Extending Inferences to a Target Population Without Positivity

Paul Zivich, PhD

Assistant Professor
Department of Epidemiology
University of North Carolina at Chapel Hill

November 17, 2023

Corresponding Publications

Zivich PN, Edwards JK, Lofgren ET, Cole SR, Shook-Sa BE, Lessler J. Transportability without positivity: a synthesis of statistical and simulation modeling. *Epidemiology* In-press 2023.

Zivich PN, Edwards JK, Shook-Sa BE, Lofgren ET, Lessler J, Cole SR. Synthesis estimators for positivity violations with a continuous covariate. *arXiv*:2311.09388

Acknowledgements

Jess Edwards

Bonnie Shook-Sa

Stephen Cole

Eric Lofgren

Justin Lessler

Funding: K01AI125087, R01AI157758, R01GM140564, P30AI050410
Disclaimer: content is responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

pzivich@unc.edu pzivich PausalZ@fediscience.org

Motivating problem[1]

Teleporting to 1995, a colleague asks for help addressing a question

- **Question**: should women with HIV be treated with two-drug or one-drug antiretroviral therapy (ART)?
- **Parameter**: average causal effect of two-drug versus one-drug ART on 20-week CD4 T cell count (cells/mm^3)

Two sources of data to answer this question

- AIDS Clinical Trial Group (ACTG) 175
  - Trial comparing two-drug versus one-drug ART
- Women's Interagency HIV Study (WIHS)
  - Assumed to be a random sample of the target population

[1] Inspired by the example in Dahabreh et al. (2023) *Stats in Med*

Notation

Y^a: potential outcome under action a

Y: outcome of interest, CD4 at 20 weeks
A: action, two-drug ($A = 1$) or one-drug ($A = 0$) ART
V: continuous covariate, baseline CD4
W: set of additional covariates

- Age, race, weight

R: indicator for target population ($R = 1$) or trial ($R = 0$)
$O = (R, W, V, (1 - R)A, (1 - R)Y)$

Average causal effect (ACE)

$$\psi = E[Y^1 - Y^0 \mid R = 1]$$

Identification Assumptions

$$E[Y^a|R=1] = E\left\{E[Y|A=a,V,W,R=0] \mid R=1\right\}$$

Causal consistency

$$Y_i = Y_i^a \text{ if } a = A_i$$

Action (in trial population)

$$Y^a \amalg A \mid V, W, R=0$$

$$\Pr(A=a \mid V=v, W=w, R=0) > 0 \; \forall \; f(v,w,R=0) > 0$$

Sampling (linking between populations)

$$Y^a \amalg R \mid V, W$$

$$\Pr(R=0 \mid V=v, W=w) > 0 \; \forall \; f(v,w,R=1) > 0$$

Identification Assumptions

Causal consistency

$$Y_i = Y_i^a \text{ if } a = A_i$$

Action (in trial population)

$$Y^a \amalg A \mid V, W, R = 0$$

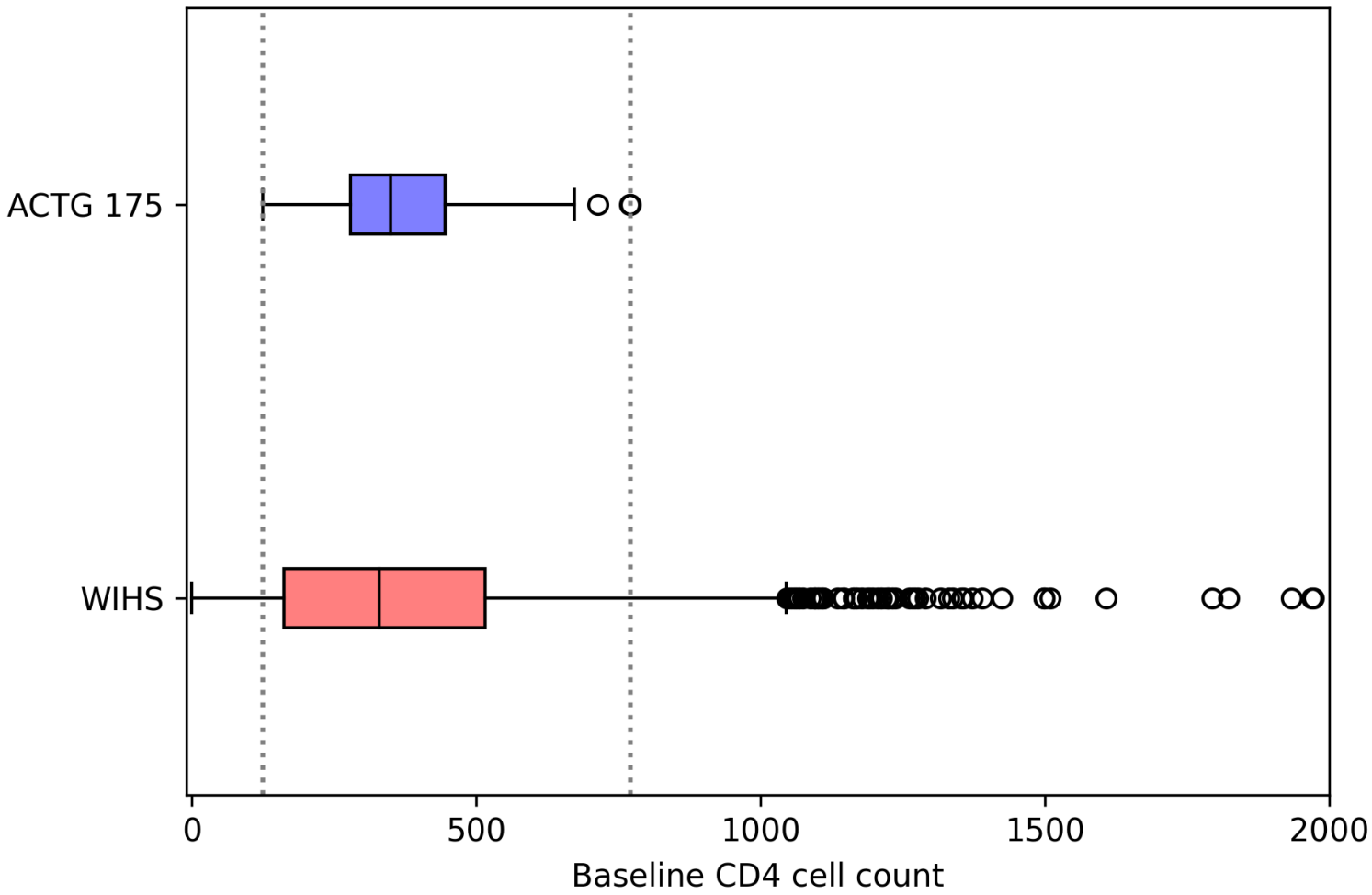
$$\Pr(A = a \mid V = v, W = w, R = 0) > 0 \; \forall \; f(v, w, R = 0) > 0$$

Sampling (linking between populations)

$$Y^a \amalg R \mid V, W$$

$$\Pr(R = 0 \mid V = v, W = w) > 0 \; \forall \; f(v, w, R = 1) > 0$$

A problem with positivity



Common solutions to non-positivity

1. Restrict the covariate set
2. Restrict the target population
3. Extrapolation

1. Restrict the covariate set

Keep parameter of interest, ψ, but modify the adjustment set

Sampling

$$Y^a \amalg R \mid W$$

Limit exchangeability to W

$$\Pr(R = 0 \mid W = w) > 0 \; \forall \; f(w, R = 1) > 0$$

No longer consider V

2. Restrict the target population

Modify the parameter of interest

$$\psi_0 = E[Y^1 - Y^0 \mid V^* = 0, R = 1]$$

where $V^* = 1 - I(v_1 \leq V \leq v_2)$

Sampling

$$Y^a \amalg R \mid V, W, V^* = 0$$

Restricting to positive region

$$\Pr(R = 0 \mid V = v, W = w) > 0 \; \forall \; f(v, w, R = 1, V^* = 0) > 0$$

Positivity for subset

3. Extrapolation

Abandon *nonparametric* identification in favor of *parametric*

- Use a parametric outcome model to extrapolate
- Requires parametric model to be valid over non-positive regions

Synthesis of statistical and mathematical models

Synthesis of statistical and mathematical models

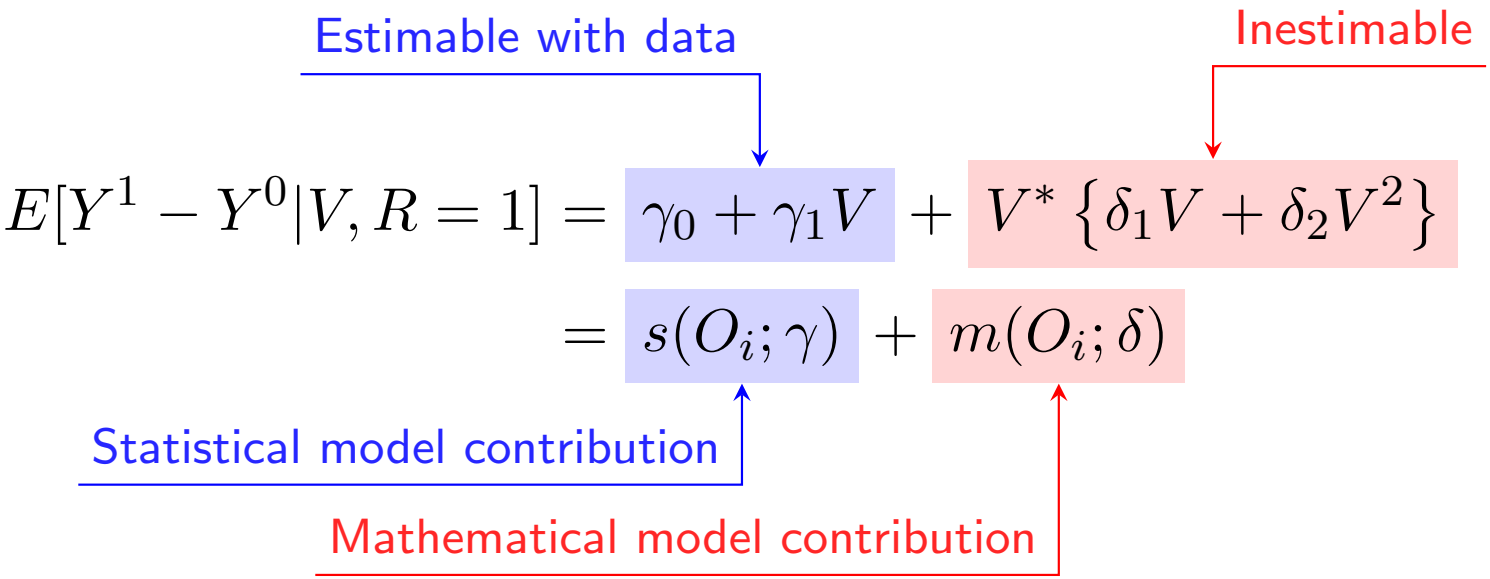
A re-expression of ψ following law of total expectation

$$\psi = \underbrace{\psi_0}_{E[Y^1 - Y^0 \mid V^* = 0, R = 1]} \Pr(V^* = 0 \mid R = 1) + \underbrace{\psi_1}_{E[Y^1 - Y^0 \mid V^* = 1, R = 1]} \Pr(V^* = 1 \mid R = 1)$$

Underlying idea: fit a statistical model for the regions with positivity, use a mathematical model to fill-in (impute) over the nonpositive region

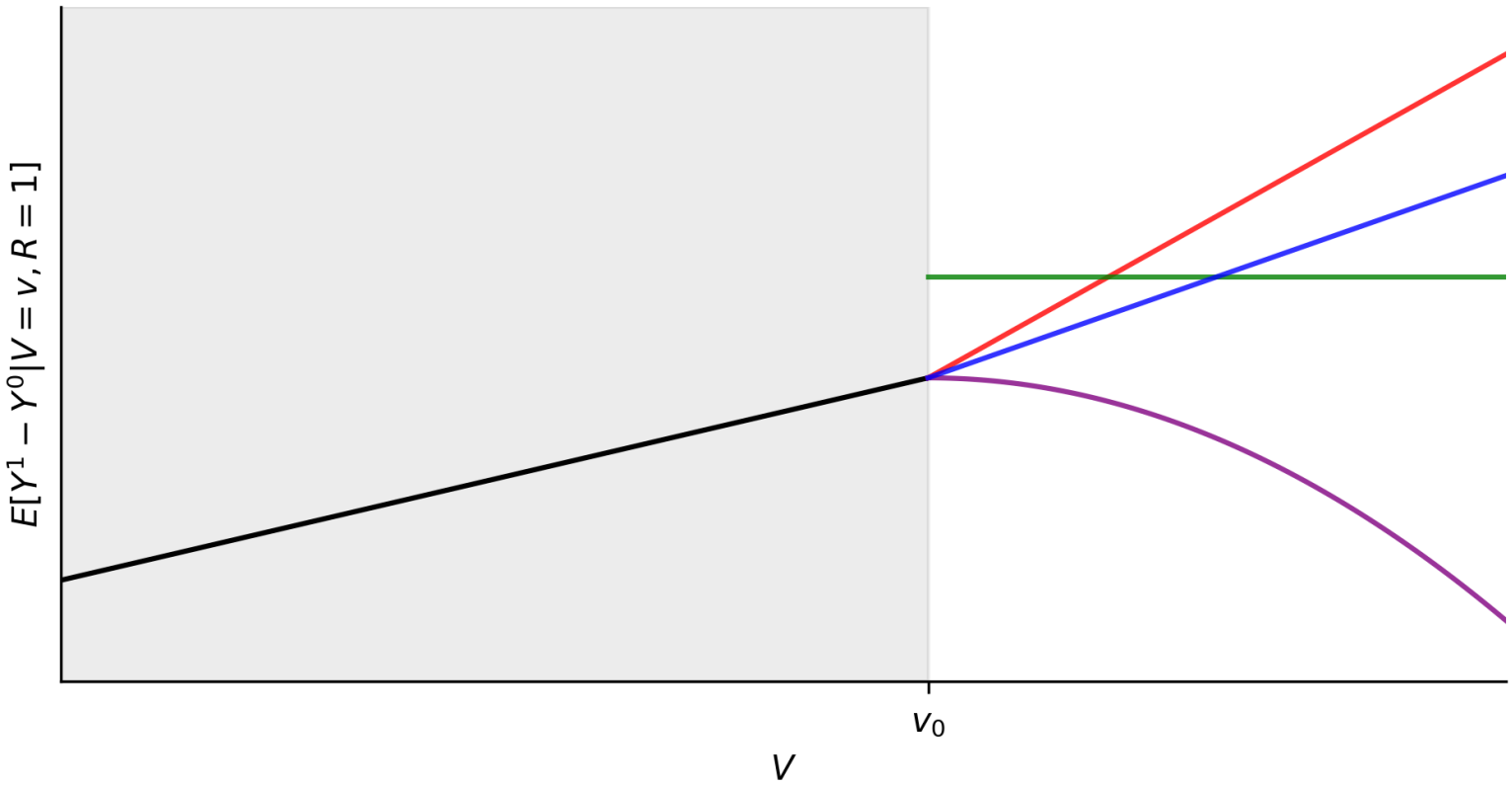
One way to combine models[2]

Model for conditional average causal effect (CACE)



[2]Other ways are considered in the *Epidemiology* and *arXiv* papers

A visualization of a synthesis CACE



Mathematical model

What do I mean by mathematical model[3]

- Mechanistic models
- Microsimulation
- Agent-based models

Informed by external information

- Studies on exposures or treatments with similar mechanisms of action, pharmacokinetic studies, animal models
- Mathematical model synthesizes this information

[3] See Roberts et al. (2012) *Med Decis Making* for general overview for constructing mathematical models

Synthesis AIPW Estimator

Estimator based on CACE model

$$\hat{\psi}_{CACE} = \frac{1}{\sum_{i=1}^{n} I(R_i = 1)} \sum_{i=1}^{n} \mathcal{G}(O_i; \hat{\gamma}, \hat{\eta}, \delta)\, I(R_i = 1)$$

where

$$E[Y^1 - Y^0 \mid V, R = 1] = \mathcal{G}(O_i; \gamma, \eta, \delta) = s(O_i; \gamma, \eta) + m(O_i; \delta)$$

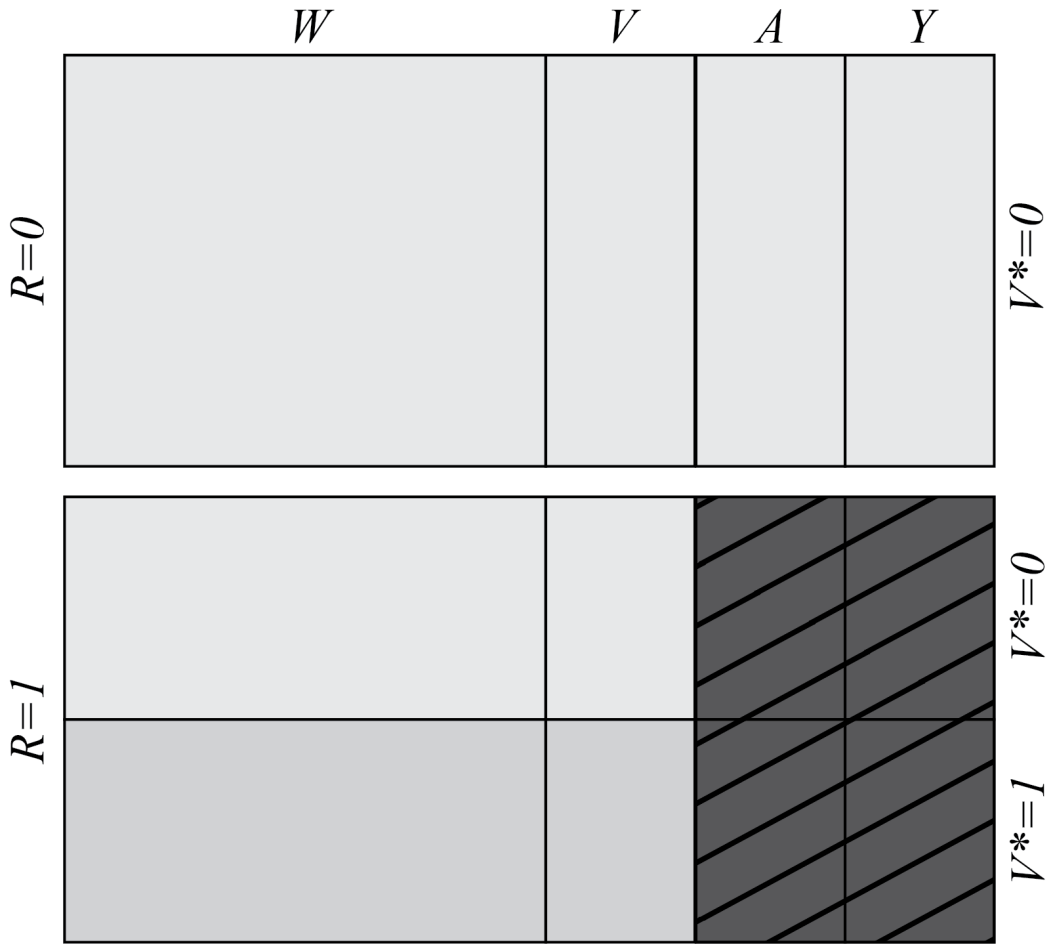
Augmented inverse probability weighting (AIPW) estimator[4]

- Weighted regression AIPW[5]

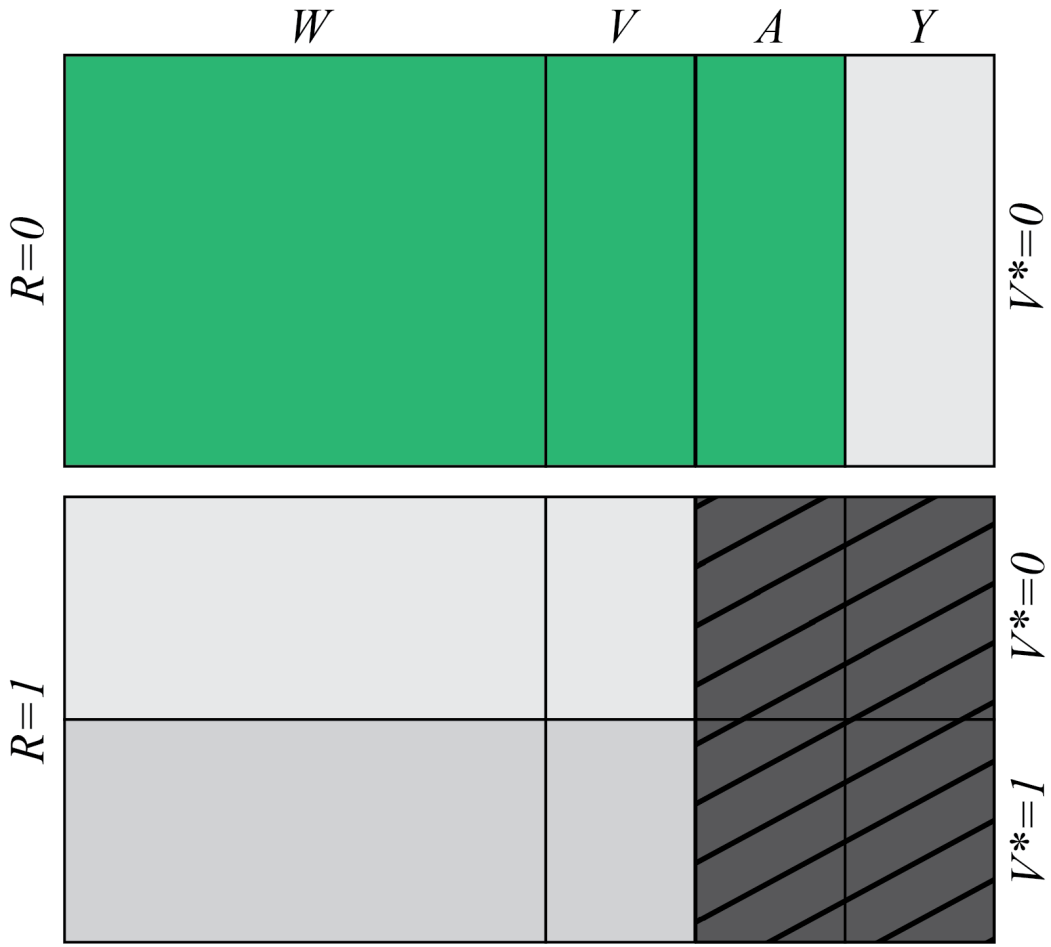
[4] Zivich et al. (2023) *Epidemiology* provides g-computation and inverse probability weighting estimators

[5] Robins et al. (2007) *Statistical Science*

Synthesis AIPW estimator

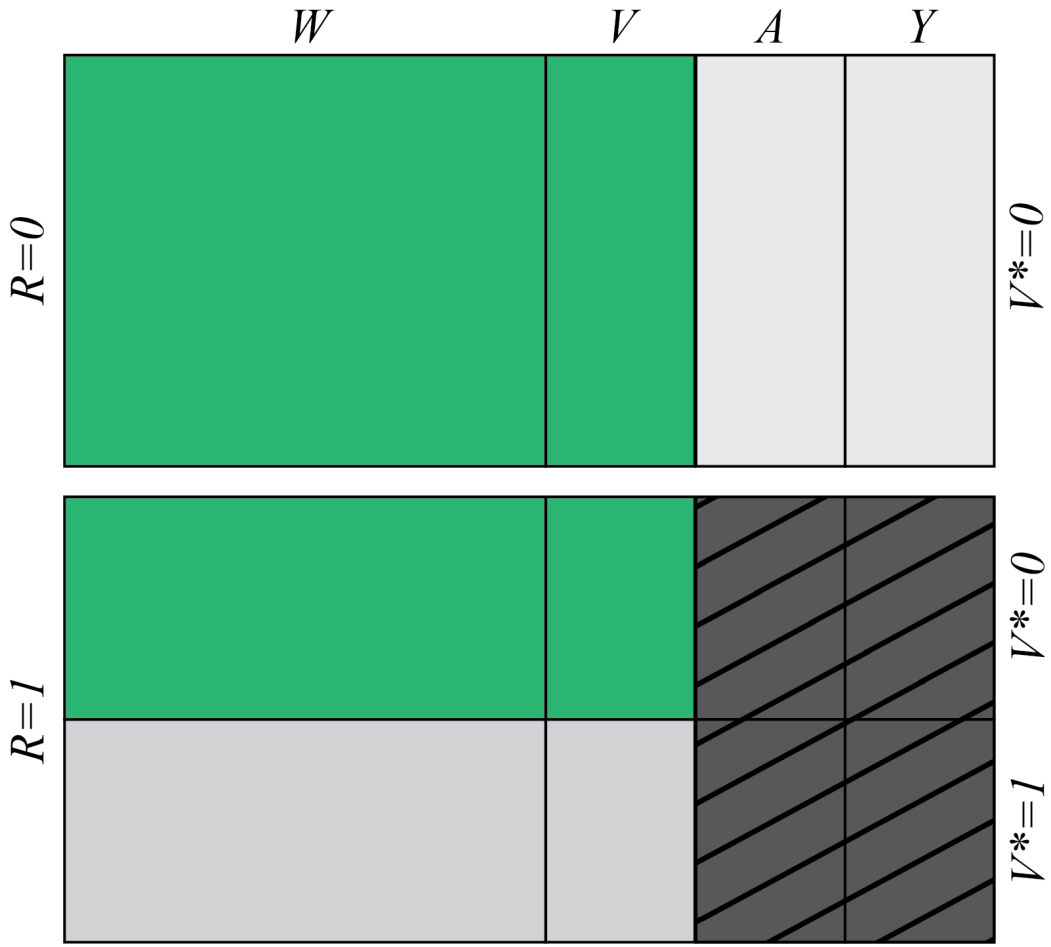


Synthesis AIPW estimator



$$\widehat{\Pr}(A = 1 \mid V, W, R = 0)$$

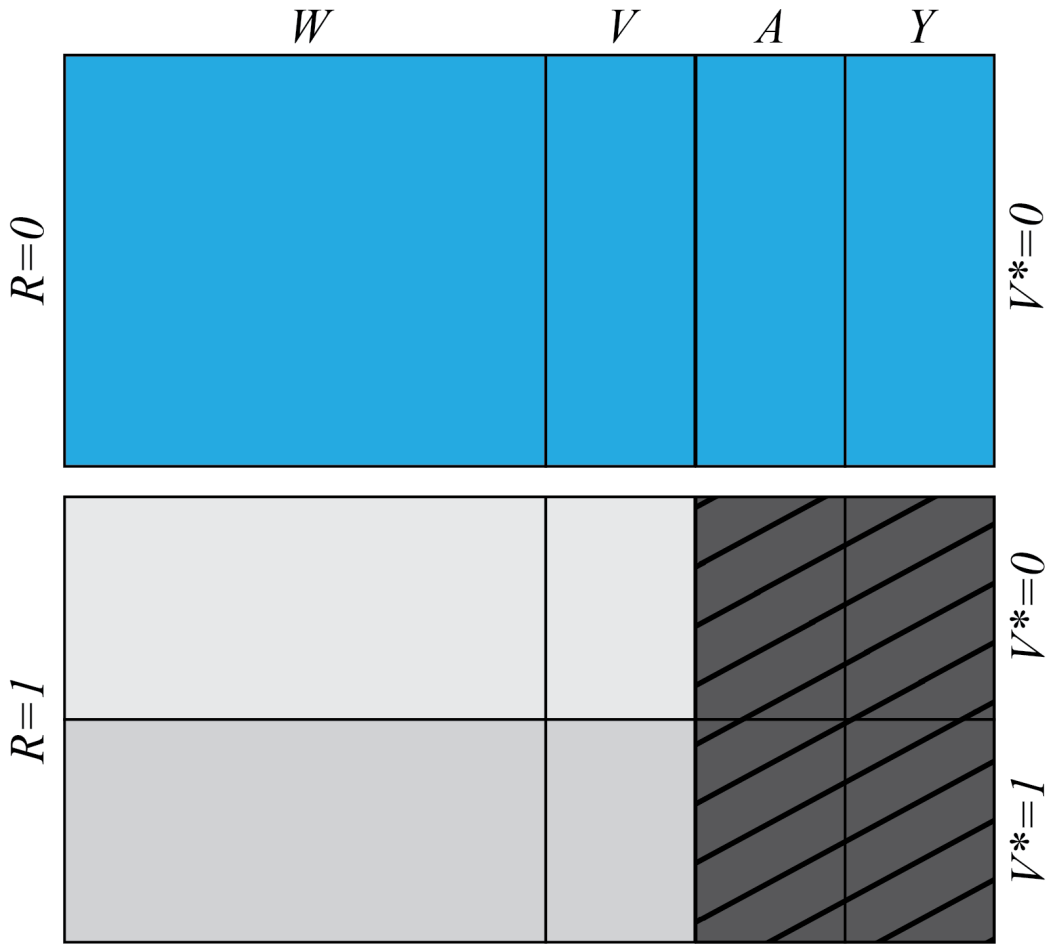
Synthesis AIPW estimator



$$\widehat{\Pr}(A = 1 \mid V, W, R = 0)$$

$$\widehat{\Pr}(R = 1 \mid V, W, V^* = 0)$$

Synthesis AIPW estimator

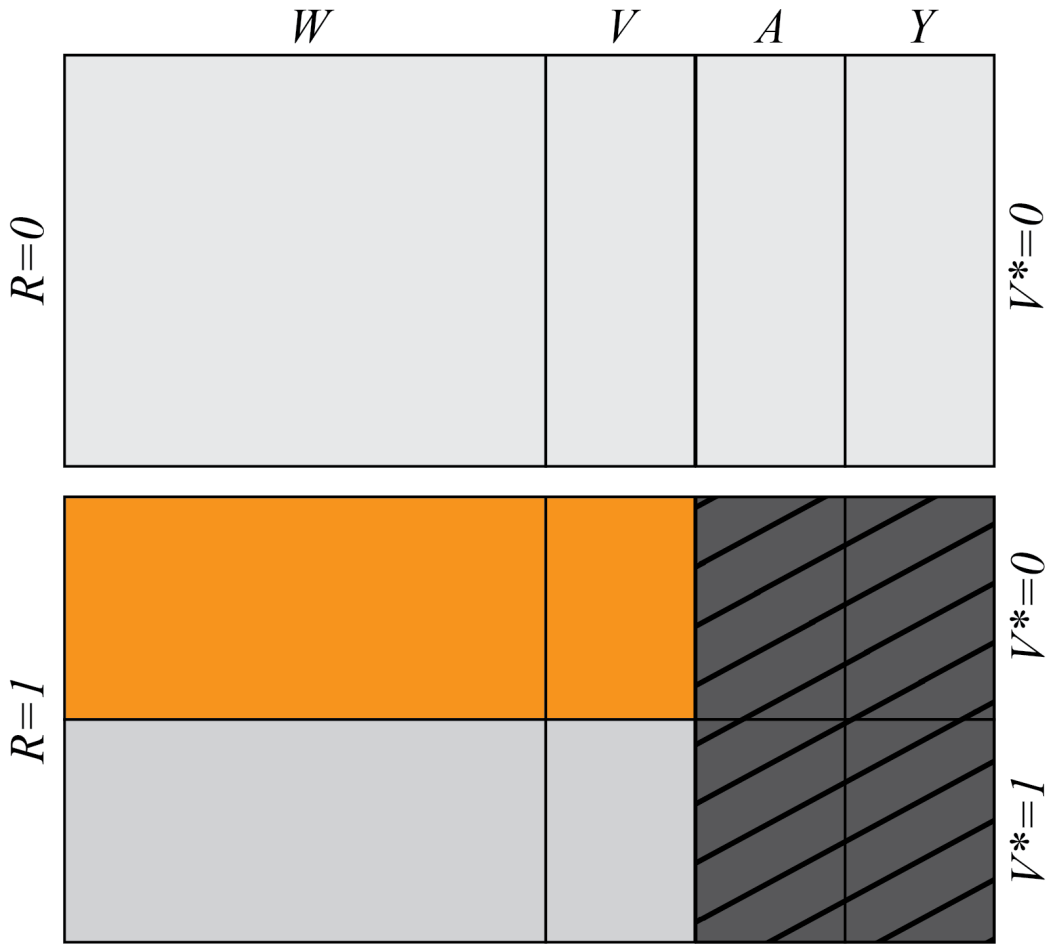


$$\widehat{\Pr}(A = 1 \mid V, W, R = 0)$$
$$\widehat{\Pr}(R = 1 \mid V, W, V^* = 0)$$

$$\downarrow$$

$$\widehat{E}[Y \mid A, V, W, R = 0]$$

Synthesis AIPW estimator



$$\widehat{\Pr}(A = 1 \mid V, W, R = 0)$$

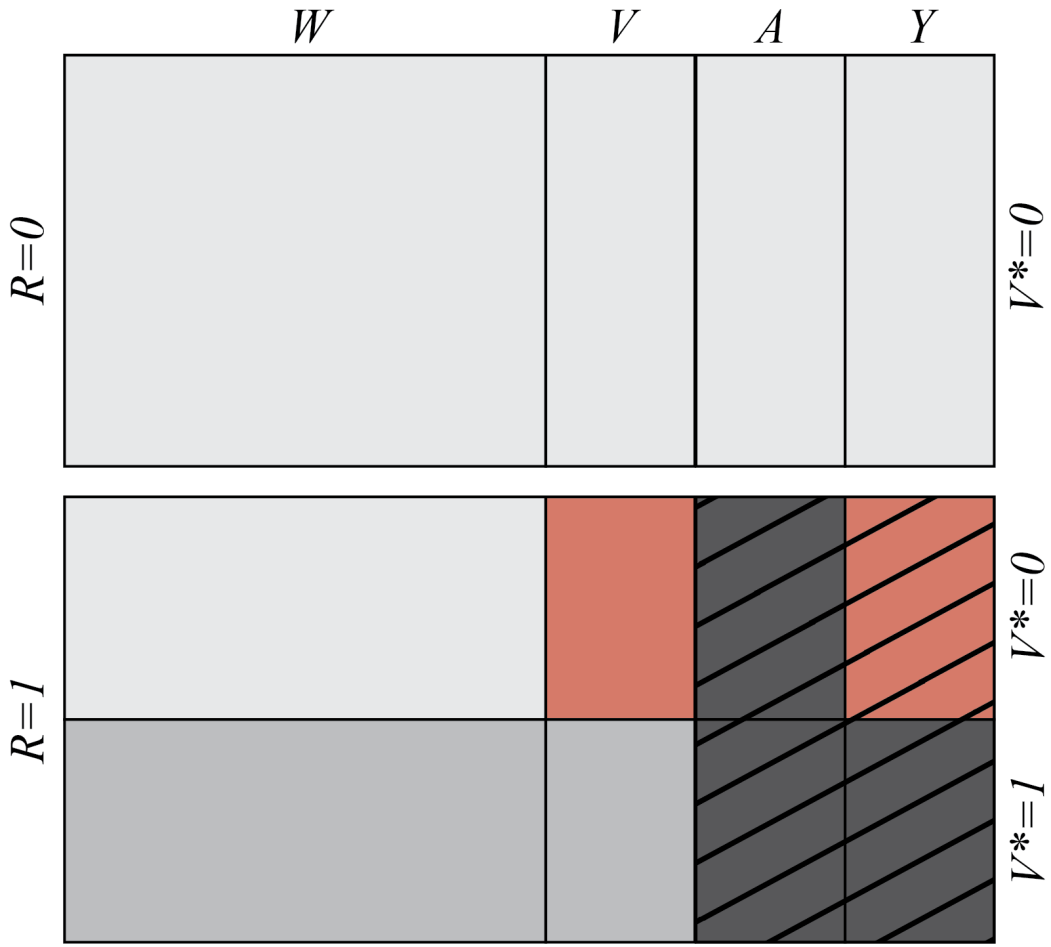
$$\widehat{\Pr}(R = 1 \mid V, W, V^* = 0)$$

$$\downarrow$$

$$\widehat{E}[Y \mid A, V, W, R = 0]$$

$$\widehat{Y}^1, \quad \widehat{Y}^0$$

Synthesis AIPW estimator



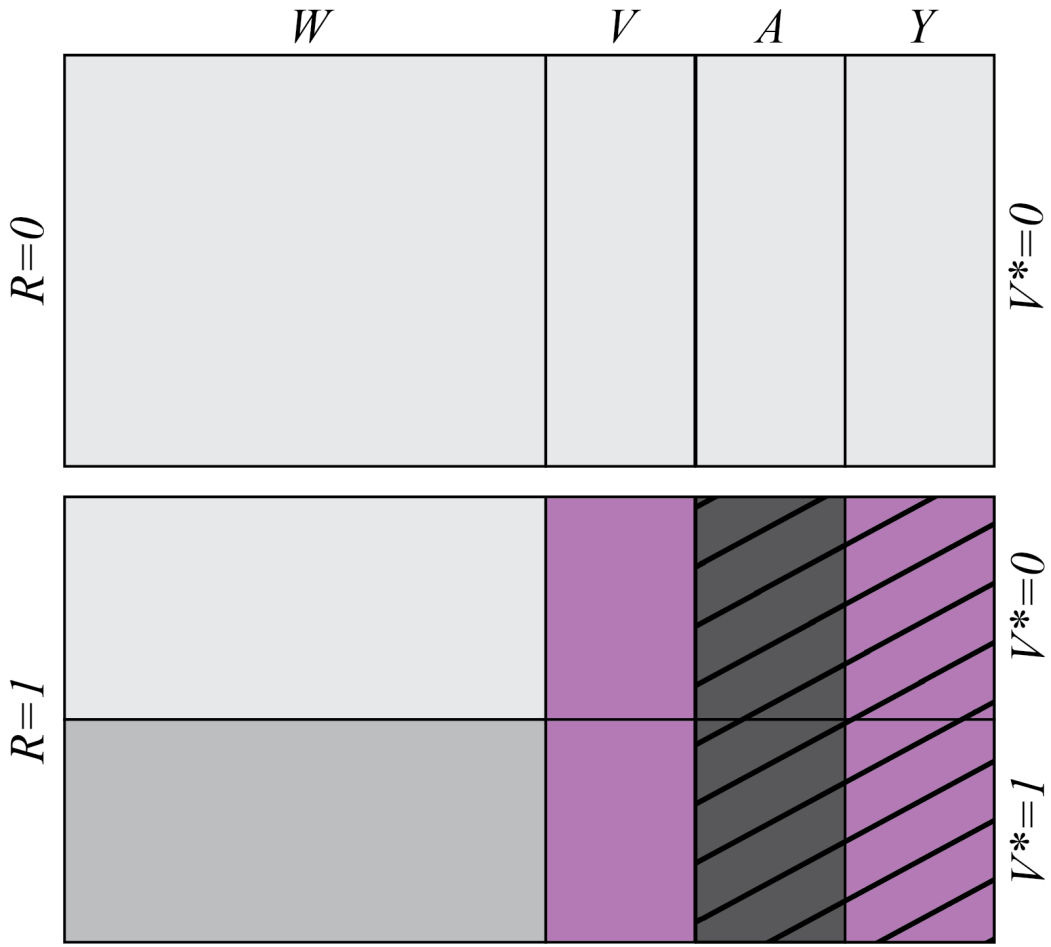
$$\widehat{\Pr}(A = 1 \mid V, W, R = 0)$$
$$\widehat{\Pr}(R = 1 \mid V, W, V^* = 0)$$

$$\widehat{E}[Y|A, V, W, R = 0]$$

$$\widehat{Y}^1, \quad \widehat{Y}^0$$

$$\widehat{E}\left[\widehat{Y}^1 - \widehat{Y}^0 | V, V^* = 0\right]$$

Synthesis AIPW estimator



$$\widehat{\Pr}(A = 1 \mid V, W, R = 0)$$
$$\widehat{\Pr}(R = 1 \mid V, W, V^* = 0)$$

$$\widehat{E}[Y|A, V, W, R = 0]$$

$$\widehat{Y}^1, \quad \widehat{Y}^0$$

$$\widehat{E}\left[\widehat{Y}^1 - \widehat{Y}^0 | V, V^* = 0\right]$$

$$m(O_i; \delta) \rightarrow \widehat{\psi}$$

Uncertainty of the Mathematical Model

Ignored uncertainty in δ

Two options

1. Range of plausible values for δ [6]
   - Bounds on ψ
2. Distribution of plausible values for δ
   - Monte Carlo procedure
   - Distribution for ψ

[6] See Vansteelandt et al. (2006) *Statistica Sinica*

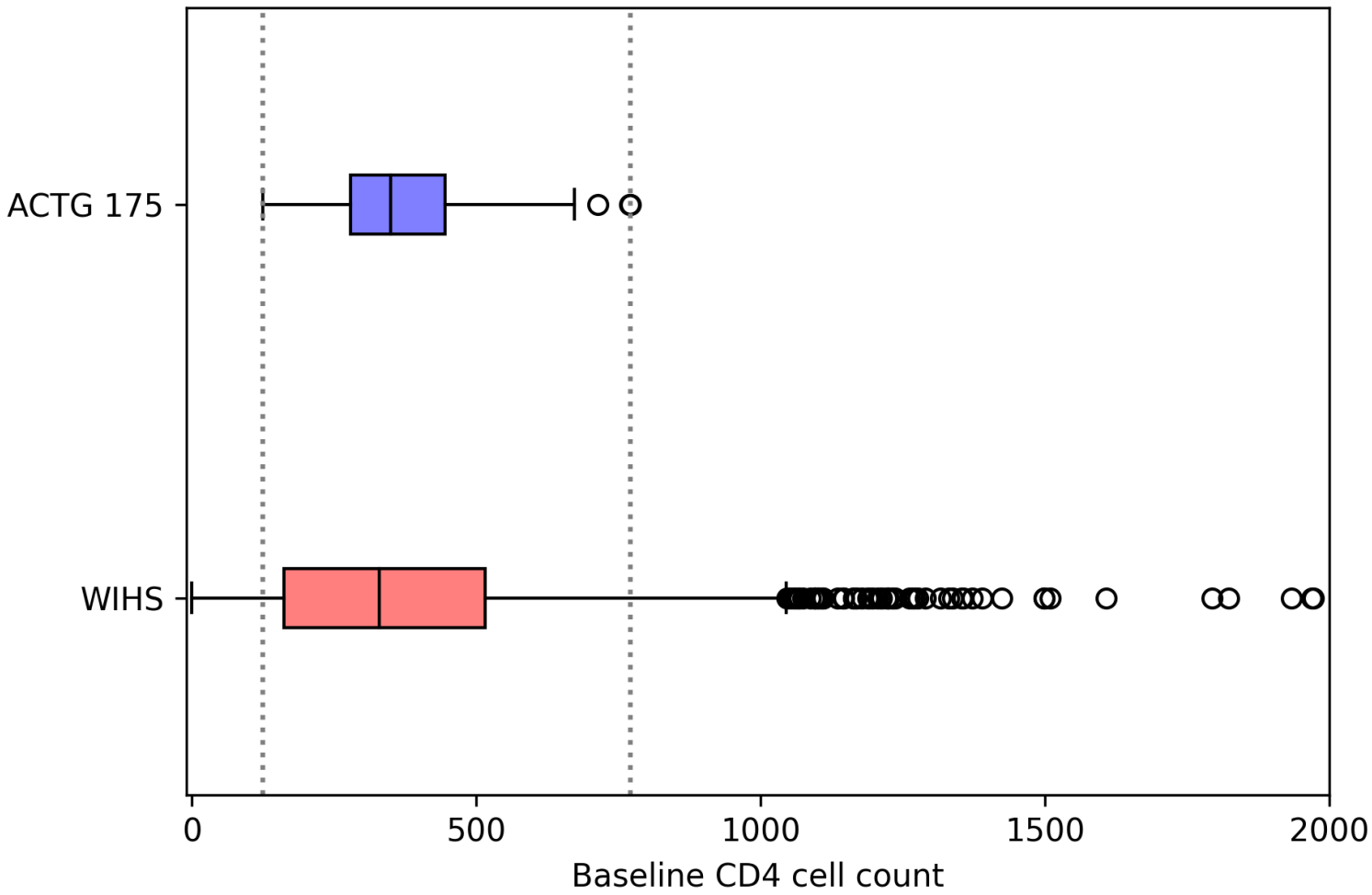
Application

Description of available data

	ACTG 175 ($n_0 = 276$)	WIHS ($n_1 = 1932$)
Age	33 [28, 39]	36 [31,41]
Baseline CD4	350 [278, 443]	330 [161, 516]
Weight (kg)	67 [59, 76]	66 [58, 78]
White	154 (56%)	390 (20%)
Two-drug ART	175 (64%)	-
CD4 20 weeks	357 [267, 480]	-

Brackets are 25^{th} and 75^{th} percentiles

A reminder of the problem



Parameter re-expression

Separating parameter into regions

$$\begin{aligned}\psi = \psi_l \ &\Pr(V < 124 | R = 1) \\ + \ \psi_m \ &\Pr(124 \leq V \leq 771 | R = 1) \\ + \ \psi_u \ &\Pr(V > 771 | R = 1)\end{aligned}$$

Synthesis model for all regions

$$\begin{aligned}\mathcal{G}(O_i; \gamma, \eta, \delta) = \delta_1 \ &I(V_i < 124) \\ + \ s(O_i; \gamma, \eta) \ &I(124 \leq V_i \leq 771) \\ + \ \delta_2 \ &I(V_i > 771)\end{aligned}$$

Mathematical model

Contemporaneous information from pharmacokinetic studies[7]

- Lower bound[8]
  - Don't expect two-drug to result in lower CD4 compared to one-drug
  - Lowest CACE would be in nonpositive regions is zero
  - $\delta_1 = \delta_2 = -20$
  - Mild antagonistic interaction between drugs
- Upper bound
  - $\delta_1 = 150$ based on largest increases observed in small-scale studies[9]
  - $\delta_2 = 100$ since no studies available (less stark but still beneficial)

[7] Wilde & Langtry *Drugs* (1993)
[8] Meng et al. *Ann Intern Med* (1992)
[9] Collier et al. (1993) *Ann Intern Med*

Statistical model

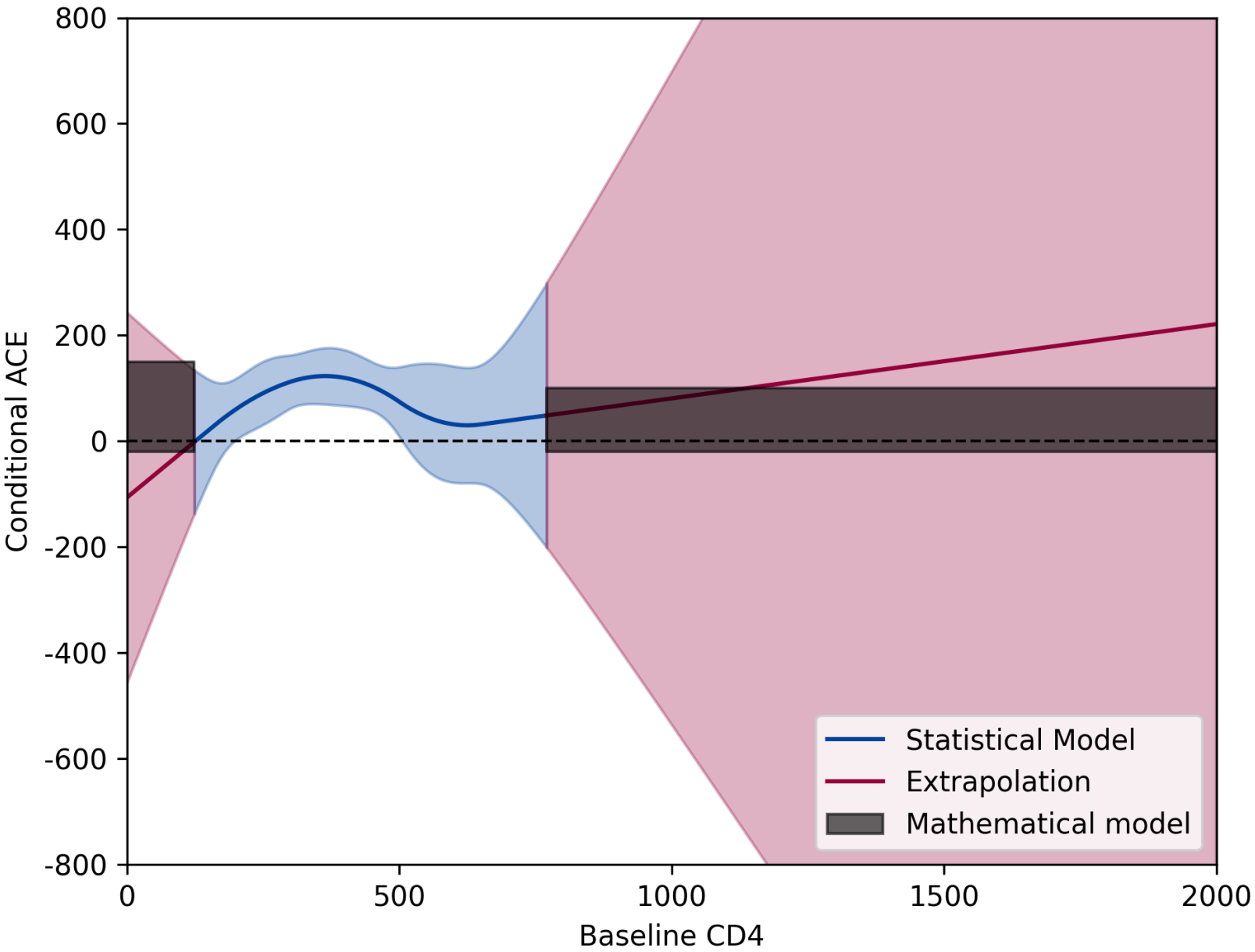
Conditional Average Causal Effect

- Weighted regression AIPW

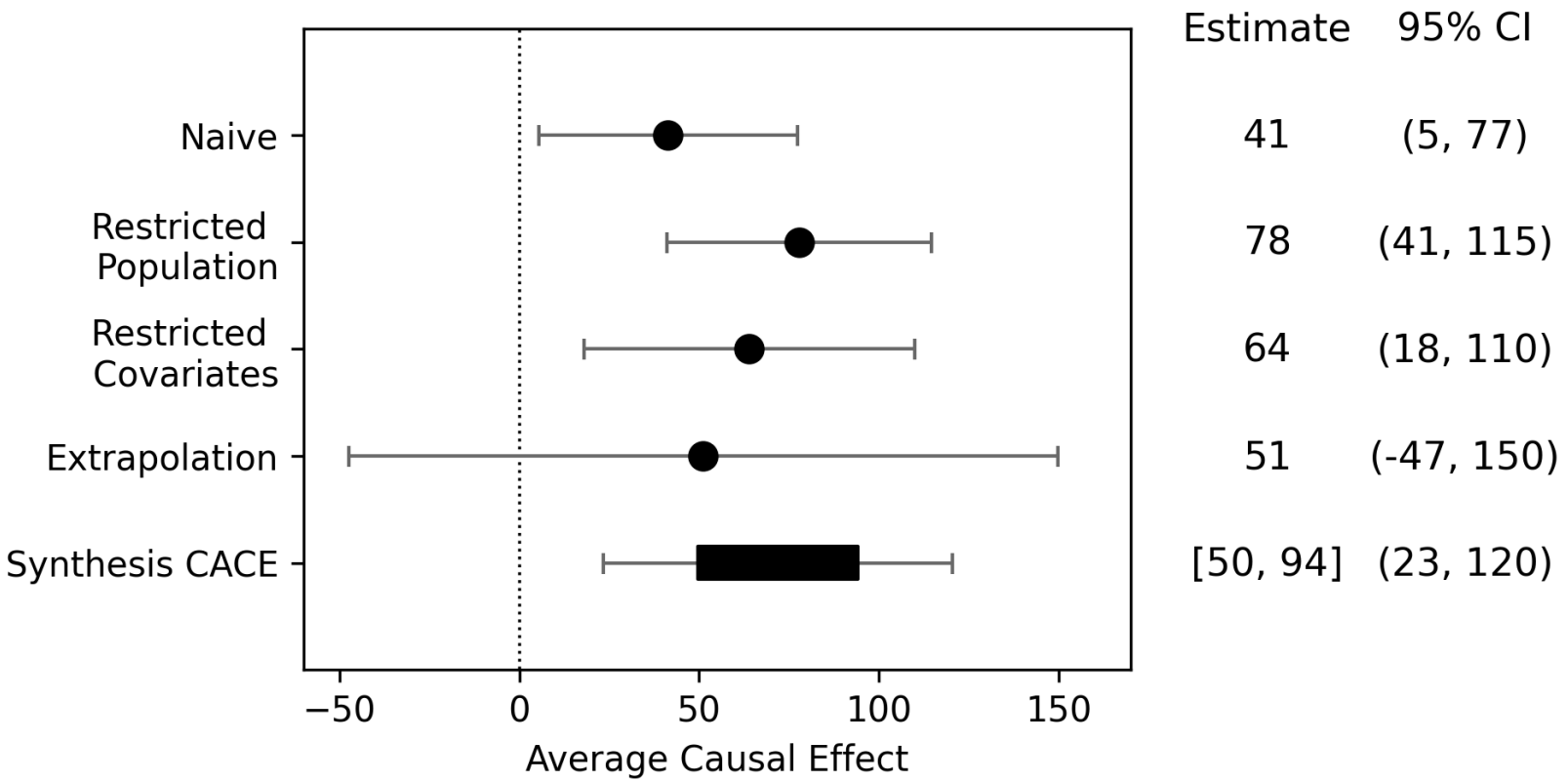
Functional forms

- Restricted quadratic splines (age, weight, baseline CD4)
  - All models
- Baseline CD4 & ART interaction terms
  - Outcome model

Estimated CACE



Results



Difference in CD4 at 20-weeks comparing two-drug to one-drug ART (higher is better)

Conclusions

Extension of inferences between populations without positivity

- Integrate external information sources
- Advantages over existing approaches

Future areas for work

- Other uses of statistical and mathematical models
  - Exchangeability paired with positivity
- Alternative estimators
- Make mathematical models more robust and reliable
  - Sensitivity analyses, diagnostics

Thanks!

Zivich PN, Edwards JK, Lofgren ET, Cole SR, Shook-Sa BE, Lessler J. Transportability without positivity: a synthesis of statistical and simulation modeling. *Epidemiology* In-press 2023.

Zivich PN, Edwards JK, Shook-Sa BE, Lofgren ET, Lessler J, Cole SR. Synthesis estimators for positivity violations with a continuous covariate. *arXiv*:2311.09388

pzivich@unc.edu pzivich PausalZ@fediscience.org

Appendix

A synthesis AIPW estimator[10]

Weighted regression AIPW as estimating equations

$$\sum_{i=1}^{n} \begin{bmatrix} (1-R_i)\left[A_i - \text{expit}(\mathbb{Z}_i\hat{\eta}_1^T)\right]\mathbb{Z}_i^T \\ (1-V_i^*)\left[R_i - \text{expit}(\mathbb{U}_i\hat{\eta}_2^T)\right]\mathbb{U}_i^T \\ (1-R_i)\pi(V_i, W_i; \hat{\eta}_1, \hat{\eta}_2)\left[Y_i - \mathbb{X}_i\hat{\eta}_3^T\right]\mathbb{X}_i^T \\ R_i(1-V_i^*)\left[(\hat{Y}_i^1 - \hat{Y}_i^0) - \mathbb{V}_i\hat{\gamma}^T\right]\mathbb{V}_i^T \\ \left(\mathbb{V}_i\hat{\gamma}^T + \mathbb{V}_i^*\delta^T\right) - \hat{\psi} \end{bmatrix} = 0$$

- $\mathbb{Z}, \mathbb{U}, \mathbb{X}, \mathbb{V}, \mathbb{V}^*$ are design matrices

[10] Estimating equations are solved using `delicatessen`, arXiv:2203.11300

Simulations

$$V \sim 375 \times \text{Weibull}(1, 1.5)$$

$$W \sim \text{Bernoulli}(0.2)$$

$$\Pr(R = 0|V, W) = \begin{cases} \text{expit}(-0.02V + 2W) & V \leq 300 \\ 0 & V > 300 \end{cases}$$

$$\Pr(A = 1|R = 0) = 0.5$$

Sample sizes

- $n_1 = 1000, n_0 = 500$
- $n_1 = 1000, n_0 = 1000$

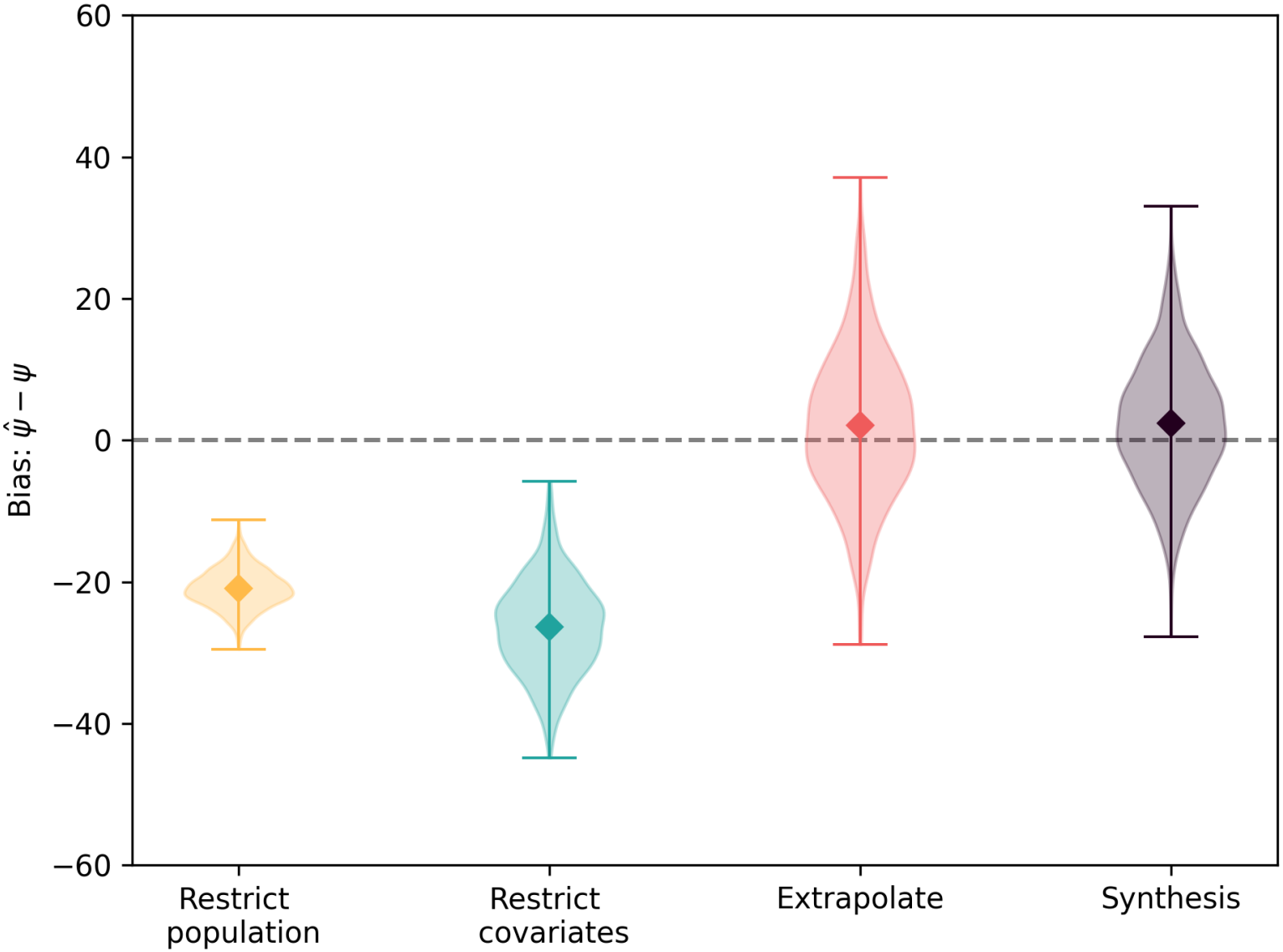
Scenario 1: Setup

$$Y^a = -20 + 70a + V + 0.12aV - 2W + 5aW + \epsilon$$

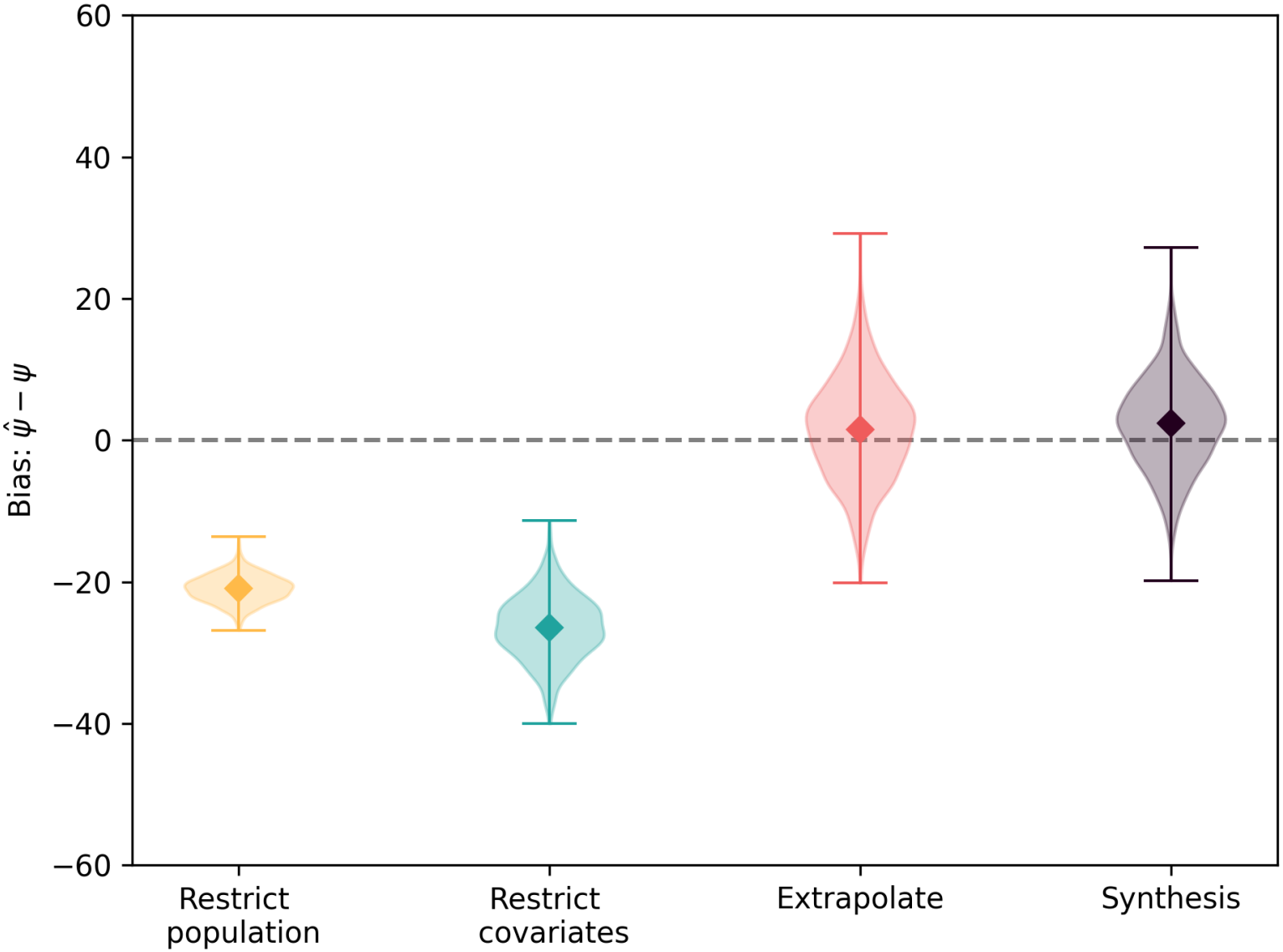
Relationship between Y^a and V doesn't change over V^*

- Extrapolation approach expected to be valid
- Synthesis with valid parameters expected to be valid
- Others are not

Scenario 1: Results, $n_1 = 1000, n_0 = 500$



Scenario 1: Results, $n_1 = 1000, n_0 = 1000$



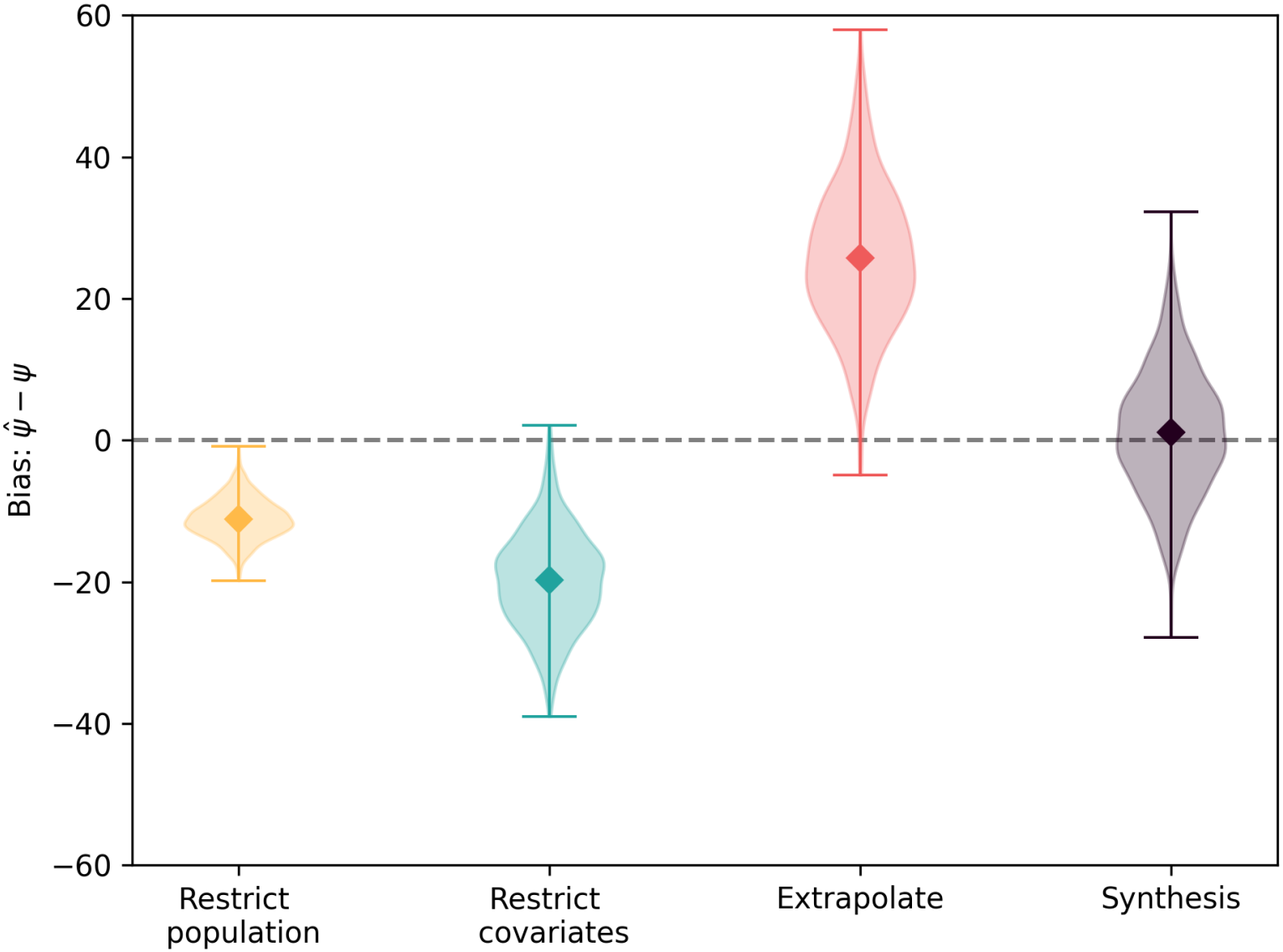
Scenario 2: Setup

$$\begin{aligned} Y^a = & -20 + 70a + V + 0.12aV - 2W + 5aW \\ & - 0.2a\{V - 300\}I(V > 300) - 0.3a\{V - 800\}I(V > 800) \\ & + \epsilon \end{aligned}$$

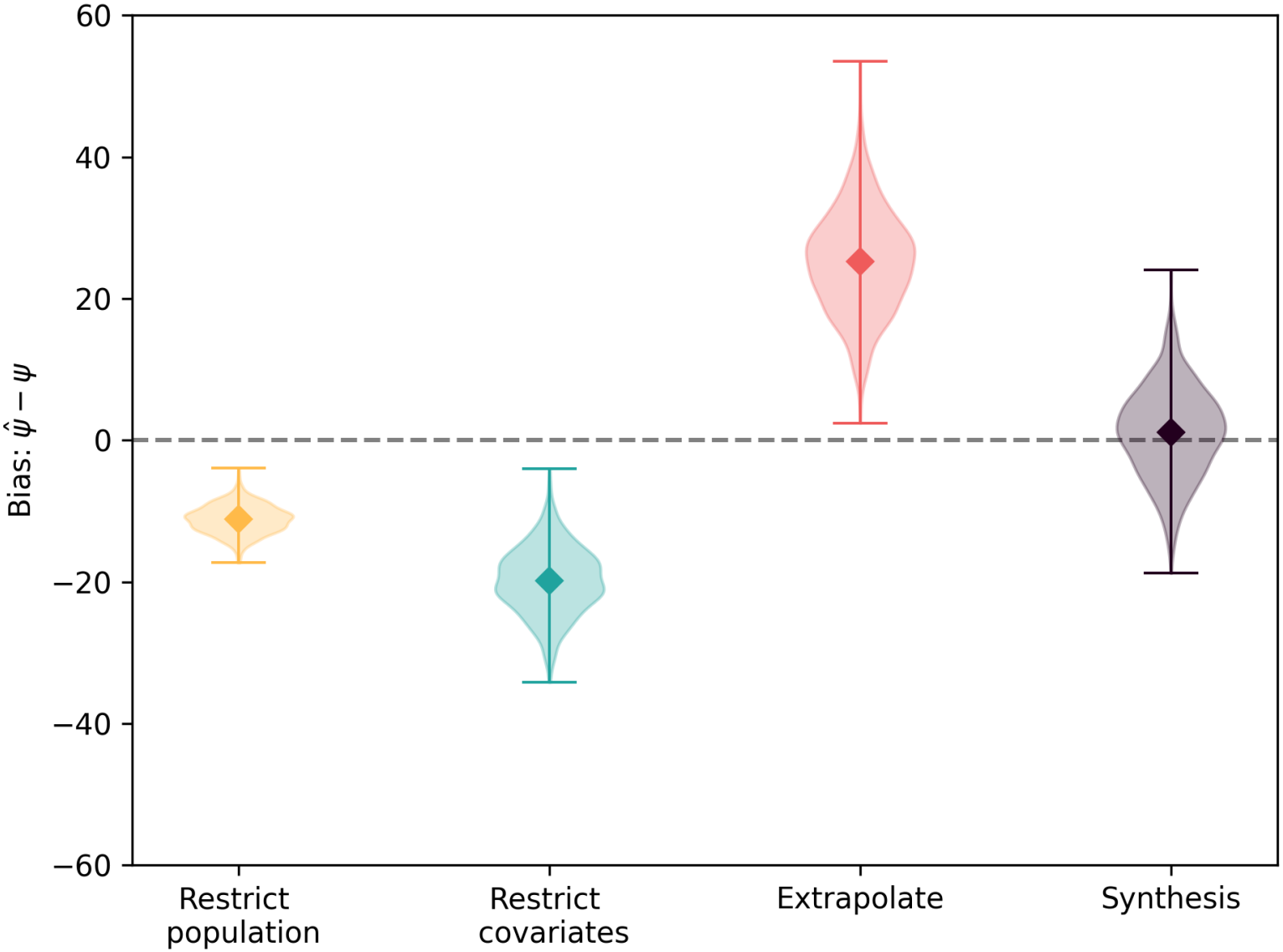
Relationship between Y^a and V changes in $V^* = 1$

- Synthesis with valid parameters expected to be valid
- Others are not

Scenario 2: Results, $n_1 = 1000, n_0 = 500$



Scenario 2: Results, $n_1 = 1000, n_0 = 1000$



Other results

In the pre-print, other items considered

- Different mathematical model parameter specifications
- Alternative estimator based on marginal structural models (MSMs)