L, U$

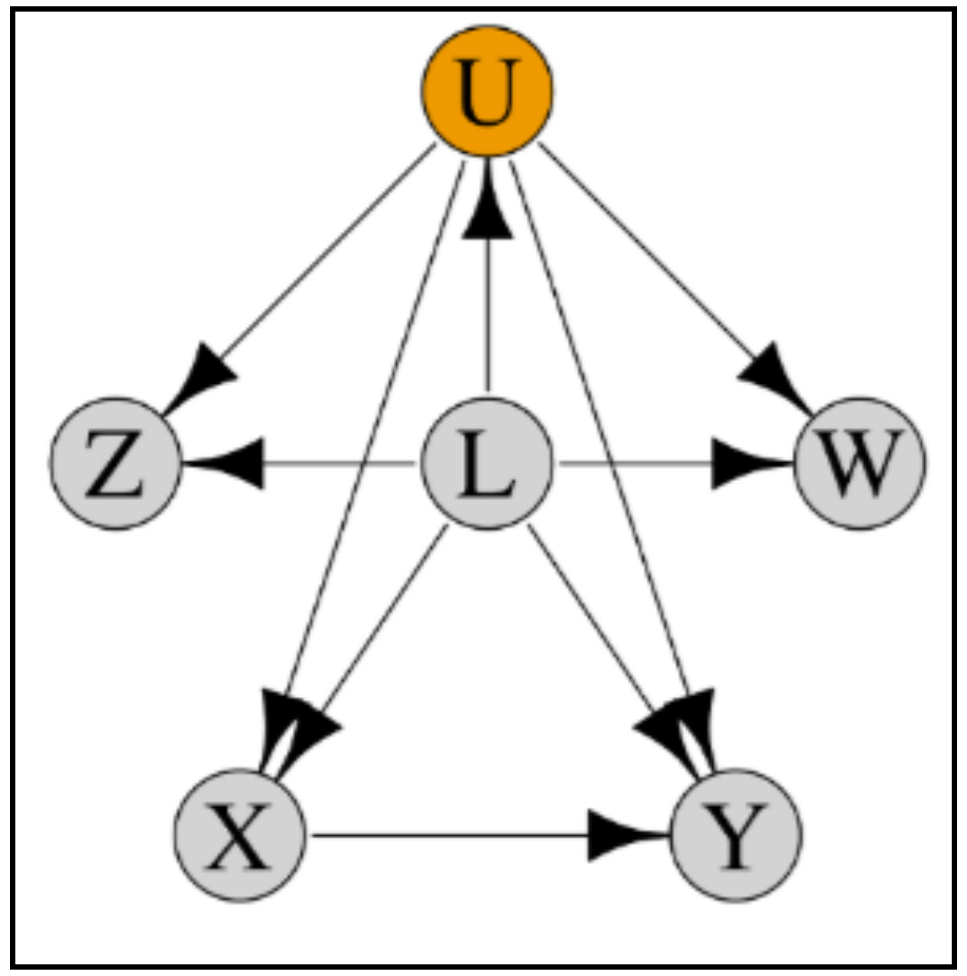
Observed trx and outcome bridge functions

Observed trx bridge function: any $g^{\dagger}$ that solves

$$\alpha_0(X, L, W) = \mathbb{E}\left[g(X, L, Z) \,|\, X, L, W\right].$$

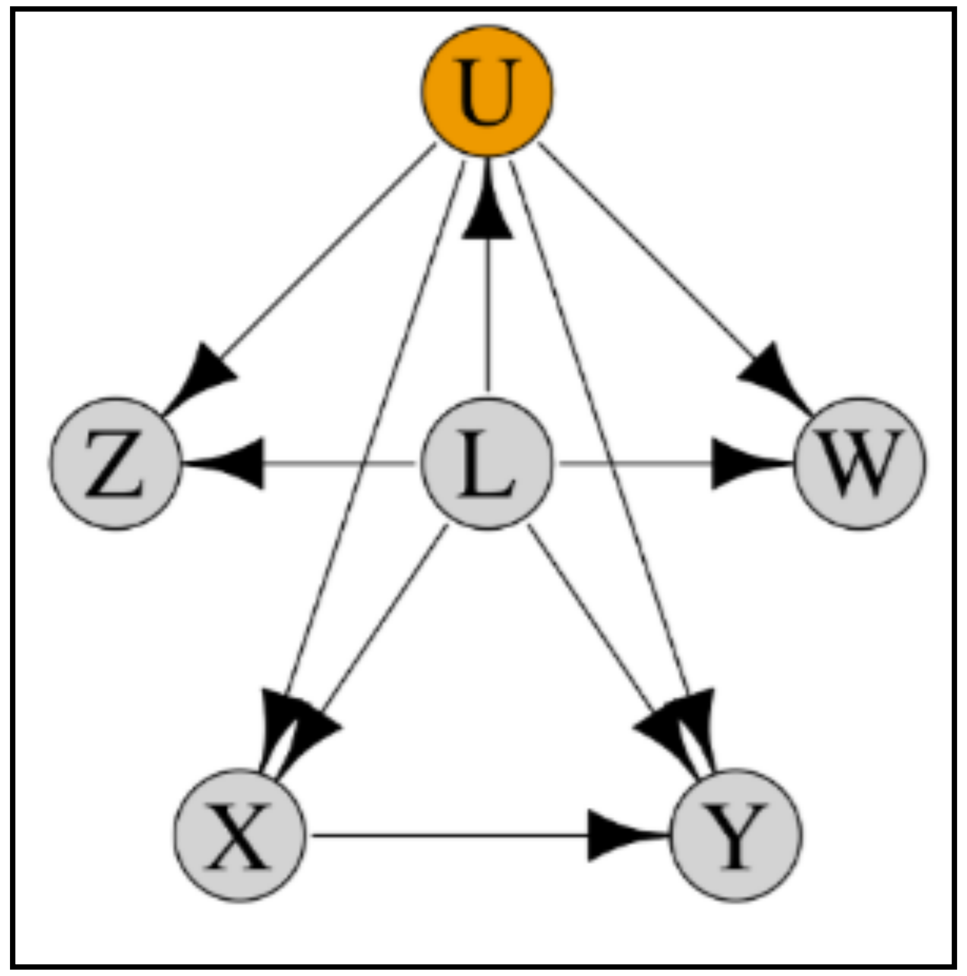
Observed outcome bridge function: any $h^{\dagger}$ that solves

$$\mathbb{E}\left[Y \,|\, X, L, Z\right] = \mathbb{E}\left[h(X, L, W) \,|\, X, L, Z\right].$$



Unlike their latent counterparts, the observed bridge functions depend on the law of the observable variables.

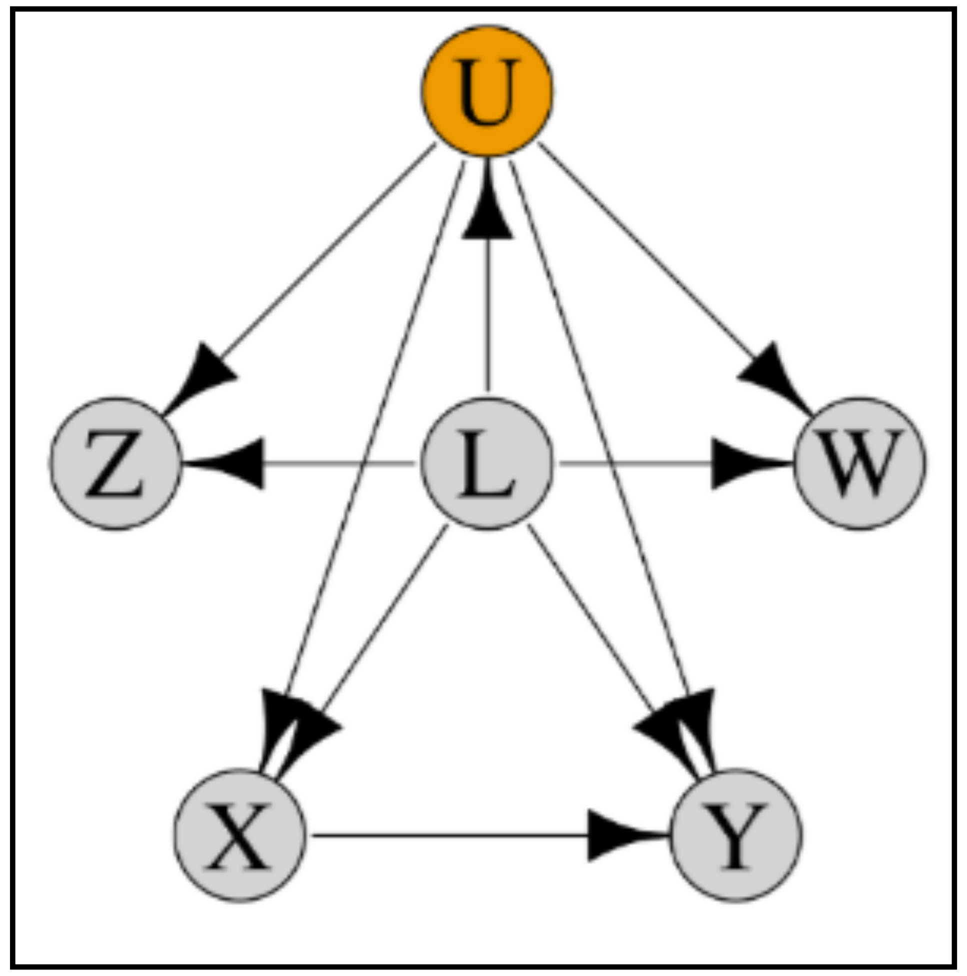
Observed trx and outcome bridge functions



<u>Result</u>: If $h^{\dagger}$ is an observed outcome bridge fcn and $g^{\dagger}$ is an observed trx bridge fcn then

$$\mathbb{E}\left[h^{\dagger}(q(X), L, W)\right] = \mathbb{E}\left[Y\, g^{\dagger}(X, L, Z)\right]$$

Observed trx and outcome bridge functions



Result: If $h^\dagger$ is an observed outcome bridge fcn and $g^\dagger$ is an observed trx bridge fcn then

$$\mathbb{E}\left[h^\dagger(q(X), L, W)\right] = \mathbb{E}\left[Y\, g^\dagger(X, L, Z)\right]$$

Proof relies on the fact that $\alpha_0(X, L, W)$ is the Riesz-representer of the bounded linear map

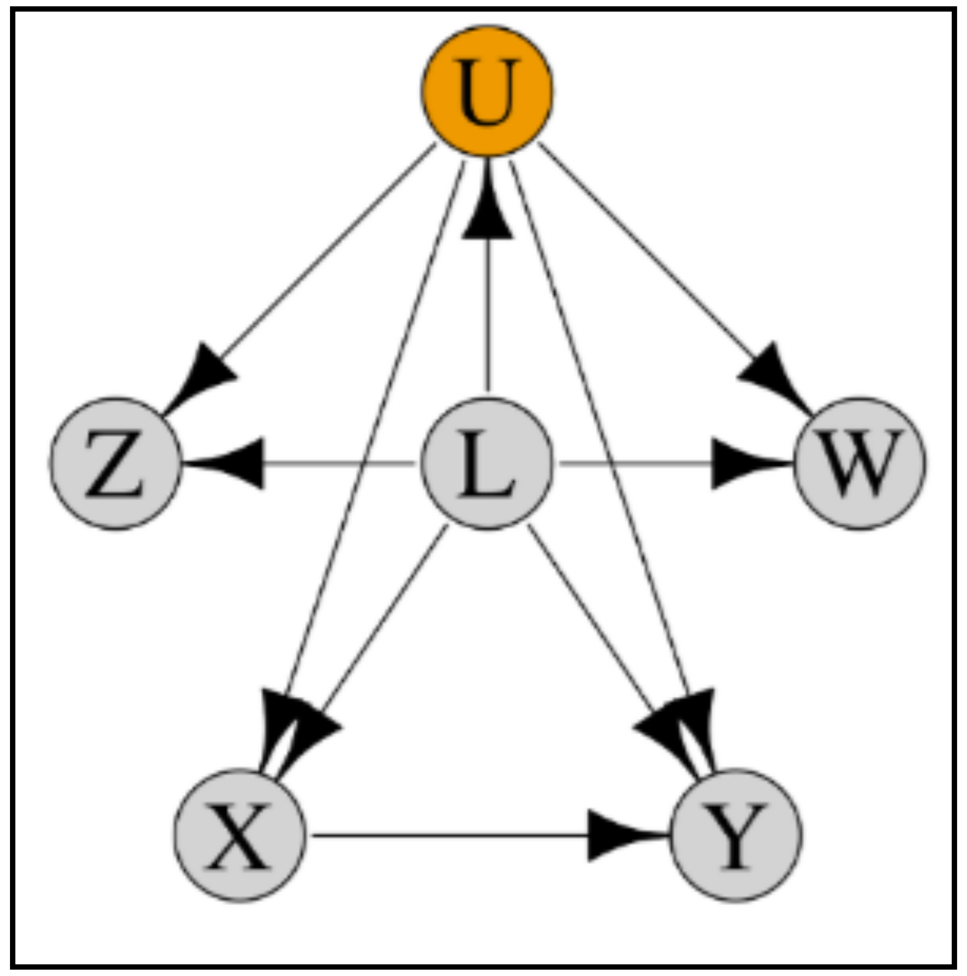
$$h \mapsto \mathbb{E}\left[h(q(X), L, W)\right]$$

Any latent bridge fcn is an observed bridge function

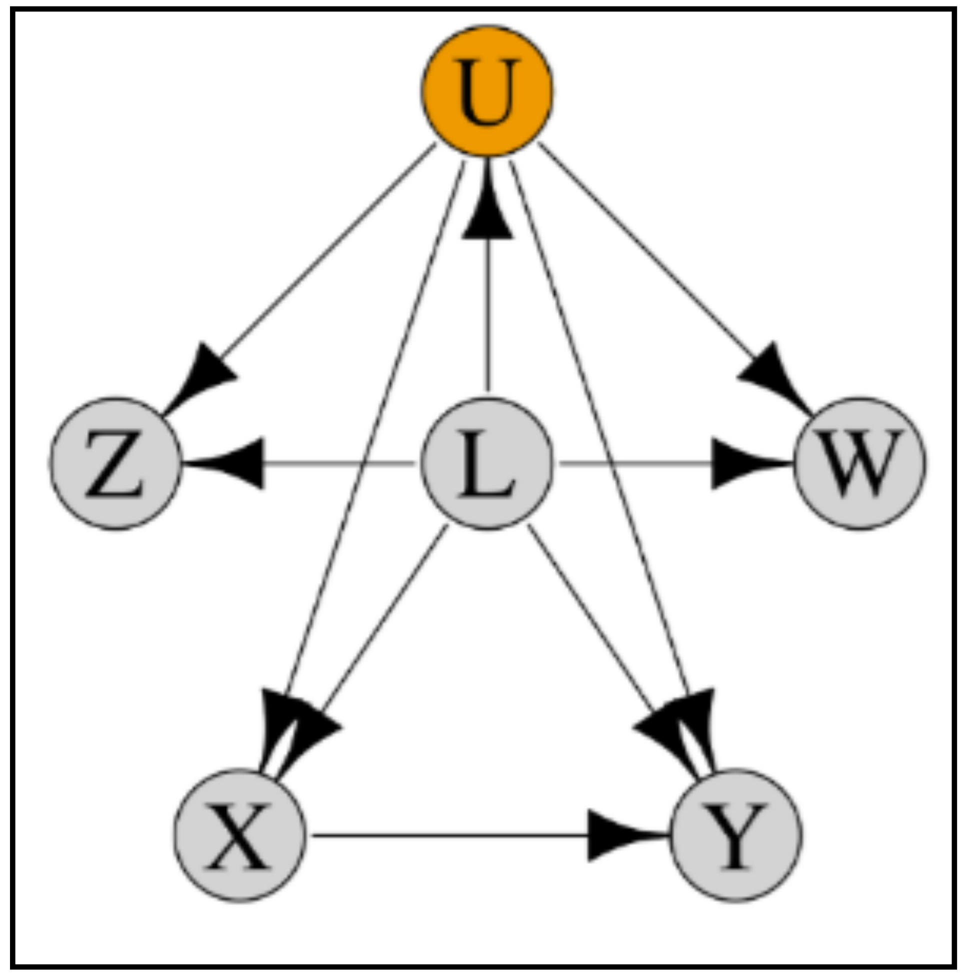
Theorem 1: If W is a NC outcome and Z is a NC trx then

(a) h_0 latent outcome br fcn $\Rightarrow$ h_0 observed outcome br fcn

(b) g_0 latent trx br fcn. $\Rightarrow$ g_0 observed trx br fcn



Any latent bridge fcn is an observed bridge function



Theorem 1: If W is a NC outcome and Z is a NC trx then

(a) h_0 latent outcome br fcn $\Rightarrow$ h_0 observed outcome br fcn

(b) g_0 latent trx br fcn. $\Rightarrow$ g_0 observed trx br fcn

Proof of (a)

$$\mathbb{E}\left[Y \,|\, X, L, U\right] \overset{h_0 \text{ latent out. br fcn}}{=} \mathbb{E}\left[h_0(X, L, W) \,|\, X, L, U\right]$$

|| (Z n.c. trx) || (W n.c. outcome)

$$\mathbb{E}\left[Y \,|\, X, L, Z, U\right] = \mathbb{E}\left[h_0(X, L, W) \,|\, X, L, Z, U\right]$$

$$\Rightarrow \quad \mathbb{E}\left[Y \,|\, X, L, Z\right] = \mathbb{E}\left[h_0(X, L, W) \,|\, X, L, Z\right]$$

Identification[3]

Theorem 2: If a latent outcome bridge fcn and an observed trx bridge fcn exist, then

$$\psi_0 = \mathbb{E}\left[h^\dagger(q(X), L, W)\right] = \mathbb{E}\left[Y\, g^\dagger(X, L, Z)\right]$$

for any $h^\dagger$ and $g^\dagger$ observed outcome and trx bridge fcns respectively

Proof: Let h_0 be a latent outcome bridge fcn, and $g^\dagger$ and $h^\dagger$ be arbitrary observed trx and outcome bridge fcns. Then

$$\psi_0 = \mathbb{E}\left[h_0(q(X), L, W)\right] \quad \text{under 3-4}$$

$$= \mathbb{E}\left[Y\, g^\dagger(X, L, Z)\right] \quad \text{by the thm in the previous slide since } h_0 \text{ is an } \textbf{observed} \text{ outcome bridge function}$$

$$= \mathbb{E}\left[h^\dagger(q(X), L, W)\right] \quad \text{by the thm in the previous slide}$$

[3]Analogues theorems were earlier derived by Miao et al. (2018), Tchetgen Tchetgen et al. (2020), Cui et al. (2023), and Kallus et al. (2020) for other causal constrasts.

Double Robust Representation

Corollary: If one of the latent bridge eqns and the remaining observed bridge eqn have solutions, then

$$\psi_0 = \mathbb{E}\left[\phi\left(O; h^{\dagger}, g^{\dagger}\right)\right],$$

where $O = (X, L, Z, W, Y)$, for any h and g

$$\phi(O; h, g) := h\left[q(X), L, W\right] + g(X, L, Z)\{Y - h(X, L, W)\}$$

and where one of h† or g† (but not necessarily both) solve the corresponding observed equation

Outline

- Biomarkers in vaccine research
- Negative control outcomes and treatments
- Modified treatment policies (MTPs): Identification in the absence of unmeasured confounding.
- **Modified treatment policies (MTPs): Proximal Causal Inference for MTP means with unmeasured confounding**
  - Identification
  - **Estimation**
- Simulation experiments
- Application to the analysis of a vaccine trial

Possible Estimators of ψ_0

Possible Estimators of ψ_0

Our identification result suggests three estimators:

$$\hat{\psi}^{OR} = \mathbb{E}_n \left\{ \hat{h} \left[q(X), L, W \right] \right\},$$

$$\hat{\psi}^{IPW} = \mathbb{E}_n \left[Y \cdot \hat{g}(X, L, Z) \right],$$

$$\hat{\psi}^{DR} = \mathbb{E}_n \left[\phi(O; \hat{h}, \hat{g}) \right]$$

Possible Estimators of ψ_0

Our identification result suggests three estimators:

$$\hat{\psi}^{OR} = \mathbb{E}_n \left\{ \hat{h} \left[q(X), L, W \right] \right\},$$

$$\hat{\psi}^{IPW} = \mathbb{E}_n \left[Y \cdot \hat{g}(X, L, Z) \right],$$

$$\boxed{\hat{\psi}^{DR} = \mathbb{E}_n \left[\phi(O; \hat{h}, \hat{g}) \right]}$$

GMM-parametric estimation of bridge functions

Outcome bridge

$$\mathbb{E}\left[Y \,|\, X, L, Z\right] = \mathbb{E}\left[h(X, L, W) \,|\, X, L, Z\right]$$

GMM-parametric estimation of bridge functions

Outcome bridge
$\mathbb{E}\left[Y \mid X, L, Z\right] = \mathbb{E}\left[h(X, L, W) \mid X, L, Z\right]$
Then $\forall g$ $\mathbb{E}\left[g(X, L, Z)\left\{Y - h^{\dagger}(X, L, W)\right\}\right] = 0$

GMM-parametric estimation of bridge functions

Outcome bridge

$$\mathbb{E}\left[Y \mid X, L, Z\right] = \mathbb{E}\left[h(X, L, W) \mid X, L, Z\right]$$

Then $\forall g$

$$\mathbb{E}\left[g(X, L, Z)\left\{Y - h^{\dagger}(X, L, W)\right\}\right] = 0$$

So, compute

$$\hat{\tau} = \arg\min_{\tau} \sum_{j=1}^{J} \left\{\mathbb{E}_n\left[g_j(X, L, Z)\{Y - h(X, L, W; \tau)\}\right]\right\}^2$$

where

$h\left(\cdots; \tau\right)$ is a parametric model for the, assumed unique, out. br fcn

GMM-parametric estimation of bridge functions

Treatment bridge

$$\alpha_0(X, L, W) = \mathbb{E}\left[g(X, L, Z) \,|\, X, L, W\right]$$

GMM-parametric estimation of bridge functions

Treatment bridge

$$\alpha_0(X, L, W) = \mathbb{E}\left[g(X, L, Z) \,|\, X, L, W\right]$$

Then $\forall h$

$$\begin{aligned}\mathbb{E}\left[h(q(X), L, W)\right] &= \mathbb{E}\left[h(X, L, W)\ \alpha_0(X, L, W)\right] \\ &= \mathbb{E}\left[h(X, L, W)\ g^{\dagger}(X, L, Z)\right]\end{aligned}$$

GMM-parametric estimation of bridge functions

Treatment bridge

$$\alpha_0(X, L, W) = \mathbb{E}\left[g(X, L, Z) \,|\, X, L, W\right]$$

Then $\forall h$

$$\mathbb{E}\left[h(q(X), L, W)\right] = \mathbb{E}\left[h(X, L, W)\ \alpha_0(X, L, W)\right]$$
$$= \mathbb{E}\left[h(X, L, W)\ g^{\dagger}(X, L, Z)\right]$$

So, compute

$$\hat{\delta} = \arg\min_{\delta} \sum_{j=1}^{J} \left\{ \mathbb{E}_n \left[h_j(q(X), L, W) - h_j(X, L, W) g(X, L, Z; \delta) \right] \right\}^2$$

where $g\,(\cdots;\delta)$ is a parametric model for the, assumed unique, trx bridge function. Note that we need not know nor estimate $\alpha_0(X, L, W)$!!!

Parametric estimation

The estimator

$$\hat{\psi}_{DR} = \mathbb{E}_n \left[\phi \left(O; h\left(X, L, W; \hat{\tau}\right), g(X, L, Z; \hat{\delta}) \right) \right]$$

is **Doubly Robust**: it is consistent for ψ_0 and asymptotically normal if one of the models is correct.

Debiased Machine Learning Estimation

Cross-Fitted Debiased Machine-Learning Estimator

Sample

$O_1 = (X_1, L_1, Z_1, W_1, Y_1)$

$\vdots$

$O_n = (X_n, L_n, Z_n, W_n, Y_n)$

K folds

$I_1 = \{O_1, \ldots, O_m\}$

$\vdots$

$I_{K-1} = \left\{O_{(K-2)m+1}, \ldots, O_{(K-1)m}\right\}$

$I_K = \left\{O_{(K-1)m+1}, \ldots, O_n\right\}$

Cross-Fitted Debiased Machine-Learning Estimator

Sample

$$O_1 = (X_1, L_1, Z_1, W_1, Y_1)$$

$$\vdots$$

$$O_n = (X_n, L_n, Z_n, W_n, Y_n)$$

K folds

$$I_1 = \{O_1, \ldots, O_m\}$$

$$\vdots$$

$$I_{K-1} = \left\{O_{(K-2)m+1}, \ldots, O_{(K-1)m}\right\}$$

$$I_K = \left\{O_{(K-1)m+1}, \ldots, O_n\right\}$$

Obtain estimators $\hat{h}^{(-K)}$ and $\hat{g}^{(-K)}$

Cross-Fitted Debiased Machine-Learning Estimator

Sample

$$O_1 = (X_1, L_1, Z_1, W_1, Y_1)$$

$$\vdots$$

$$O_n = (X_n, L_n, Z_n, W_n, Y_n)$$

K folds

$$I_1 = \{O_1, \ldots, O_m\}$$

$$\vdots$$

$$I_{K-1} = \left\{O_{(K-2)m+1}, \ldots, O_{(K-1)m}\right\}$$

$$I_K = \left\{O_{(K-1)m+1}, \ldots, O_n\right\}$$

Obtain estimators $\hat{h}^{(-K)}$ and $\hat{g}^{(-K)}$

$$\hat{\psi}^K = \mathbb{E}_n^K\left[\phi\left(O; \hat{h}^{(-K)}, \hat{g}^{(-K)}\right)\right]$$

Cross-Fitted Debiased Machine-Learning Estimator

Sample

$$O_1 = \left(X_1, L_1, Z_1, W_1, Y_1\right)$$

$$\vdots$$

$$O_n = \left(X_n, L_n, Z_n, W_n, Y_n\right)$$

K folds

$$I_1 = \left\{O_1, \ldots, O_m\right\}$$

$$\vdots$$

$$I_{K-1} = \left\{O_{(K-2)m+1}, \ldots, O_{(K-1)m}\right\}$$

$$I_K = \left\{O_{(K-1)m+1}, \ldots, O_n\right\}$$

Obtain estimators $\hat{h}^{(-K)}$ and $\hat{g}^{(-K)}$

$$\hat{\psi}^K = \mathbb{E}_n^K\left[\phi\left(O; \hat{h}^{(-K)}, \hat{g}^{(-K)}\right)\right]$$

The final estimator is

$$\hat{\psi}_{CF}^{DR} = \frac{1}{K}\sum_{k=1}^{K}\hat{\psi}^k$$

Asymptotic Analysis

Let

$$\left(\mathcal{T} h\right)(X, L, Z) = \mathbb{E}\left[h(X, L, W) \mid X, L, Z\right] \quad \text{and} \quad \left(\mathcal{T}^{*} g\right)(X, L, W) = \mathbb{E}\left[g(X, L, Z) \mid X, L, W\right]$$

Theorem: if $|Y| \leq B$, $|\alpha_0(X, L, W)| \leq B$ for some B, and

$$\min\left\{ \left\| \hat{h}^{(-k)} - h^{\dagger} \right\|_2 \times \left\| \mathcal{T}^{*}\left(\hat{g}^{(-k)} - g^{\dagger}\right) \right\|_2, \left\| \mathcal{T}\left(\hat{h}^{(-k)} - h^{\dagger}\right) \right\|_2 \times \left\| \hat{g}^{(-k)} - g^{\dagger} \right\|_2 \right\} = o_p\left(n^{-1/2}\right),$$

Asymptotic Analysis

Let

$$\left(\mathcal{T}h\right)(X,L,Z) = \mathbb{E}\left[h(X,L,W)\,|\,X,L,Z\right] \quad \text{and} \quad \left(\mathcal{T}^*g\right)(X,L,W) = \mathbb{E}\left[g(X,L,Z)\,|\,X,L,W\right]$$

Theorem: If $|Y| \leq B$, $|\alpha_0(X, L, W)| \leq B$ for some B, and

$$\min\left\{ \left\| \hat{h}^{(-k)} - h^\dagger \right\|_2 \times \left\| \mathcal{T}^*\left(\hat{g}^{(-k)} - g^\dagger\right) \right\|_2, \left\| \mathcal{T}\left(\hat{h}^{(-k)} - h^\dagger\right) \right\|_2 \times \left\| \hat{g}^{(-k)} - g^\dagger \right\|_2 \right\} = o_p\left(n^{-1/2}\right),$$

where $h^\dagger$ and $g^\dagger$ are observed outcome and treatment bridge functions, then

$$\sqrt{n}\left(\psi_{CF}^{DR} - \psi_0\right) \longrightarrow \mathcal{N}(0,\tau^2),$$

where $\tau^2 = \mathbb{E}\left\{\left[\phi(O; h^\dagger, g^\dagger) - \psi_0\right]^2\right\}$.

Recall GMM-parametric estimation of bridge functions

Treatment bridge

$$\alpha_0(X, L, W) = \mathbb{E}\left[g(X, L, Z) \,|\, X, L, W\right]$$

Then $\forall h$

$$\mathbb{E}\left[h(q(X), L, W)\right] = \mathbb{E}\left[h(X, L, W)\ \alpha_0(X, L, W)\right]$$
$$= \mathbb{E}\left[h(X, L, W)\ g^{\dagger}(X, L, Z)\right]$$

So, compute

$$\hat{\delta} = \arg\min_{\delta} \sum_{j=1}^{J} \left\{ \mathbb{E}_n \left[h_j(q(X), L, W) - h_j(X, L, W) g(X, L, Z; \delta) \right] \right\}^2$$

where $g\left(\cdots; \delta\right)$ is a parametric model for the, assumed unique, trx bridge function. Note that we need not know nor estimate $\alpha_0(X, L, W)$!!!

Non-parametric estimation of trx bridge function

Lemma: if $\mathscr{H}$ is a **linear subspace of** $\mathscr{L}^2\left(P_{X,L,W}\right)$ and $h_1, h_2, \ldots$ is an orthonormal basis of $\mathscr{H}$ then

$$\sum_{j=1}^{\infty}\left\{\mathbb{E}\left[h_j\left(q(X), L, W\right) - h_j(X, L, W)g(X, L, Z)\right]\right\}^2 = \max_{h\in\mathscr{H}} \mathbb{E}\left[h\left(q(X), L, W\right) - h(X, L, W)g(X, L, Z) - h(X, L, W)^2\right]$$

Non-parametric estimation of trx bridge function

Lemma: if $\mathscr{H}$ is a **linear subspace of** $\mathscr{L}^2\left(P_{X,L,W}\right)$ and $h_1, h_2, \ldots$ is an orthonormal basis of $\mathscr{H}$ then

$$\sum_{j=1}^{\infty}\left\{\mathbb{E}\left[h_j\left(q(X), L, W\right) - h_j(X, L, W)g(X, L, Z)\right]\right\}^2 = \max_{h\in\mathscr{H}} \mathbb{E}\left[h\left(q(X), L, W\right) - h(X, L, W)g(X, L, Z) - h(X, L, W)^2\right]$$

This motivates estimators of the form

$$\hat{g}^{(-k)} = \arg\min_{g\in\mathscr{G}}\max_{h\in\mathscr{H}'} \mathbb{E}_n^{(-k)}\left[h(q(X), L, W) - h(X, L, W)g(X, L, Z) - h(X, L, W)^2\right]$$

This is analogous to the estimators proposed by Bennett, et. al. 2019, where $\mathscr{G}$ and $\mathscr{H}'$ are neural nets. Assuming a unique solution to the integral, the authors showed consistency. However, to achieve convergence rates, regularization is needed due to the ill-possedness of the problem.

Min-max regularized estimator of the trx bridge function

Following Dikkala et al. 2020 and Ghassami et al. 2022, we considered estimators of the form

Min-max regularized estimator of the trx bridge function

Following Dikkala et al. 2020 and Ghassami et al. 2022, we considered estimators of the form

$$\hat{g}^{(-k)} = \arg\min_{g \in \mathcal{G}} \max_{h \in \mathcal{H}'} \mathbb{E}_n^{(-k)} \left[h(q(X), L, W) - h(X, L, W) g(X, L, Z) - h(X, L, W)^2 \right] - \lambda_{\mathcal{H}'} \|h\|^2_{\mathcal{H}'} + \lambda_{\mathcal{G}} \|g\|^2_{\mathcal{G}},$$

where $\mathcal{G}$ and $\mathcal{H}'$ are RKHS function classes and the penalization parameters are selected by cross-validation.

Convergence analysis of trx bridge fcn estimator[4]

[4]Adaptation of the convergence result established by Bennett et al. (2023).

Convergence analysis of trx bridge fcn estimator[4]

<u>Theorem</u>: Let $\mathscr{G}$ and $\mathscr{H}'$ be Gaussian RKHS satisfying certain regularity conditions. Suppose that $\mathscr{G}$ contains an observed trx bridge fcn and let $g^{\dagger}$ be the one with minimum $||\cdot||_{\mathscr{G}}$ norm. If $g^{\dagger}$ satisfies a compounded β-**source condition,** then, with high probability, $\hat{g}^{(-k)}$ satisfies

[4]Adaptation of the convergence result established by Bennett et al. (2023).

Convergence analysis of trx bridge fcn estimator[4]

Theorem: Let $\mathscr{G}$ and $\mathscr{H}'$ be Gaussian RKHS satisfying certain regularity conditions. Suppose that $\mathscr{G}$ contains an observed trx bridge fcn and let $g^\dagger$ be the one with minimum $\|\cdot\|_{\mathscr{G}}$ norm. If $g^\dagger$ satisfies a compounded β**-source condition,** then, with high probability, $\hat{g}^{(-k)}$ satisfies

$$\left\| \hat{g}^{(-k)} - g^\dagger \right\|_2^2 = O\left(\frac{\log(n)/n}{\lambda_{\mathscr{G}}} + \frac{\lambda_{\mathscr{H}'}}{\lambda_{\mathscr{G}}} + \lambda_{\mathscr{G}}^{\min\{\beta,1\}} \right),$$

[4]Adaptation of the convergence result established by Bennett et al. (2023).

Convergence analysis of trx bridge fcn estimator[4]

Theorem: Let $\mathscr{G}$ and $\mathscr{H}'$ be Gaussian RKHS satisfying certain regularity conditions. Suppose that $\mathscr{G}$ contains an observed trx bridge fcn and let $g^\dagger$ be the one with minimum $\|\cdot\|_{\mathscr{G}}$ norm. If $g^\dagger$ satisfies a compounded β**-source condition,** then, with high probability, $\hat{g}^{(-k)}$ satisfies

$$\left\| \hat{g}^{(-k)} - g^\dagger \right\|_2^2 = O\left(\frac{\log(n)/n}{\lambda_{\mathscr{G}}} + \frac{\lambda_{\mathscr{H}'}}{\lambda_{\mathscr{G}}} + \lambda_{\mathscr{G}}^{\min\{\beta,1\}} \right),$$

and

[4]Adaptation of the convergence result established by Bennett et al. (2023).

Convergence analysis of trx bridge fcn estimator[4]

Theorem: Let $\mathscr{G}$ and $\mathscr{H}'$ be Gaussian RKHS satisfying certain regularity conditions. Suppose that $\mathscr{G}$ contains an observed trx bridge fcn and let $g^\dagger$ be the one with minimum $\|\cdot\|_{\mathscr{G}}$ norm. If $g^\dagger$ satisfies a compounded β**-source condition,** then, with high probability, $\hat{g}^{(-k)}$ satisfies

$$\left\| \hat{g}^{(-k)} - g^\dagger \right\|_2^2 = O\left(\frac{\log(n)/n}{\lambda_{\mathscr{G}}} + \frac{\lambda_{\mathscr{H}'}}{\lambda_{\mathscr{G}}} + \lambda_{\mathscr{G}}^{\min\{\beta,1\}} \right),$$

and

$$\left\| \mathscr{T}^* \left(\hat{g}^{(-k)} - g^\dagger \right) \right\|_2^2 = O\left(\frac{\log(n)}{n} + \lambda_{\mathscr{H}'} + \lambda_{\mathscr{G}}^{\min\{\beta+1,2\}} \right).$$

[4]Adaptation of the convergence result established by Bennett et al. (2023).

Gaussian Kernels Guarantee Convergence Rates

For example, if $\beta \geq 1$, taking $\lambda_{\mathscr{G}} = \left(\log(n)/n\right)^{1/2}$ and $\lambda_{\mathscr{H}'} = \lambda_{\mathscr{G}}^2$ in the previous Theorem ensures that:

$$\left\| \hat{g}^{(-k)} - g_0 \right\|_2^2 = O_p\left(\sqrt{\frac{\log(n)}{n}}\right),$$

and

$$\left\| \mathscr{T}^* \left(\hat{g}^{(-k)} - g_0\right) \right\|_2^2 = O_p\left(\frac{\log(n)}{n}\right).$$

Min-max regularized estimator of the outcome bridge function

Likewise, to estimate the outcome bridge function, following Dikkala et al. 2020 and Ghassami et al. 2022, we consider estimators of the form

Min-max regularized estimator of the outcome bridge function

Likewise, to estimate the outcome bridge function, following Dikkala et al. 2020 and Ghassami et al. 2022, we consider estimators of the form

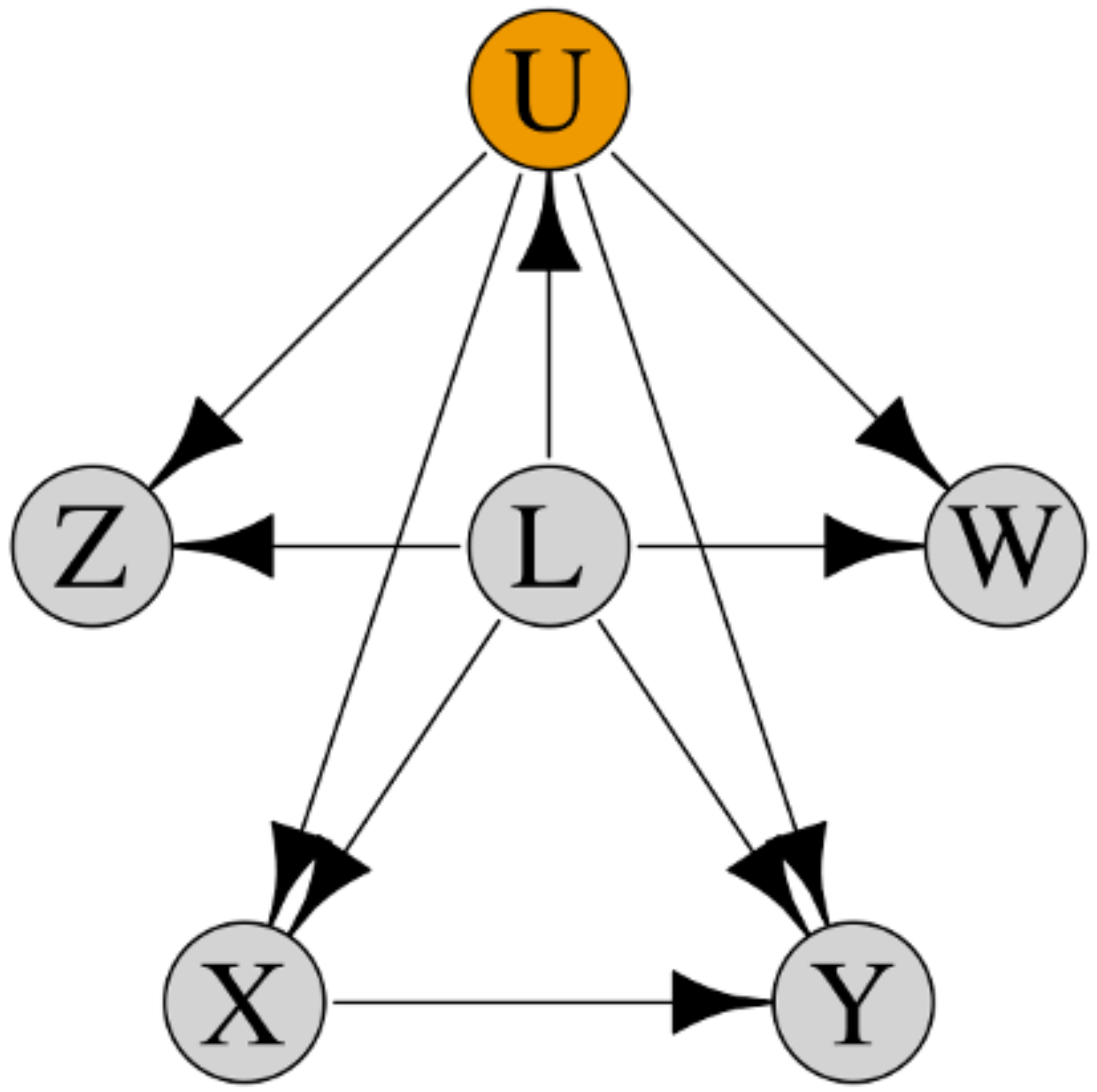
$$\hat{h}^{(-k)} = \arg\min_{h \in \mathscr{H}} \max_{g \in \mathscr{G}'} \mathbb{E}_n^{(-k)} \left[g(X, L, Z)\{Y - h(X, L, W)\} - g(X, L, Z)^2 \right] - \lambda_{\mathscr{G}'} \|g\|_{\mathscr{G}'}^2 + \lambda_{\mathscr{H}} \|h\|_{\mathscr{H}}^2,$$

where $\mathscr{H}$ and $\mathscr{G}'$ are RKHS function classes and the penalization parameters are selected by cross-validation.

Outline

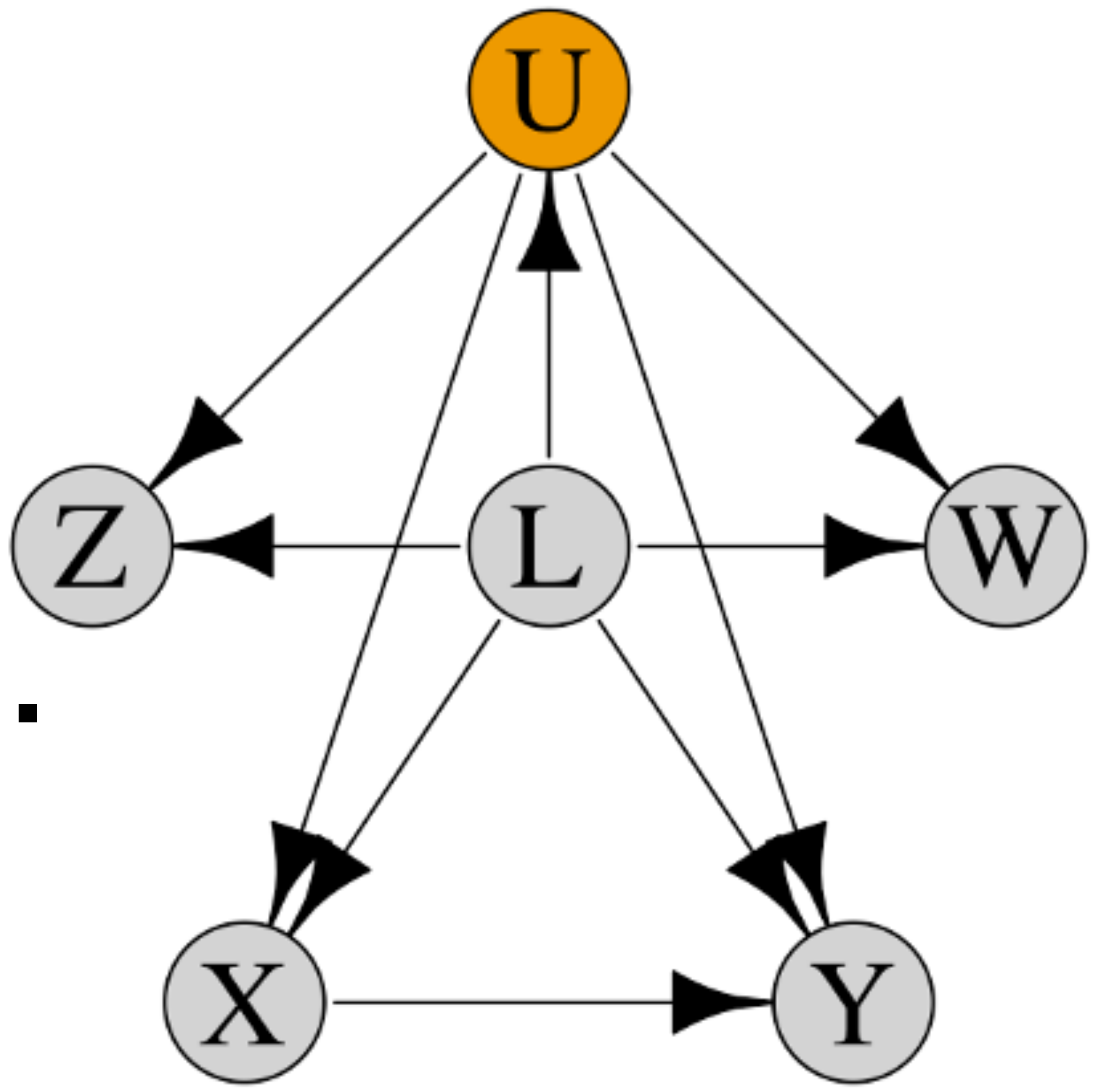
- Biomarkers in vaccine research
- Negative control outcomes and treatments
- Modified treatment policies (MTPs): Identification in the absence of unmeasured confounding.
- Modified treatment policies (MTPs): Proximal Causal Inference for MTP means with unmeasured confounding
  - Identification
  - Estimation
- **Simulation experiments**
- Application to the analysis of a vaccine trial

Numerical Experiments

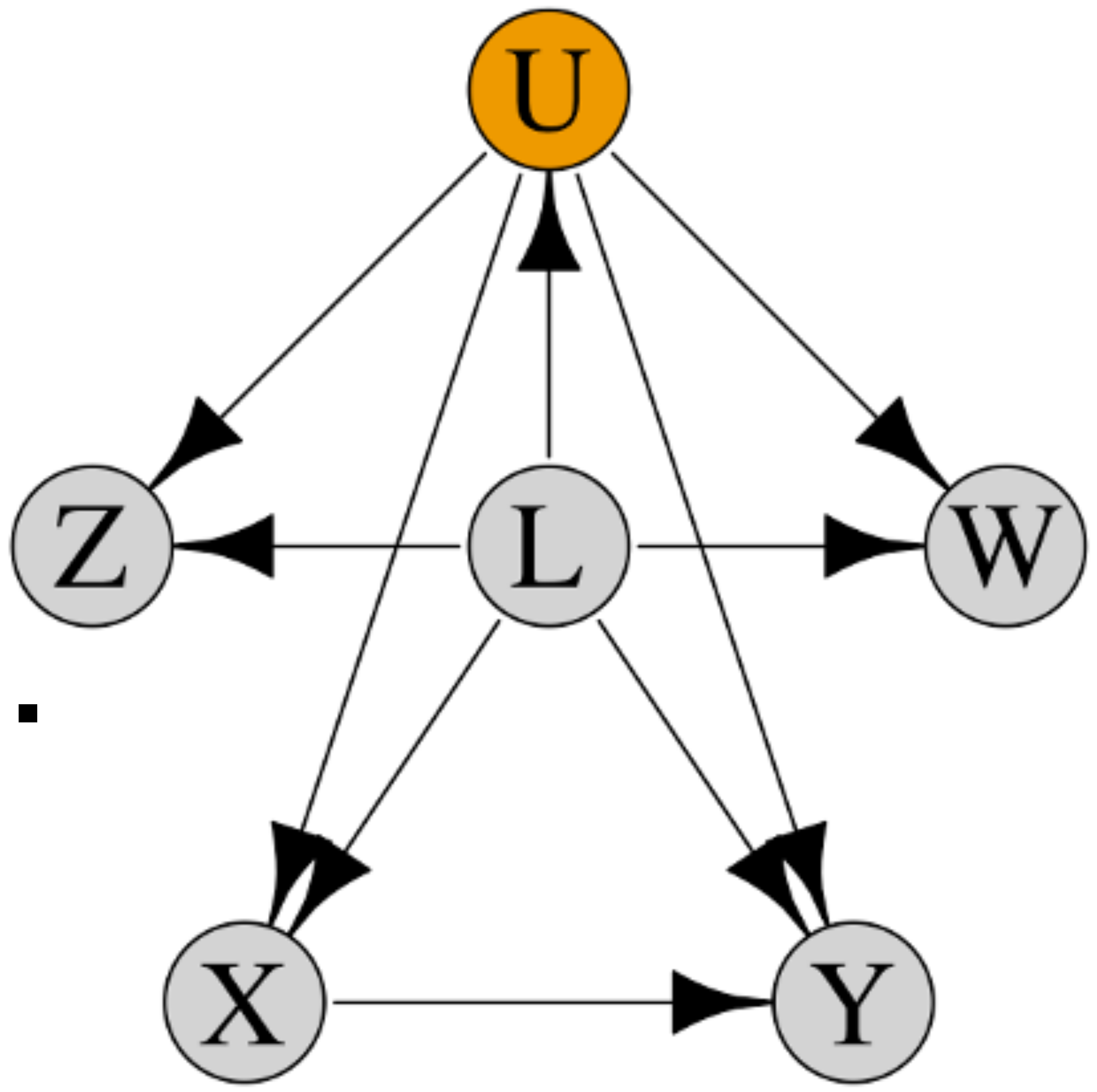


Numerical Experiments

1. **Bounded** support data for $n \in \{750, 1500, 3000, 6000\}$.



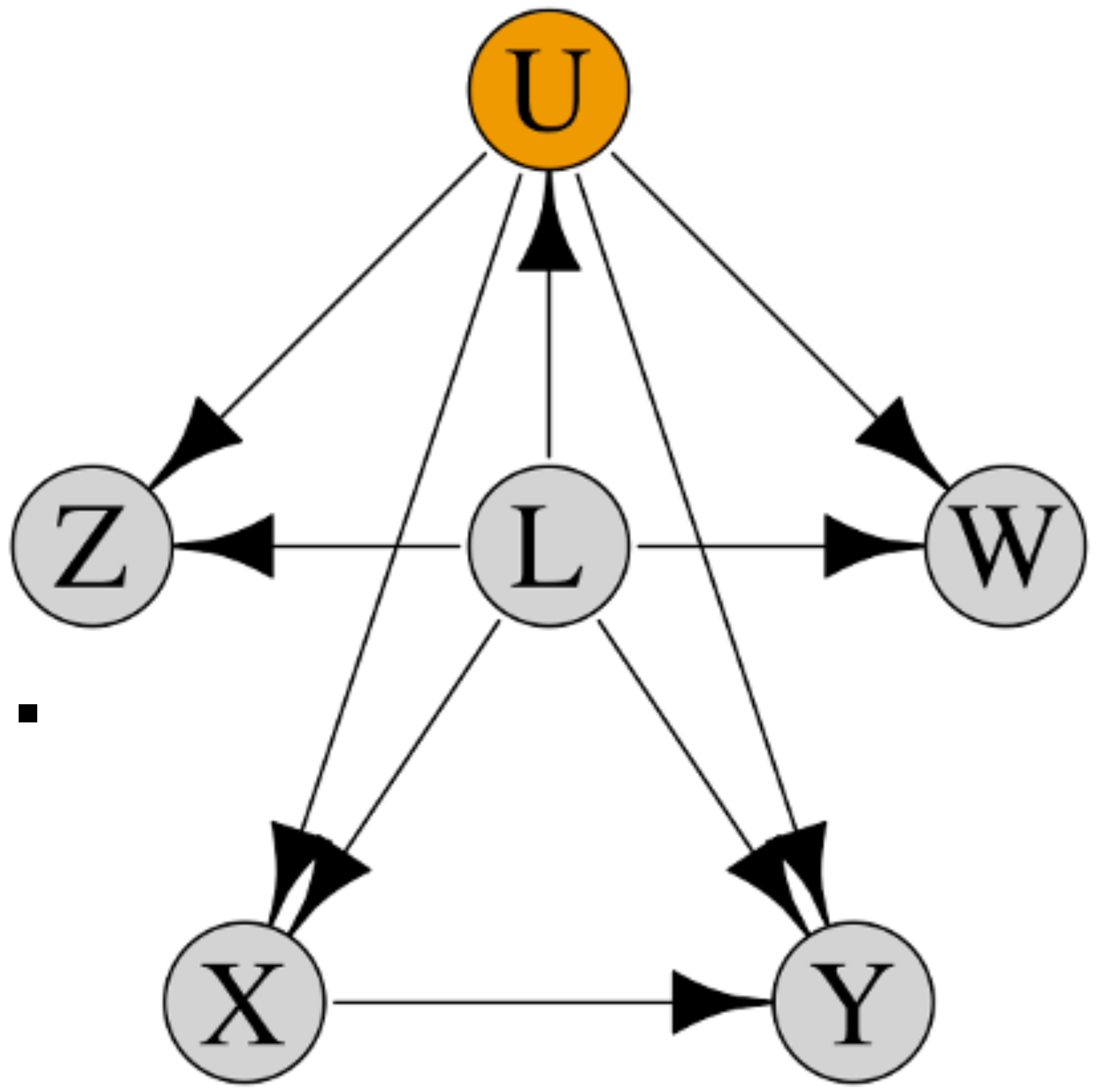
Numerical Experiments



1. **Bounded** support data for $n \in \{750, 1500, 3000, 6000\}$.

2. Nine data generating mechanisms defined by:
   a) Three levels of Cor $[Z, U \mid L] \in \{0.894, 0.707, 0.447\}$.
   b) Three levels of Cor $[W, U \mid L] \in \{-0.894, -0.707, -0.447\}$.

Numerical Experiments

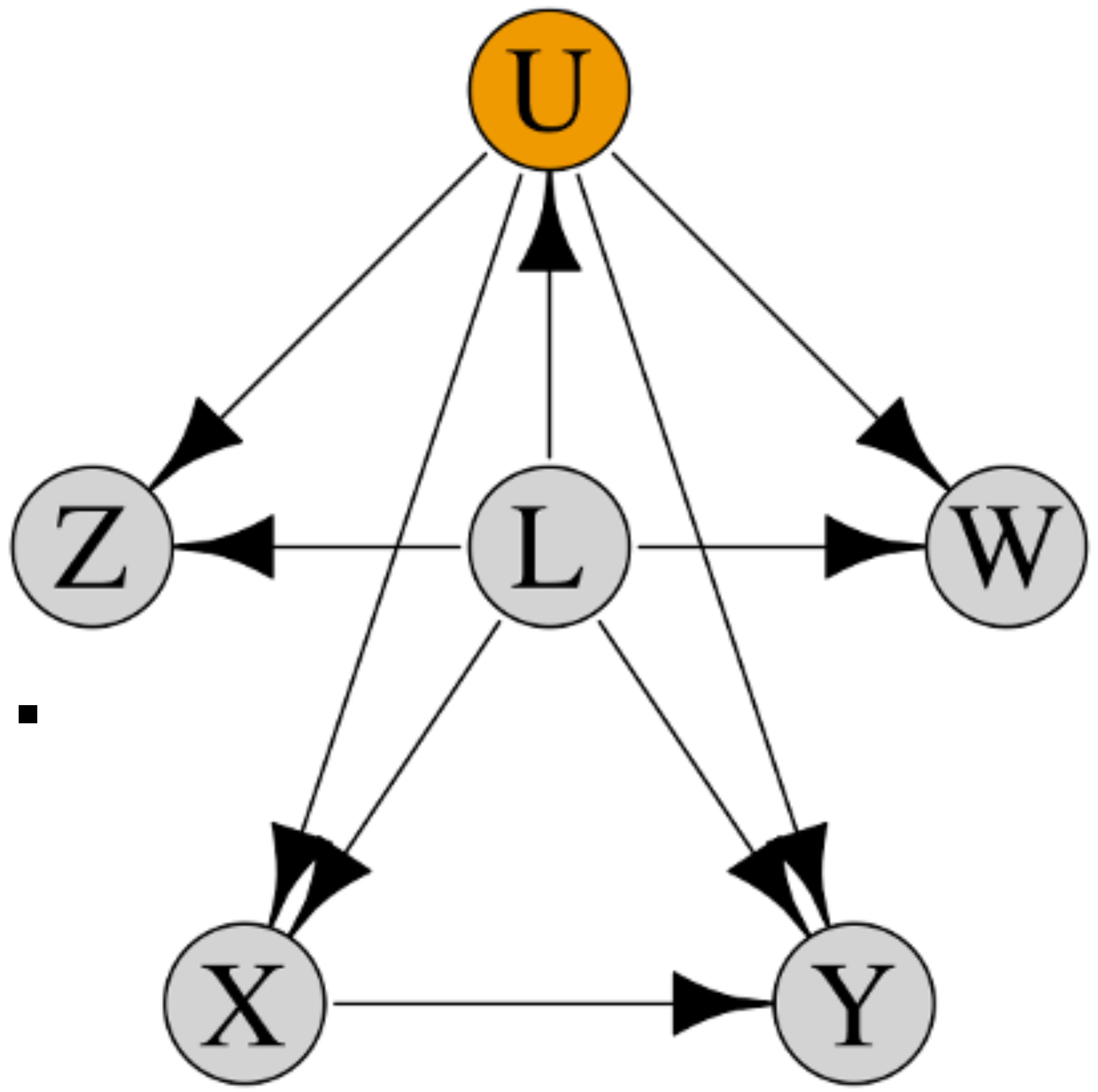


1. **Bounded** support data for $n \in \{750, 1500, 3000, 6000\}$.

2. Nine data generating mechanisms defined by:
   a) Three levels of Cor $[Z, U \mid L] \in \{0.894, 0.707, 0.447\}$.
   b) Three levels of Cor $[W, U \mid L] \in \{-0.894, -0.707, -0.447\}$.

3. **Policy:** Shift intervention under strategy 2 with $\delta = 0.4$ and $\varepsilon = 1$.

Numerical Experiments



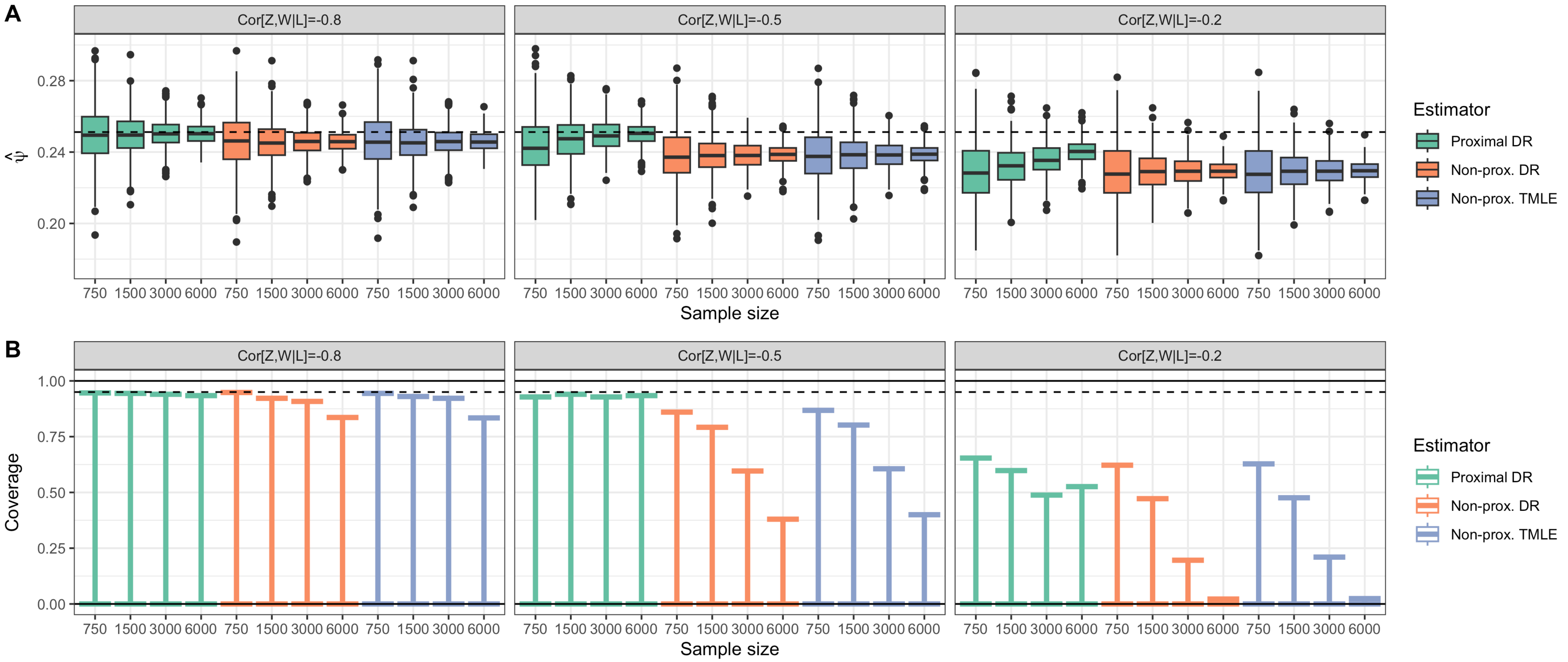
1. **Bounded** support data for $n \in \{750, 1500, 3000, 6000\}$.

2. Nine data generating mechanisms defined by:
   a) Three levels of Cor $[Z, U \mid L] \in \{0.894, 0.707, 0.447\}$.
   b) Three levels of Cor $[W, U \mid L] \in \{-0.894, -0.707, -0.447\}$.

3. **Policy:** Shift intervention under strategy 2 with $\delta = 0.4$ and $\varepsilon = 1$.

4. **Function classes:** RKHS generated by Gaussian kernels.

Numerical Experiments

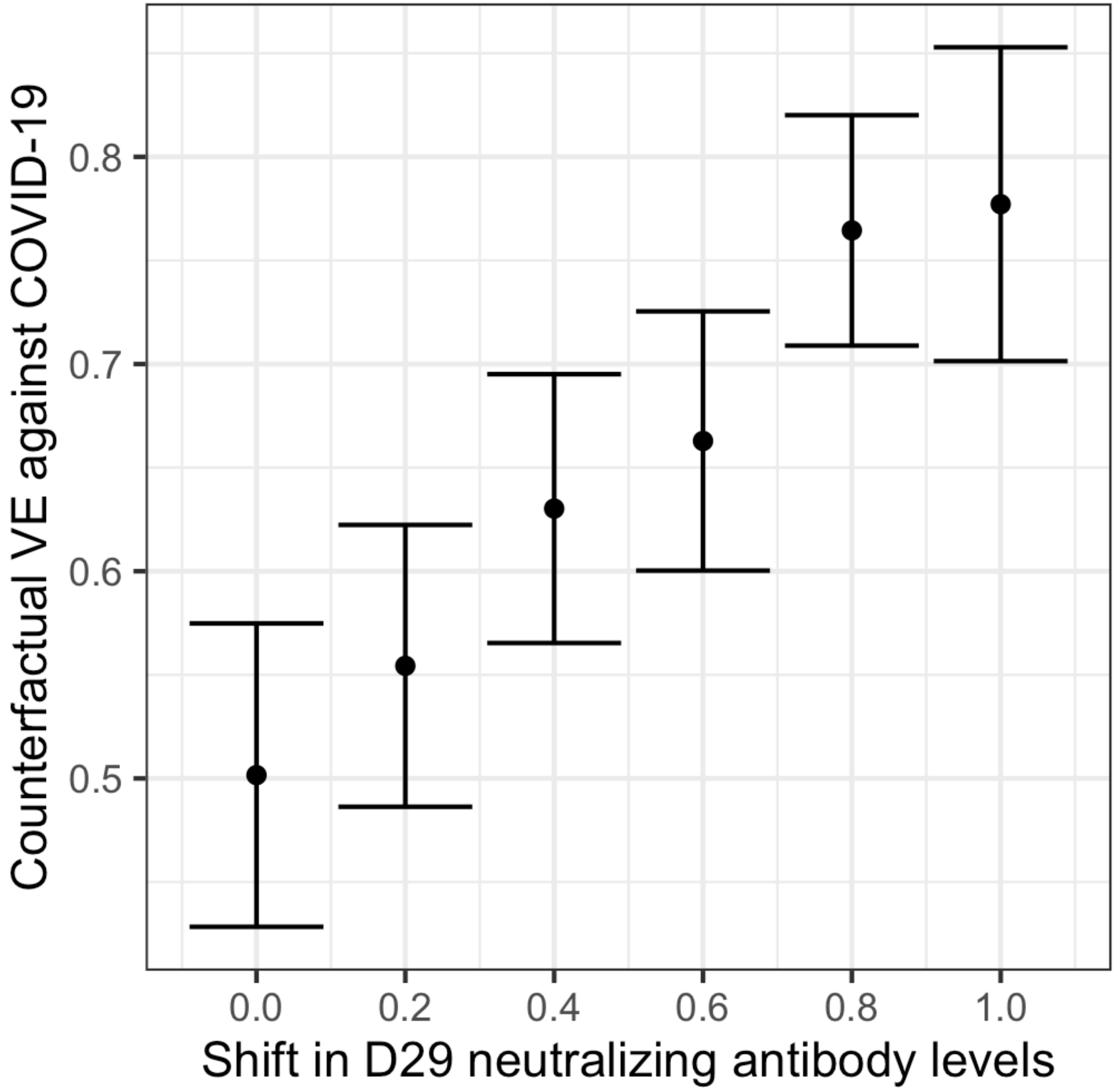


Non-proximal estimators are the ones proposed by Diaz-Munoz & van Der Laan (2012).

Outline

- Biomarkers in vaccine research
- Negative control outcomes and treatments
- Modified treatment policies (MTPs): Identification in the absence of unmeasured confounding.
- Modified treatment policies (MTPs): Proximal Causal Inference for MTP means with unmeasured confounding
  - Identification
  - Estimation
- Simulation experiments
- **Application to the analysis of a vaccine trial**

ENSEMBLE Trial



We applied a modified **shift intervention** with $\varepsilon = 1$, and $\delta \in \{0.2, 0.4, 0.6, 0.8, 1.0\}$.[5]

Potential observed confounders: Age. Region, Risk score of COVID-19 infection

Using data from the placebo arm, we estimated $\mathbb{E}\left[Y|\text{Placebo}\right]$ and with data in the trx arm we estimated $\mathbb{E}\left[Y(q^{\delta})\right]$

We then estimated the vaccine efficacy:

$$VE(\delta) = 1 - \frac{\mathbb{E}\left[Y(q^{\delta})\right]}{\mathbb{E}\left[Y \mid \text{Placebo}\right]}$$

VE observed in the trial: 50.2%.

$VE(\delta = 1)$: 77.7%.

[5] Exposure is in log10 scale. Estimation accounting for two-phase sampling.

Discussion

- **Limitations of proximal causal inference.**
  - Requires large sample sizes when negative controls are weak proxies.
  - Negative control treatment and outcomes are not easy to come by.

Discussion

❖**Limitations of proximal causal inference.**

- Requires large sample sizes when negative controls are weak proxies.
- Negative control treatment and outcomes are not easy to come by.

❖**Current areas of active research**

- Non-parametric estimation of bridge functions: Tuning parameter selection? Ensemble methods?
- Doubly-robust estimation with one bridge function estimator inconsistent and the other non-parametrically estimated.
- For MTP interventions: extensions to account for time dependent MTP's and time to event outcomes

Backup Slides

Completeness Conditions

Completeness Conditions

Define $\mathscr{A} : \mathscr{L}^2(X, L, W) \to \mathscr{L}^2(X, L, U)$ and $\mathscr{A}^* : \mathscr{L}^2(X, L, U) \to \mathscr{L}^2(X, L, W)$ as:

$$\mathscr{A}h := \mathbb{E}\left[h(X, L, W) \,|\, X, L, U\right] \quad \text{and} \quad \mathscr{A}^*g := \mathbb{E}\left[g(X, L, U) \,|\, X, L, W\right]$$

Completeness Conditions

Outcome bridge equation:

$$\mathbb{E}\left[Y \mid X, L, U\right] = \mathscr{A}h$$

Define $\mathscr{A} : \mathscr{L}^2(X, L, W) \to \mathscr{L}^2(X, L, U)$ and $\mathscr{A}^* : \mathscr{L}^2(X, L, U) \to \mathscr{L}^2(X, L, W)$ as:

$$\mathscr{A}h := \mathbb{E}\left[h(X, L, W) \mid X, L, U\right] \quad \text{and} \quad \mathscr{A}^*g := \mathbb{E}\left[g(X, L, U) \mid X, L, W\right]$$

Completeness Conditions

Outcome bridge equation:

$$\mathbb{E}\left[Y \,|\, X, L, U\right] = \mathscr{A}h$$

Define $\mathscr{A} : \mathscr{L}^2(X, L, W) \to \mathscr{L}^2(X, L, U)$ and $\mathscr{A}^* : \mathscr{L}^2(X, L, U) \to \mathscr{L}^2(X, L, W)$ as:

$$\mathscr{A}h := \mathbb{E}\left[h(X, L, W) \,|\, X, L, U\right] \quad \text{and} \quad \mathscr{A}^*g := \mathbb{E}\left[g(X, L, U) \,|\, X, L, W\right]$$

- The **completeness condition:**

Completeness Conditions

Outcome bridge equation:

$$\mathbb{E}\left[Y \mid X, L, U\right] = \mathscr{A}h$$

Define $\mathscr{A} : \mathscr{L}^2(X, L, W) \to \mathscr{L}^2(X, L, U)$ and $\mathscr{A}^* : \mathscr{L}^2(X, L, U) \to \mathscr{L}^2(X, L, W)$ as:

$$\mathscr{A}h := \mathbb{E}\left[h(X, L, W) \mid X, L, U\right] \quad \text{and} \quad \mathscr{A}^*g := \mathbb{E}\left[g(X, L, U) \mid X, L, W\right]$$

- The **completeness condition:**

$$\mathbb{E}\left[\eta(X, L, U) \mid X, L, W\right] = 0 \quad \Rightarrow \quad \eta(X, L, U) = 0 \quad \text{a.e.}$$

Completeness Conditions

Outcome bridge equation:

$$\mathbb{E}\left[Y \mid X, L, U\right] = \mathscr{A}h$$

Define $\mathscr{A} : \mathscr{L}^2(X, L, W) \to \mathscr{L}^2(X, L, U)$ and $\mathscr{A}^* : \mathscr{L}^2(X, L, U) \to \mathscr{L}^2(X, L, W)$ as:

$$\mathscr{A}h := \mathbb{E}\left[h(X, L, W) \mid X, L, U\right] \quad \text{and} \quad \mathscr{A}^*g := \mathbb{E}\left[g(X, L, U) \mid X, L, W\right]$$

- The **completeness condition:**

$$\mathbb{E}\left[\eta(X, L, U) \mid X, L, W\right] = 0 \quad \Rightarrow \quad \eta(X, L, U) = 0 \quad \text{a.e.}$$

implies that

Completeness Conditions

Outcome bridge equation:

$$\mathbb{E}\left[Y \,|\, X, L, U\right] = \mathscr{A}h$$

Define $\mathscr{A} : \mathscr{L}^2(X, L, W) \to \mathscr{L}^2(X, L, U)$ and $\mathscr{A}^* : \mathscr{L}^2(X, L, U) \to \mathscr{L}^2(X, L, W)$ as:

$$\mathscr{A}h := \mathbb{E}\left[h(X, L, W) \,|\, X, L, U\right] \quad \text{and} \quad \mathscr{A}^*g := \mathbb{E}\left[g(X, L, U) \,|\, X, L, W\right]$$

- The **completeness condition:**

$$\mathbb{E}\left[\eta(X, L, U) \,|\, X, L, W\right] = 0 \quad \Rightarrow \quad \eta(X, L, U) = 0 \quad \text{a.e.}$$

  implies that

$$\ker\left(\mathscr{A}^*\right) := \left\{g \in \mathscr{L}^2(X, L, U) : \mathscr{A}^*g = 0\right\} = \{0\}\,.$$

Completeness Conditions

Outcome bridge equation:

$$\mathbb{E}\left[Y \mid X, L, U\right] = \mathscr{A}h$$

Define $\mathscr{A} : \mathscr{L}^2(X, L, W) \to \mathscr{L}^2(X, L, U)$ and $\mathscr{A}^* : \mathscr{L}^2(X, L, U) \to \mathscr{L}^2(X, L, W)$ as:

$$\mathscr{A}h := \mathbb{E}\left[h(X, L, W) \,|\, X, L, U\right] \quad \text{and} \quad \mathscr{A}^*g := \mathbb{E}\left[g(X, L, U) \,|\, X, L, W\right]$$

- The **completeness condition:**

$$\mathbb{E}\left[\eta(X, L, U) \,|\, X, L, W\right] = 0 \quad \Rightarrow \quad \eta(X, L, U) = 0 \quad \text{a.e.}$$

  implies that

$$\ker(\mathscr{A}^*) := \left\{g \in \mathscr{L}^2(X, L, U) : \mathscr{A}^*g = 0\right\} = \{0\}\,.$$

  Then

Completeness Conditions

Outcome bridge equation:

$$\mathbb{E}\left[Y \mid X, L, U\right] = \mathscr{A}h$$

Define $\mathscr{A} : \mathscr{L}^2(X, L, W) \to \mathscr{L}^2(X, L, U)$ and $\mathscr{A}^* : \mathscr{L}^2(X, L, U) \to \mathscr{L}^2(X, L, W)$ as:

$$\mathscr{A}h := \mathbb{E}\left[h(X, L, W) \mid X, L, U\right] \quad \text{and} \quad \mathscr{A}^*g := \mathbb{E}\left[g(X, L, U) \mid X, L, W\right]$$

- The **completeness condition:**

$$\mathbb{E}\left[\eta(X, L, U) \mid X, L, W\right] = 0 \quad \Rightarrow \quad \eta(X, L, U) = 0 \quad \text{a.e.}$$

implies that

$$\ker\left(\mathscr{A}^*\right) := \left\{g \in \mathscr{L}^2(X, L, U) : \mathscr{A}^*g = 0\right\} = \{0\}\,.$$

Then

$$\overline{\mathscr{R}\left(\mathscr{A}\right)} = \left[\ker\left(\mathscr{A}^*\right)\right]^{\perp} = \mathscr{L}^2(X, L, U).$$

Discrete Bridge Equations

X = Exposure
Y = Outcome
L = Observed confounders
U = Unobserved confounders
Z = Negative control Treatment
W = Negative control Outcome

Discrete Bridge Equations

X = Exposure
Y = Outcome
L = Observed confounders
U = Unobserved confounders
Z = Negative control Treatment
W = Negative control Outcome

- $W \in \{w_1, \ldots, w_m\}$, $Z \in \{z_1, \ldots, z_n\}$, and $U \in \{u_1, \ldots, u_r\}$.

Discrete Bridge Equations

$X =$ Exposure
$Y =$ Outcome
$L =$ Observed confounders
$U =$ Unobserved confounders
$Z =$ Negative control Treatment
$W =$ Negative control Outcome

- $W \in \{w_1, \ldots, w_m\}, Z \in \{z_1, \ldots, z_n\}$, and $U \in \{u_1, \ldots, u_r\}$.

- For given $x, l \in \mathscr{X} \times \mathscr{L}$, define the $m \times r$ matix:

Discrete Bridge Equations

X = Exposure
Y = Outcome
L = Observed confounders
U = Unobserved confounders
Z = Negative control Treatment
W = Negative control Outcome

- $W \in \{w_1, \ldots, w_m\}, Z \in \{z_1, \ldots, z_n\}$, and $U \in \{u_1, \ldots, u_r\}$.

- For given $x, l \in \mathscr{X} \times \mathscr{L}$, define the $m \times r$ matix:

$$P\left(\mathbf{W} \,|\, \mathbf{U}, x, l\right)_{m \times r} = \begin{bmatrix} \mathbb{P}\left(W = w_1 \,|\, U = u_1, x, l\right) & \cdots & \mathbb{P}\left(W = w_1 \,|\, U = u_r, x, l\right) \\ \vdots & \ddots & \vdots \\ \mathbb{P}\left(W = w_m \,|\, U = u_1, x, l\right) & \cdots & \mathbb{P}\left(W = w_m \,|\, U = u_r, x, l\right) \end{bmatrix}$$

Discrete Bridge Equations

X = Exposure
Y = Outcome
L = Observed confounders
U = Unobserved confounders
Z = Negative control Treatment
W = Negative control Outcome

- $W \in \{w_1, \ldots, w_m\}, Z \in \{z_1, \ldots, z_n\}$, and $U \in \{u_1, \ldots, u_r\}$.

- For given $x, l \in \mathcal{X} \times \mathcal{L}$, define the $m \times r$ matix:

$$P\left(\mathbf{W} \,|\, \mathbf{U}, x, l\right)_{m \times r} = \begin{bmatrix} \mathbb{P}\left(W = w_1 \,|\, U = u_1, x, l\right) & \cdots & \mathbb{P}\left(W = w_1 \,|\, U = u_r, x, l\right) \\ \vdots & \ddots & \vdots \\ \mathbb{P}\left(W = w_m \,|\, U = u_1, x, l\right) & \cdots & \mathbb{P}\left(W = w_m \,|\, U = u_r, x, l\right) \end{bmatrix}$$

Discrete Bridge Equations

X = Exposure
Y = Outcome
L = Observed confounders
U = Unobserved confounders
Z = Negative control Treatment
W = Negative control Outcome

- $W \in \{w_1, \ldots, w_m\}, Z \in \{z_1, \ldots, z_n\}$, and $U \in \{u_1, \ldots, u_r\}$.

- For given $x, l \in \mathcal{X} \times \mathcal{L}$, define the $m \times r$ matix:

$$P\left(\mathbf{W} \,|\, \mathbf{U}, x, l\right)_{m \times r} = \begin{bmatrix} \mathbb{P}\left(W = w_1 \,|\, U = u_1, x, l\right) & \cdots & \mathbb{P}\left(W = w_1 \,|\, U = u_r, x, l\right) \\ \vdots & \ddots & \vdots \\ \mathbb{P}\left(W = w_m \,|\, U = u_1, x, l\right) & \cdots & \mathbb{P}\left(W = w_m \,|\, U = u_r, x, l\right) \end{bmatrix}$$

- Define $P\left(\mathbf{Z} \,|\, \mathbf{U}, x, l\right)_{n \times r}$ similarly.

Discrete Bridge Equations

X = Exposure
Y = Outcome
L = Observed confounders
U = Unobserved confounders
Z = Negative control Treatment
W = Negative control Outcome

Discrete Bridge Equations

X = Exposure
Y = Outcome
L = Observed confounders
U = Unobserved confounders
Z = Negative control Treatment
W = Negative control Outcome

1. Outcome bridge equation:

Discrete Bridge Equations

$X =$ Exposure
$Y =$ Outcome
$L =$ Observed confounders
$U =$ Unobserved confounders
$Z =$ Negative control Treatment
$W =$ Negative control Outcome

1. Outcome bridge equation:

$$\begin{bmatrix} \mathbb{E}\left(Y \mid x, l, U = u_1\right) \\ \vdots \\ \mathbb{E}\left(Y \mid x, l, U = u_r\right) \end{bmatrix} = \left[P\left(\mathbf{W} \mid \mathbf{U}, x, l\right)_{m \times r}\right]^T \begin{bmatrix} h(x, l, W = w_1) \\ \vdots \\ h(x, l, W = w_m) \end{bmatrix}$$

Discrete Bridge Equations

$X =$ Exposure
$Y =$ Outcome
$L =$ Observed confounders
$U =$ Unobserved confounders
$Z =$ Negative control Treatment
$W =$ Negative control Outcome

1. Outcome bridge equation:

$$\begin{bmatrix} \mathbb{E}\left(Y \mid x, l, U = u_1\right) \\ \vdots \\ \mathbb{E}\left(Y \mid x, l, U = u_r\right) \end{bmatrix} = \left[P\left(\mathbf{W} \mid \mathbf{U}, x, l\right)_{m \times r}\right]^T \begin{bmatrix} h(x, l, W = w_1) \\ \vdots \\ h(x, l, W = w_m) \end{bmatrix}$$

2. Treatment bridge equation:

Discrete Bridge Equations

$X =$ Exposure
$Y =$ Outcome
$L =$ Observed confounders
$U =$ Unobserved confounders
$Z =$ Negative control Treatment
$W =$ Negative control Outcome

1. Outcome bridge equation:

$$\begin{bmatrix} \mathbb{E}\left(Y|x,l,U=u_1\right) \\ \vdots \\ \mathbb{E}\left(Y|x,l,U=u_r\right) \end{bmatrix} = \left[P\left(\mathbf{W}|\mathbf{U},x,l\right)_{m\times r}\right]^T \begin{bmatrix} h(x,l,W=w_1) \\ \vdots \\ h(x,l,W=w_m) \end{bmatrix}$$

2. Treatment bridge equation:

$$\begin{bmatrix} \alpha_0\left(x,l,U=u_1\right) \\ \vdots \\ \alpha_0\left(x,l,U=u_r\right) \end{bmatrix} = \left[P\left(\mathbf{Z}|\mathbf{U},x,l\right)_{n\times r}\right]^T \begin{bmatrix} g(x,l,Z=z_1) \\ \vdots \\ g(x,l,Z=z_n) \end{bmatrix}$$

Discrete Bridge Equations

X = Exposure
Y = Outcome
L = Observed confounders
U = Unobserved confounders
Z = Negative control Treatment
W = Negative control Outcome

Treatment bridge equation:

$$\left[\alpha_0\left(x, l, u_i\right)\right]_{r\times 1} = \left[P\left(\mathbf{Z} \mid \mathbf{U}, x, l\right)_{n\times r}\right]^T \left[g\left(x, l, z_j\right)\right]_{n\times 1}$$

Discrete Bridge Equations

$X =$ Exposure
$Y =$ Outcome
$L =$ Observed confounders
$U =$ Unobserved confounders
$Z =$ Negative control Treatment
$W =$ Negative control Outcome

Treatment bridge equation:

$$\left[\alpha_0\left(x, l, u_i\right)\right]_{r\times 1} = \left[P\left(\mathbf{Z} \,|\, \mathbf{U}, x, l\right)_{n\times r}\right]^T \left[g\left(x, l, z_j\right)\right]_{n\times 1}$$

- **Existence of solutions:**

Discrete Bridge Equations

X = Exposure
Y = Outcome
L = Observed confounders
U = Unobserved confounders
Z = Negative control Treatment
W = Negative control Outcome

Treatment bridge equation:

$$\left[\alpha_0\left(x, l, u_i\right)\right]_{r\times 1} = \left[P\left(\mathbf{Z} \mid \mathbf{U}, x, l\right)_{n\times r}\right]^T \left[g\left(x, l, z_j\right)\right]_{n\times 1}$$

- **Existence of solutions:**

✓ $\left[\alpha_0\left(x, l, u_i\right)\right]_{r\times 1}$ lies in the space generated by the rows of $P\left(\mathbf{Z} \mid \mathbf{U}, x, l\right)_{n\times r}$

Discrete Bridge Equations

$X =$ Exposure
$Y =$ Outcome
$L =$ Observed confounders
$U =$ Unobserved confounders
$Z =$ Negative control Treatment
$W =$ Negative control Outcome

Treatment bridge equation:

$$\left[\alpha_0\left(x, l, u_i\right)\right]_{r\times 1} = \left[P\left(\mathbf{Z} \mid \mathbf{U}, x, l\right)_{n\times r}\right]^T \left[g\left(x, l, z_j\right)\right]_{n\times 1}$$

- **Existence of solutions:**

✓ $\left[\alpha_0\left(x, l, u_i\right)\right]_{r\times 1}$ lies in the space generated by the rows of $P\left(\mathbf{Z} \mid \mathbf{U}, x, l\right)_{n\times r}$

- **Uniqueness of solutions:**

Discrete Bridge Equations

$X =$ Exposure
$Y =$ Outcome
$L =$ Observed confounders
$U =$ Unobserved confounders
$Z =$ Negative control Treatment
$W =$ Negative control Outcome

Treatment bridge equation:

$$\left[\alpha_0\left(x, l, u_i\right)\right]_{r\times 1} = \left[P\left(\mathbf{Z} \mid \mathbf{U}, x, l\right)_{n\times r}\right]^T \left[g\left(x, l, z_j\right)\right]_{n\times 1}$$

- **Existence of solutions:**

✓ $\left[\alpha_0\left(x, l, u_i\right)\right]_{r\times 1}$ lies in the space generated by the rows of $P\left(\mathbf{Z} \mid \mathbf{U}, x, l\right)_{n\times r}$

- **Uniqueness of solutions:**

✓ $n \geq r$ and $P\left(\mathbf{Z} \mid \mathbf{U}, x, l\right)$ is full row-rank, where $n = |\mathscr{Z}|$ and $r = |\mathscr{U}|$.

$X =$ Exposure
$Y =$ Outcome
$L =$ Observed confounders
$U =$ Unobserved confounders
$Z =$ Negative control treatment
$W =$ Negative control outcome

Observed Bridge Equations

Theorem 1: Under assumption 3 (negative controls):

1. Any solution h_0 to the **latent** outcome bridge equation also solves the equation:[2]

$$\mathbb{E}\left[Y \mid X, L, Z\right] = \mathbb{E}\left[h(X, L, W) \mid X, L, Z\right].$$

[2]This result was established by Miao et al. (2018), Tchetgen Tchetgen et al. (2020), and Kallus et al. (2020).

$X =$ Exposure
$Y =$ Outcome
$L =$ Observed confounders
$U =$ Unobserved confounders
$Z =$ Negative control treatment
$W =$ Negative control outcome

Observed Bridge Equations

Theorem 1: Under assumption 3 (negative controls):

1. Any solution h_0 to the **latent** outcome bridge equation also solves the equation:[2]

$$\mathbb{E}\left[Y \mid X, L, Z\right] = \mathbb{E}\left[h(X, L, W) \mid X, L, Z\right].$$

2. Any solution g_0 to the **latent** treatment bridge equation also solves the equation:

[2] This result was established by Miao et al. (2018), Tchetgen Tchetgen et al. (2020), and Kallus et al. (2020).

Observed Bridge Equations

X = Exposure
Y = Outcome
L = Observed confounders
U = Unobserved confounders
Z = Negative control treatment
W = Negative control outcome

Theorem 1: Under assumption 3 (negative controls):

1. Any solution h_0 to the **latent** outcome bridge equation also solves the equation:[2]

$$\mathbb{E}\left[Y \mid X, L, Z\right] = \mathbb{E}\left[h(X, L, W) \mid X, L, Z\right].$$

2. Any solution g_0 to the **latent** treatment bridge equation also solves the equation:

$$\alpha_0(X, L, W) = \mathbb{E}\left[g(X, L, Z) \mid X, L, W\right].$$

[2] This result was established by Miao et al. (2018), Tchetgen Tchetgen et al. (2020), and Kallus et al. (2020).

Bridge Equations Solution Sets

- Define

$$\mathbb{H}_0^{lat} = \{h : h \text{ solves the } \textbf{latent} \text{ outcome bridge equation}\}$$

Bridge Equations Solution Sets

- Define

$$\mathbb{H}_0^{lat} = \{h : h \text{ solves the } \textbf{latent} \text{ outcome bridge equation}\}$$

$$\mathbb{H}_0^{obs} = \{h : h \text{ solves the } \textbf{observed} \text{ outcome bridge equation}\}$$

Bridge Equations Solution Sets

- Define

$$\mathbb{H}_0^{lat} = \{h : h \text{ solves the } \textbf{latent} \text{ outcome bridge equation}\}$$

$$\mathbb{H}_0^{obs} = \{h : h \text{ solves the } \textbf{observed} \text{ outcome bridge equation}\}$$

$$\mathbb{G}_0^{lat} = \left\{g : g \text{ solves the } \textbf{latent} \text{ treatment bridge equation}\right\}$$

Bridge Equations Solution Sets

- Define

$$\mathbb{H}_0^{lat} = \{h : h \text{ solves the } \textbf{latent} \text{ outcome bridge equation}\}$$

$$\mathbb{H}_0^{obs} = \{h : h \text{ solves the } \textbf{observed} \text{ outcome bridge equation}\}$$

$$\mathbb{G}_0^{lat} = \left\{g : g \text{ solves the } \textbf{latent} \text{ treatment bridge equation}\right\}$$

$$\mathbb{G}_0^{obs} = \left\{g : g \text{ solves the } \textbf{observed} \text{ treatment bridge equation}\right\}$$

Bridge Equations Solution Sets

- Define

$$\mathbb{H}_0^{lat} = \{h : h \text{ solves the } \textbf{latent} \text{ outcome bridge equation}\}$$

$$\mathbb{H}_0^{obs} = \{h : h \text{ solves the } \textbf{observed} \text{ outcome bridge equation}\}$$

$$\mathbb{G}_0^{lat} = \left\{g : g \text{ solves the } \textbf{latent} \text{ treatment bridge equation}\right\}$$

$$\mathbb{G}_0^{obs} = \left\{g : g \text{ solves the } \textbf{observed} \text{ treatment bridge equation}\right\}$$

- Assumption 6 means: $\mathbb{H}_0^{lat} \neq \varnothing$

Bridge Equations Solution Sets

- Define

$$\mathbb{H}_0^{lat} = \{h : h \text{ solves the } \textbf{latent} \text{ outcome bridge equation}\}$$

$$\mathbb{H}_0^{obs} = \{h : h \text{ solves the } \textbf{observed} \text{ outcome bridge equation}\}$$

$$\mathbb{G}_0^{lat} = \left\{g : g \text{ solves the } \textbf{latent} \text{ treatment bridge equation}\right\}$$

$$\mathbb{G}_0^{obs} = \left\{g : g \text{ solves the } \textbf{observed} \text{ treatment bridge equation}\right\}$$

- Assumption 6 means: $\mathbb{H}_0^{lat} \neq \varnothing$
- Assumption 7 means: $\mathbb{G}_0^{lat} \neq \varnothing$

Bridge Equations Solution Sets

- Define

$$\mathbb{H}_0^{lat} = \{h : h \text{ solves the } \textbf{latent} \text{ outcome bridge equation}\}$$

$$\mathbb{H}_0^{obs} = \{h : h \text{ solves the } \textbf{observed} \text{ outcome bridge equation}\}$$

$$\mathbb{G}_0^{lat} = \{g : g \text{ solves the } \textbf{latent} \text{ treatment bridge equation}\}$$

$$\mathbb{G}_0^{obs} = \{g : g \text{ solves the } \textbf{observed} \text{ treatment bridge equation}\}$$

- Assumption 6 means: $\mathbb{H}_0^{lat} \neq \varnothing$
- Assumption 7 means: $\mathbb{G}_0^{lat} \neq \varnothing$
- Theorem 1 implies that: $\mathbb{H}_0^{lat} \subseteq \mathbb{H}^{obs}$ and $\mathbb{G}_0^{lat} \subseteq \mathbb{G}^{obs}$.

Existing Identification Strategies

Strategy	Assumptions	Representation
Our proposal		
Kallus et al. 2022		
Miao et al. 2018, Cui et al., 2023		

Existing Identification Strategies

Strategy	Assumptions	Representation
Our proposal	$\mathbb{H}_0^{lat} \neq \emptyset$ and $\mathbb{G}_0^{obs} \neq \emptyset$ or $\mathbb{G}_0^{lat} \neq \emptyset$ and $\mathbb{H}_0^{obs} \neq \emptyset$	
Kallus et al. 2022		
Miao et al. 2018, Cui et al., 2023		

Existing Identification Strategies

Strategy	Assumptions		Representation
Our proposal	$\mathbb{H}_0^{lat} \neq \varnothing$ and $\mathbb{G}_0^{obs} \neq \varnothing$ or $\mathbb{G}_0^{lat} \neq \varnothing$ and $\mathbb{H}_0^{obs} \neq \varnothing$	➡	1. Outcome bridge 2. Treatment bridge 3. Double Robust
Kallus et al. 2022			
Miao et al. 2018, Cui et al., 2023			

Existing Identification Strategies

Strategy	Assumptions		Representation
Our proposal	$\mathbb{H}_0^{lat} \neq \varnothing$ and $\mathbb{G}_0^{obs} \neq \varnothing$ or $\mathbb{G}_0^{lat} \neq \varnothing$ and $\mathbb{H}_0^{obs} \neq \varnothing$	$\Rightarrow$	1. Outcome bridge 2. Treatment bridge 3. Double Robust
Kallus et al. 2022	$\mathbb{H}_0^{lat} \neq \varnothing$ and $\mathbb{G}_0^{obs} \neq \varnothing$		
Miao et al. 2018, Cui et al., 2023			

Existing Identification Strategies

Strategy	Assumptions		Representation
Our proposal	$\mathbb{H}_0^{lat} \neq \varnothing$ and $\mathbb{G}_0^{obs} \neq \varnothing$ or $\mathbb{G}_0^{lat} \neq \varnothing$ and $\mathbb{H}_0^{obs} \neq \varnothing$	$\Rightarrow$	1. Outcome bridge 2. Treatment bridge 3. Double Robust
Kallus et al. 2022	$\mathbb{H}_0^{lat} \neq \varnothing$ and $\mathbb{G}_0^{obs} \neq \varnothing$	$\Rightarrow$	Outcome bridge
Miao et al. 2018, Cui et al., 2023			

Existing Identification Strategies

Strategy	Assumptions		Representation
Our proposal	$\mathbb{H}_0^{lat} \neq \varnothing$ and $\mathbb{G}_0^{obs} \neq \varnothing$ or $\mathbb{G}_0^{lat} \neq \varnothing$ and $\mathbb{H}_0^{obs} \neq \varnothing$	$\Rightarrow$	1. Outcome bridge 2. Treatment bridge 3. Double Robust
Kallus et al. 2022	$\mathbb{H}_0^{lat} \neq \varnothing$ and $\mathbb{G}_0^{obs} \neq \varnothing$	$\Rightarrow$	Outcome bridge
	$\mathbb{G}_0^{lat} \neq \varnothing$ and $\mathbb{H}_0^{obs} \neq \varnothing$		
Miao et al. 2018, Cui et al., 2023			

Existing Identification Strategies

Strategy	Assumptions	Representation
Our proposal	$\mathbb{H}_0^{lat} \neq \emptyset$ and $\mathbb{G}_0^{obs} \neq \emptyset$ or $\mathbb{G}_0^{lat} \neq \emptyset$ and $\mathbb{H}_0^{obs} \neq \emptyset$	1. Outcome bridge 2. Treatment bridge 3. Double Robust
Kallus et al. 2022	$\mathbb{H}_0^{lat} \neq \emptyset$ and $\mathbb{G}_0^{obs} \neq \emptyset$	Outcome bridge
	$\mathbb{G}_0^{lat} \neq \emptyset$ and $\mathbb{H}_0^{obs} \neq \emptyset$	Treatment bridge
Miao et al. 2018, Cui et al., 2023		

Existing Identification Strategies

Strategy	Assumptions	Representation
Our proposal	$\mathbb{H}_0^{lat} \neq \varnothing$ and $\mathbb{G}_0^{obs} \neq \varnothing$ or $\mathbb{G}_0^{lat} \neq \varnothing$ and $\mathbb{H}_0^{obs} \neq \varnothing$	1. Outcome bridge 2. Treatment bridge 3. Double Robust
Kallus et al. 2022	$\mathbb{H}_0^{lat} \neq \varnothing$ and $\mathbb{G}_0^{obs} \neq \varnothing$	Outcome bridge
		Double Robust
	$\mathbb{G}_0^{lat} \neq \varnothing$ and $\mathbb{H}_0^{obs} \neq \varnothing$	Treatment bridge
Miao et al. 2018, Cui et al., 2023		

Existing Identification Strategies

Strategy	Assumptions	Representation
Our proposal	$\mathbb{H}_0^{lat} \neq \varnothing$ and $\mathbb{G}_0^{obs} \neq \varnothing$ or $\mathbb{G}_0^{lat} \neq \varnothing$ and $\mathbb{H}_0^{obs} \neq \varnothing$	1. Outcome bridge 2. Treatment bridge 3. Double Robust
Kallus et al. 2022	$\mathbb{H}_0^{lat} \neq \varnothing$ and $\mathbb{G}_0^{obs} \neq \varnothing$	Outcome bridge
		Double Robust
	$\mathbb{G}_0^{lat} \neq \varnothing$ and $\mathbb{H}_0^{obs} \neq \varnothing$	Treatment bridge
Miao et al. 2018, Cui et al., 2023	$\mathbb{H}_0^{obs} \neq \varnothing$ and completeness	

Existing Identification Strategies

Strategy	Assumptions	Representation
Our proposal	$\mathbb{H}_0^{lat} \neq \varnothing$ and $\mathbb{G}_0^{obs} \neq \varnothing$ or $\mathbb{G}_0^{lat} \neq \varnothing$ and $\mathbb{H}_0^{obs} \neq \varnothing$	1. Outcome bridge 2. Treatment bridge 3. Double Robust
Kallus et al. 2022	$\mathbb{H}_0^{lat} \neq \varnothing$ and $\mathbb{G}_0^{obs} \neq \varnothing$	Outcome bridge
		Double Robust
	$\mathbb{G}_0^{lat} \neq \varnothing$ and $\mathbb{H}_0^{obs} \neq \varnothing$	Treatment bridge
Miao et al. 2018, Cui et al., 2023	$\mathbb{H}_0^{obs} \neq \varnothing$ and completeness	Outcome bridge

Existing Identification Strategies

Strategy	Assumptions	Representation
Our proposal	$\mathbb{H}_0^{lat} \neq \varnothing$ and $\mathbb{G}_0^{obs} \neq \varnothing$ or $\mathbb{G}_0^{lat} \neq \varnothing$ and $\mathbb{H}_0^{obs} \neq \varnothing$	1. Outcome bridge 2. Treatment bridge 3. Double Robust
Kallus et al. 2022	$\mathbb{H}_0^{lat} \neq \varnothing$ and $\mathbb{G}_0^{obs} \neq \varnothing$	Outcome bridge
		Double Robust
	$\mathbb{G}_0^{lat} \neq \varnothing$ and $\mathbb{H}_0^{obs} \neq \varnothing$	Treatment bridge
Miao et al. 2018, Cui et al., 2023	$\mathbb{H}_0^{obs} \neq \varnothing$ and completeness	Outcome bridge
	$\mathbb{G}_0^{obs} \neq \varnothing$ and completeness	

Existing Identification Strategies

Strategy	Assumptions		Representation
Our proposal	$\mathbb{H}_0^{lat} \neq \varnothing$ and $\mathbb{G}_0^{obs} \neq \varnothing$ or $\mathbb{G}_0^{lat} \neq \varnothing$ and $\mathbb{H}_0^{obs} \neq \varnothing$	$\rightarrow$	1. Outcome bridge 2. Treatment bridge 3. Double Robust
Kallus et al. 2022	$\mathbb{H}_0^{lat} \neq \varnothing$ and $\mathbb{G}_0^{obs} \neq \varnothing$	$\rightarrow$	Outcome bridge
			Double Robust
	$\mathbb{G}_0^{lat} \neq \varnothing$ and $\mathbb{H}_0^{obs} \neq \varnothing$	$\rightarrow$	Treatment bridge
Miao et al. 2018, Cui et al., 2023	$\mathbb{H}_0^{obs} \neq \varnothing$ and completeness	$\rightarrow$	Outcome bridge
	$\mathbb{G}_0^{obs} \neq \varnothing$ and completeness	$\rightarrow$	Treatment bridge

Existing Identification Strategies

Strategy	Assumptions	Representation
Our proposal	$\mathbb{H}_0^{lat} \neq \varnothing$ and $\mathbb{G}_0^{obs} \neq \varnothing$ or $\mathbb{G}_0^{lat} \neq \varnothing$ and $\mathbb{H}_0^{obs} \neq \varnothing$	1. Outcome bridge 2. Treatment bridge 3. Double Robust
Kallus et al. 2022	$\mathbb{H}_0^{lat} \neq \varnothing$ and $\mathbb{G}_0^{obs} \neq \varnothing$	Outcome bridge, Double Robust
	$\mathbb{G}_0^{lat} \neq \varnothing$ and $\mathbb{H}_0^{obs} \neq \varnothing$	Treatment bridge
Miao et al. 2018, Cui et al., 2023	$\mathbb{H}_0^{obs} \neq \varnothing$ and completeness	Outcome bridge, Double Robust
	$\mathbb{G}_0^{obs} \neq \varnothing$ and completeness	Treatment bridge

Asymptotic Analysis

Theorem 3: If $|Y| \leq B$ and $|\alpha_0(X, L, W)| \leq B$ for some B, and $\left\| \hat{h}^{(-k)} - h_0 \right\| = o_p(1)$ and $\left\| \hat{g}^{(-k)} - g_0 \right\| = o_p(1)$ for all $k \in \{1, \ldots, K\}$,

then

Asymptotic Analysis

Theorem 3: If $|Y| \leq B$ and $|\alpha_0(X, L, W)| \leq B$ for some B, and $\left\| \hat{h}^{(-k)} - h_0 \right\| = o_p(1)$ and $\left\| \hat{g}^{(-k)} - g_0 \right\| = o_p(1)$ for all $k \in \{1, \ldots, K\}$,

then

$$\sqrt{n}\left(\hat{\psi}^{DR}_{CF} - \psi_0\right) = \sqrt{n}\mathbb{E}_n\left[\phi(O; h_0, g_0) - \psi_0\right] + \sqrt{n}R_n + o_p(1),$$

where

$$R_n = \frac{1}{K}\sum_{k=1}^{K}\mathbb{E}\left[\left(\hat{h}^{(-k)} - h_0\right)\left(g_0 - \hat{g}^{(-k)}\right)\right].$$

Asymptotic Analysis

If $R_n = o_p\left(n^{-1/2}\right)$ and $\mathbb{E}\left[\phi^2(O; h_0, g_0)\right] < \infty$, then

$$\sqrt{n}\left(\psi_{CF}^{DR} - \psi_0\right) \xrightarrow{d} \mathcal{N}(0, \tau^2),$$

where $\tau^2 = \mathbb{E}\left\{\left[\phi(O; h_0, g_0) - \psi_0\right]^2\right\}$.

Asymptotic Analysis

If $R_n = o_p\left(n^{-1/2}\right)$ and $\mathbb{E}\left[\phi^2(O; h_0, g_0)\right] < \infty$, then

$$\sqrt{n}\left(\psi_{CF}^{DR} - \psi_0\right) \xrightarrow{d} \mathcal{N}(0,\tau^2),$$

where $\tau^2 = \mathbb{E}\left\{\left[\phi(O; h_0, g_0) - \psi_0\right]^2\right\}$.

Remark: The following condition suffices for $R_n = o_p\left(n^{-1/2}\right)$:

$$\min\left\{ \left\| \hat{h} - h_0 \right\|_2 \left\| \mathbb{E}\left[(\hat{g} - g_0)(X, Z) \,|\, X, W\right] \right\|_2, \left\| \mathbb{E}\left[(\hat{h} - h_0)(X, W) \,|\, X, Z\right] \right\|_2 \left\| \hat{g} - g_0 \right\|_2 \right\} = o_p\left(n^{-1/2}\right).$$

Estimation of the Bridge Functions

Estimation of the Bridge Functions

The following result avoids estimation of conditional mean operators:

Estimation of the Bridge Functions

The following result avoids estimation of conditional mean operators:

Lemma:[5] Suppose that $h_0(X, W)$ satisfies

Estimation of the Bridge Functions

The following result avoids estimation of conditional mean operators:

Lemma:[5] Suppose that $h_0(X, W)$ satisfies

[5]This insightful result was introduced by Bennett et al. (2019) and Muandet et al. (2020)

Estimation of the Bridge Functions

The following result avoids estimation of conditional mean operators:

Lemma:[5] Suppose that $h_0(X, W)$ satisfies

$$\mathbb{E}\left[Y \mid X, Z\right] = \mathbb{E}\left[h_0(X, W) \mid X, Z\right].$$

[5]This insightful result was introduced by Bennett et al. (2019) and Muandet et al. (2020)

Estimation of the Bridge Functions

The following result avoids estimation of conditional mean operators:

Lemma:[5] Suppose that $h_0(X, W)$ satisfies

$$\mathbb{E}\left[Y \,|\, X, Z\right] = \mathbb{E}\left[h_0(X, W) \,|\, X, Z\right].$$

Then, for any $c > 0$ and any $h(X, W)$, it holds that

[5]This insightful result was introduced by Bennett et al. (2019) and Muandet et al. (2020)

Estimation of the Bridge Functions

The following result avoids estimation of conditional mean operators:

Lemma:[5] Suppose that $h_0(X, W)$ satisfies

$$\mathbb{E}\left[Y \,|\, X, Z\right] = \mathbb{E}\left[h_0(X, W) \,|\, X, Z\right].$$

Then, for any $c > 0$ and any $h(X, W)$, it holds that

$$\max_g \mathbb{E}\left[g(X,Z)\{Y - h(X,W)\} - c\,g(X,Z)^2\right] = \frac{1}{4c}\mathbb{E}\left\{\left[\mathbb{E}\left[(h_0 - h)(X,W) \,|\, X, Z\right]\right]^2\right\}.$$

[5]This insightful result was introduced by Bennett et al. (2019) and Muandet et al. (2020)

Estimation of the Bridge Functions

The following result avoids estimation of conditional mean operators:

Lemma:[5] Suppose that $h_0(X, W)$ satisfies

$$\mathbb{E}\left[Y \mid X, Z\right] = \mathbb{E}\left[h_0(X, W) \mid X, Z\right].$$

Then, for any $c > 0$ and any $h(X, W)$, it holds that

$$\max_g \boxed{\mathbb{E}\left[g(X, Z)\{Y - h(X, W)\}\right.} \left. - c g(X, Z)^2\right] = \frac{1}{4c}\mathbb{E}\left\{\left[\mathbb{E}\left[(h_0 - h)(X, W) \mid X, Z\right]\right]^2\right\}.$$

[5]This insightful result was introduced by Bennett et al. (2019) and Muandet et al. (2020)

Estimation of the Bridge Functions

The following result avoids estimation of conditional mean operators:

Lemma:[5] Suppose that $h_0(X, W)$ satisfies

$$\mathbb{E}\left[Y \,|\, X, Z\right] = \mathbb{E}\left[h_0(X, W) \,|\, X, Z\right].$$

Then, for any $c > 0$ and any $h(X, W)$, it holds that

$$\max_g \boxed{\mathbb{E}\left[g(X, Z)\{Y - h(X, W)\}\right.} \left. - c g(X, Z)^2\right] = \frac{1}{4c} \mathbb{E}\left\{\left[\mathbb{E}\left[(h_0 - h)(X, W) \,|\, X, Z\right]\right]^2\right\}.$$

$$\frac{1}{2c} \cdot 2\mathbb{E}\left[g(X, Z) \cdot \mathbb{E}\left[Y - h(X, W) \,|\, X, Z\right]\right]$$

[5]This insightful result was introduced by Bennett et al. (2019) and Muandet et al. (2020)

Estimation of the Bridge Functions

The following result avoids estimation of conditional mean operators:

Lemma:[5] Suppose that $h_0(X, W)$ satisfies

$$\mathbb{E}\left[Y \mid X, Z\right] = \mathbb{E}\left[h_0(X, W) \mid X, Z\right].$$

Then, for any $c > 0$ and any $h(X, W)$, it holds that

$$\max_g \boxed{\mathbb{E}\left[g(X, Z)\{Y - h(X, W)\}\right.} \left. - c g(X, Z)^2\right] = \frac{1}{4c}\mathbb{E}\left\{\left[\mathbb{E}\left[(h_0 - h)(X, W) \mid X, Z\right]\right]^2\right\}.$$

[5]This insightful result was introduced by Bennett et al. (2019) and Muandet et al. (2020)

Estimation of the Bridge Functions

The following result avoids estimation of conditional mean operators:

Lemma:[5] Suppose that $h_0(X, W)$ satisfies

$$\mathbb{E}\left[Y \mid X, Z\right] = \mathbb{E}\left[h_0(X, W) \mid X, Z\right].$$

Then, for any $c > 0$ and any $h(X, W)$, it holds that

$$\max_g \mathbb{E}\left[\boxed{g(X, Z)\{Y - h(X, W)\}} - c g(X, Z)^2\right] = \frac{1}{4c}\mathbb{E}\left\{\left[\mathbb{E}\left[(h_0 - h)(X, W) \mid X, Z\right]\right]^2\right\}.$$

Consequently,

[5] This insightful result was introduced by Bennett et al. (2019) and Muandet et al. (2020)

Estimation of the Bridge Functions

The following result avoids estimation of conditional mean operators:

Lemma:[5] Suppose that $h_0(X, W)$ satisfies

$$\mathbb{E}\left[Y \mid X, Z\right] = \mathbb{E}\left[h_0(X, W) \mid X, Z\right].$$

Then, for any $c > 0$ and any $h(X, W)$, it holds that

$$\max_g \boxed{\mathbb{E}\left[g(X, Z)\{Y - h(X, W)\}\right.} \left. - c g(X, Z)^2\right] = \frac{1}{4c}\mathbb{E}\left\{\left[\mathbb{E}\left[(h_0 - h)(X, W) \mid X, Z\right]\right]^2\right\}.$$

Consequently,

$$h_0 \in \arg\min_h \max_g \mathbb{E}\left[g(X, Z)\{Y - h(X, W)\} - g(X, Z)^2\right].$$

[5]This insightful result was introduced by Bennett et al. (2019) and Muandet et al. (2020)

Kernel Integral Operator

- Let $T_{K_{\mathscr{H}}}$ denote the kernel integral operator:

$$T_{K_{\mathscr{H}}}h := \int K_{\mathscr{H}}\left(x, l, w; x', l', w'\right) h(x', l', w')p(x', l', w')dx'dl'dw' \, .$$

Kernel Integral Operator

- Let $T_{K_{\mathscr{H}}}$ denote the kernel integral operator:

$$T_{K_{\mathscr{H}}}h := \int K_{\mathscr{H}}\left(x, l, w; x', l', w'\right) h(x', l', w')p(x', l', w')dx'dl'dw' .$$

- **Assumption:** $T_{K_{\mathscr{H}}}$ is an injective Hilbert-Schmith operator.

Kernel Integral Operator

- Let $T_{K_{\mathscr{H}}}$ denote the kernel integral operator:

$$T_{K_{\mathscr{H}}}h := \int K_{\mathscr{H}}\left(x, l, w; x', l', w'\right) h(x', l', w')p(x', l', w')dx'dl'dw'\,.$$

- **Assumption:** $T_{K_{\mathscr{H}}}$ is an injective Hilbert-Schmith operator.
- **Implication:** $T_{K_{\mathscr{H}}}$ admits a SVD: eigenfunctions $(\varphi_j)_{j=1}^{\infty}$ forming an orthonormal basis of $\mathscr{L}^2(X, L, W)$ and positive eigenvalues $(\eta_j)_{j=1}^{\infty}$ converging to 0.

Kernel Integral Operator

- Let $T_{K_{\mathscr{H}}}$ denote the kernel integral operator:

$$T_{K_{\mathscr{H}}}h := \int K_{\mathscr{H}}\left(x, l, w; x', l', w'\right) h(x', l', w')p(x', l', w')dx'dl'dw' \,.$$

- **Assumption:** $T_{K_{\mathscr{H}}}$ is an injective Hilbert-Schmith operator.

- **Implication:** $T_{K_{\mathscr{H}}}$ admits a SVD: eigenfunctions $(\varphi_j)_{j=1}^{\infty}$ forming an orthonormal basis of $\mathscr{L}^2(X, L, W)$ and positive eigenvalues $(\eta_j)_{j=1}^{\infty}$ converging to 0.

- Define $T_{\mathscr{H}}^{1/2} : \mathscr{L}^2(X, L, W) \to \mathscr{H}$ as:

$$T_{\mathscr{H}}^{1/2}h := \sum_{j=1}^{\infty} \sqrt{\eta_j}\langle h, \varphi_j\rangle_2 \varphi_j$$

RKHS-specific source condition

- Let $\mathscr{T}$ and $\mathscr{T}^*$ denote the conditional expectation operators:

$$\mathscr{T}h := \mathbb{E}\left[h(X, L, W) \,|\, X, L, Z\right] \quad \text{and} \quad \mathscr{T}^*g := \mathbb{E}\left[g(X, L, Z) \,|\, X, L, W\right]$$

[6]Bennett et al. (2023) proposed a similar condition to control regularization bias of estimators of solutions to inverse problems.

RKHS-specific source condition

- Let $\mathcal{T}$ and $\mathcal{T}^*$ denote the conditional expectation operators:

$$\mathcal{T}h := \mathbb{E}\left[h(X,L,W)\,|\,X,L,Z\right] \quad \text{and} \quad \mathcal{T}^*g := \mathbb{E}\left[g(X,L,Z)\,|\,X,L,W\right]$$

- **$\beta-$source condition:**[6] There exists $w_0 \in \mathcal{L}^2(X,L,W)$ such that the minimum-norm solution h_0 satisfies

$$h_0 = T_{\mathcal{H}}^{1/2} \circ \left(\widetilde{\mathcal{T}}^* \circ \widetilde{\mathcal{T}}\right)^{\beta/2} w_0,$$

where $\widetilde{\mathcal{T}} := \frac{1}{2}\mathcal{T} \circ T_{\mathcal{H}}^{1/2}$ and $\widetilde{\mathcal{T}}^* := \frac{1}{2}T_{\mathcal{H}}^{1/2} \circ \mathcal{T}$.

[6] Bennett et al. (2023) proposed a similar condition to control regularization bias of estimators of solutions to inverse problems.

Additional assumptions on $\mathscr{H}$ and $\mathscr{G}'$

1. **Closedness of $\mathscr{H}$ with respect to $\mathscr{G}'$.**

   For any $h \in \mathscr{H}$, $\mathscr{T}(h_0 - h) \in \mathscr{G}'$.

2. **Lipschitz condition between $\mathscr{H}$ and $\mathscr{G}'$.**

   There exists $L > 0$ such that for all $h \in \mathscr{H}$, $||\mathscr{T}h||_{\mathscr{G}'} \leq L||h||_{\mathscr{H}}$

Convergence Analysis of Bridge Functions[7]

[7]Adaptation of the convergence result established by Bennett et al. (2023).

Convergence Analysis of Bridge Functions[7]

Theorem 4: Suppose h_0 satisfies an RKHS-specific β**-source condition** and the function classes $\mathscr{H}$ and $\mathscr{G}'$ satisfy some regularity conditions. Let δ_n be an upper bound on the critical radius of $\mathscr{H}$, $\mathscr{G}'$, $\mathscr{H} \cdot \mathscr{G}'$ and $Y \cdot \mathscr{G}'$.

[7]Adaptation of the convergence result established by Bennett et al. (2023).

Convergence Analysis of Bridge Functions[7]

Theorem 4: Suppose h_0 satisfies an RKHS-specific β**-source condition** and the function classes $\mathscr{H}$ and $\mathscr{G}'$ satisfy some regularity conditions. Let δ_n be an upper bound on the critical radius of $\mathscr{H}$, $\mathscr{G}'$, $\mathscr{H} \cdot \mathscr{G}'$ and $Y \cdot \mathscr{G}'$.

Then, with high probability, $\hat{h}$ satisfies

[7]Adaptation of the convergence result established by Bennett et al. (2023).

Convergence Analysis of Bridge Functions[7]

Theorem 4: Suppose h_0 satisfies an RKHS-specific β**-source condition** and the function classes $\mathscr{H}$ and $\mathscr{G}'$ satisfy some regularity conditions. Let δ_n be an upper bound on the critical radius of $\mathscr{H}$, $\mathscr{G}'$, $\mathscr{H} \cdot \mathscr{G}'$ and $Y \cdot \mathscr{G}'$.

Then, with high probability, $\hat{h}$ satisfies

$$\left\| \hat{h} - h_0 \right\|_2^2 = O\left(\frac{\delta_n^2}{\lambda_{\mathscr{H}}} + \frac{\lambda_{\mathscr{G}'}}{\lambda_{\mathscr{H}}} + \lambda_{\mathscr{H}}^{\min\{\beta,1\}} \right),$$

[7]Adaptation of the convergence result established by Bennett et al. (2023).

Convergence Analysis of Bridge Functions[7]

Theorem 4: Suppose h_0 satisfies an RKHS-specific β**-source condition** and the function classes $\mathcal{H}$ and $\mathcal{G}'$ satisfy some regularity conditions. Let δ_n be an upper bound on the critical radius of $\mathcal{H}$, $\mathcal{G}'$, $\mathcal{H} \cdot \mathcal{G}'$ and $Y \cdot \mathcal{G}'$.

Then, with high probability, $\hat{h}$ satisfies

$$\left\| \hat{h} - h_0 \right\|_2^2 = O\left(\frac{\delta_n^2}{\lambda_{\mathcal{H}}} + \frac{\lambda_{\mathcal{G}'}}{\lambda_{\mathcal{H}}} + \lambda_{\mathcal{H}}^{\min\{\beta,1\}} \right),$$

and

[7]Adaptation of the convergence result established by Bennett et al. (2023).

Convergence Analysis of Bridge Functions[7]

Theorem 4: Suppose h_0 satisfies an RKHS-specific β**-source condition** and the function classes $\mathscr{H}$ and $\mathscr{G}'$ satisfy some regularity conditions. Let δ_n be an upper bound on the critical radius of $\mathscr{H}$, $\mathscr{G}'$, $\mathscr{H} \cdot \mathscr{G}'$ and $Y \cdot \mathscr{G}'$.

Then, with high probability, $\hat{h}$ satisfies

$$\left\| \hat{h} - h_0 \right\|_2^2 = O\left(\frac{\delta_n^2}{\lambda_{\mathscr{H}}} + \frac{\lambda_{\mathscr{G}'}}{\lambda_{\mathscr{H}}} + \lambda_{\mathscr{H}}^{\min\{\beta,1\}} \right),$$

and

$$\left\| \mathbb{E}\left[\left(\hat{h} - h_0\right)(X, L, W) \,|\, X, L, Z \right] \right\|_2^2 = O\left(\delta_n^2 + \lambda_{\mathscr{G}'} + \lambda_{\mathscr{H}}^{\min\{\beta+1,2\}} \right).$$

[7]Adaptation of the convergence result established by Bennett et al. (2023).

Alternative estimation strategies

Bennett et al. 2023:

$$\hat{h}^{(-k)} = \arg\min_{h\in\mathcal{H}} \max_{g\in\mathcal{G}'} \mathbb{E}_n^{(-k)} \left[g(X,L,Z)\{Y - h(X,L,W)\} - g(X,L,Z)^2\right] + \lambda_H \|h\|_2^2.$$

Alternative estimation strategies

Bennett et al. 2023:

$$\hat{h}^{(-k)} = \arg\min_{h \in \mathscr{H}} \max_{g \in \mathscr{G}'} \mathbb{E}_n^{(-k)} \left[g(X, L, Z)\{Y - h(X, L, W)\} - g(X, L, Z)^2 \right] + \lambda_H \|h\|_2^2.$$

- <u>Challenge</u>: It requires the inversion of nearly singular matrices.

Alternative estimation strategies

Bennett et al. 2023:

$$\hat{h}^{(-k)} = \arg\min_{h \in \mathscr{H}} \max_{g \in \mathscr{G}'} \mathbb{E}_n^{(-k)} \left[g(X, L, Z)\{Y - h(X, L, W)\} - g(X, L, Z)^2 \right] + \lambda_H \|h\|_2^2.$$

- Challenge: It requires the inversion of nearly singular matrices.

Mastouri et al. 2021:

$$\hat{h}^{(-k)} = \arg\min_{h \in \mathscr{H}} \sum_{i,j=1}^{n} \left[Y_i - h(X_i, L_i, W_i) \right] K_{ij} \left[Y_j - h(X_j, L_j, W_j) \right] + \lambda_{\mathscr{H}} \|h\|_{\mathscr{H}}^2,$$

Alternative estimation strategies

Bennett et al. 2023:

$$\hat{h}^{(-k)} = \arg\min_{h\in\mathscr{H}}\max_{g\in\mathscr{G}'} \mathbb{E}_n^{(-k)}\left[g(X,L,Z)\{Y - h(X,L,W)\} - g(X,L,Z)^2\right] + \lambda_H \|h\|_2^2.$$

- Challenge: It requires the inversion of nearly singular matrices.

Mastouri et al. 2021:

$$\hat{h}^{(-k)} = \arg\min_{h\in\mathscr{H}} \sum_{i,j=1}^{n} \left[Y_i - h(X_i, L_i, W_i)\right] K_{ij} \left[Y_j - h(X_j, L_j, W_j)\right] + \lambda_{\mathscr{H}} \|h\|_{\mathscr{H}}^2,$$

- Challenge: Not applicable for estimating the treatment bridge function in the MTP setting.

Closed Form Estimators using RKHS

Notation:

- $K_{\mathscr{H}}$ and $K_{\mathscr{G}'}$ denote the kernel functions associated with the RKHS $\mathscr{H}$ and $\mathscr{G}'$
- $K_{\mathscr{H},n}$ and $K_{\mathscr{G}',n}$ denote the empirical $n \times n$ kernel matrices.

By the representer Theorem, $\hat{h}$ has the following closed form expression:

$$\hat{h}(x, l, w) = \sum_{j=1}^{n} \gamma_j K_{\mathscr{H}} \left[(x_j, l_j, w_j), (x, l, z) \right]$$

where

$$\gamma = \left(K_{\mathscr{G}',n} \Gamma K_{\mathscr{H},n} + n\lambda_{\mathscr{H}} I_n \right)^{-1} \Gamma K_{\mathscr{G}',n} \tilde{Y}_n$$

and

$$\Gamma = \frac{1}{4} \left[\frac{1}{n} K_{\mathscr{G}',n} + \lambda_{\mathscr{G}'} I_n \right]^{-1}.$$

Gaussian Kernels Guarantee Convergence Rates

Critical radius for a RKHS generated by a Gaussian kernel is

$$\delta_n = O\left(\sqrt{\frac{\log(n)}{n}}\right).$$

Taking $\lambda_{\mathscr{H}} = \delta_n$ and $\lambda_{\mathscr{G}'} = \delta_n^2$ in Theorem 4 ensures that:

$$\left\| \hat{h} - h_0 \right\|_2 = O_p\left(\sqrt{\frac{\log(n)}{n}}\right),$$

and

$$\left\| \mathbb{E}\left[\left(\hat{h} - h_0\right)(X, L, W) \,|\, X, L, Z\right] \right\|_2 = O_p\left(\frac{\log(n)}{n}\right).$$

Parameter configuration for Cross-Validation

For estimating h_0:

1. Bandwidth $\sigma^2_{\mathscr{G}'} =$ 1/4 of the median heuristic.
2. Bandwidth $\sigma^2_{\mathscr{H}} =$ median heuristic scaled by factors $c_1 \in \{1/4, 1/2, 1, 2, 4\}$.
3. $\lambda_{\mathscr{H}} = c_2 \left(\frac{\log(n)}{n}\right)^{1/2}$ with $c_2 \in \{10^{-5}, 10^{-4}, 10^{-3}, 10^{-2}, 10^{-1}\}$.
4. $\lambda_{\mathscr{G}'} = c_3 \frac{\log(n)}{n}$, with $c_3 \in \{0.1, 1, 10, 100\}$.

Similar configurations were applied for estimating g_0.

Loss Function for Cross-Validation

For estimating h_0, we use the following empirical loss function:

$$\ell_n(h) = \frac{1}{4n^2}\left[\tilde{Y}_n - h(O_i)\right]^T \left[\frac{1}{n}K_{\mathcal{G}',n} + \frac{\log(n)}{n}I_n\right]^{-1}\left[\tilde{Y}_n - h(O_i)\right]$$

Where the bandwidth $\sigma^2_{\mathcal{G}'} =$ 1/4 of the median heuristic.

As $n \to \infty$, this loss converges to the population loss

$$\ell(h) := \mathbb{E}\left\{\left[\mathbb{E}\left[(h_0 - h)(X, L, W) \,|\, X, L, Z\right]\right]^2\right\}$$

Details on the ENSEMBLE Trial

Details on the ENSEMBLE Trial

✓ The trial enrolled 17,835 vaccine recipients and 18,163 placebo recipients, all meeting the criteria of being per-protocol and baseline SARS-CoV-2 sero-negative.

Details on the ENSEMBLE Trial

✓ The trial enrolled 17,835 vaccine recipients and 18,163 placebo recipients, all meeting the criteria of being per-protocol and baseline SARS-CoV-2 sero-negative.

✓ Analyses focuses on 1,164 vaccine recipients with antibody levels at Day 29. Additionally, all placebo recipients, setting their exposure to 0 based on structural knowledge.

Details on the ENSEMBLE Trial

✓ The trial enrolled 17,835 vaccine recipients and 18,163 placebo recipients, all meeting the criteria of being per-protocol and baseline SARS-CoV-2 sero-negative.

✓ Analyses focuses on 1,164 vaccine recipients with antibody levels at Day 29. Additionally, all placebo recipients, setting their exposure to 0 based on structural knowledge.

✓ COVID-19 symptomatic infection was assessed between 7 and 210 days post Day 29.

Details on the ENSEMBLE Trial

✓ The trial enrolled 17,835 vaccine recipients and 18,163 placebo recipients, all meeting the criteria of being per-protocol and baseline SARS-CoV-2 sero-negative.

✓ Analyses focuses on 1,164 vaccine recipients with antibody levels at Day 29. Additionally, all placebo recipients, setting their exposure to 0 based on structural knowledge.

✓ COVID-19 symptomatic infection was assessed between 7 and 210 days post Day 29.

✓ During the follow-up period, there were 366 symptomatic COVID-19 cases among vaccine recipients with antibody markers measured at Day 29, and 805 cases among placebo recipients