Improving Modern Techniques of Causal Inference:

Finite Sample Performance of ATM and ATO Doubly Robust Estimators, Variance Estimation for ATO Estimators, and Contextualized Tipping Point Sensitivity Analyses for Unmeasured Confounding

Lucy D'Agostino McGowan

DISSERTATION DEFENSE

Thank you!

What method should I use?

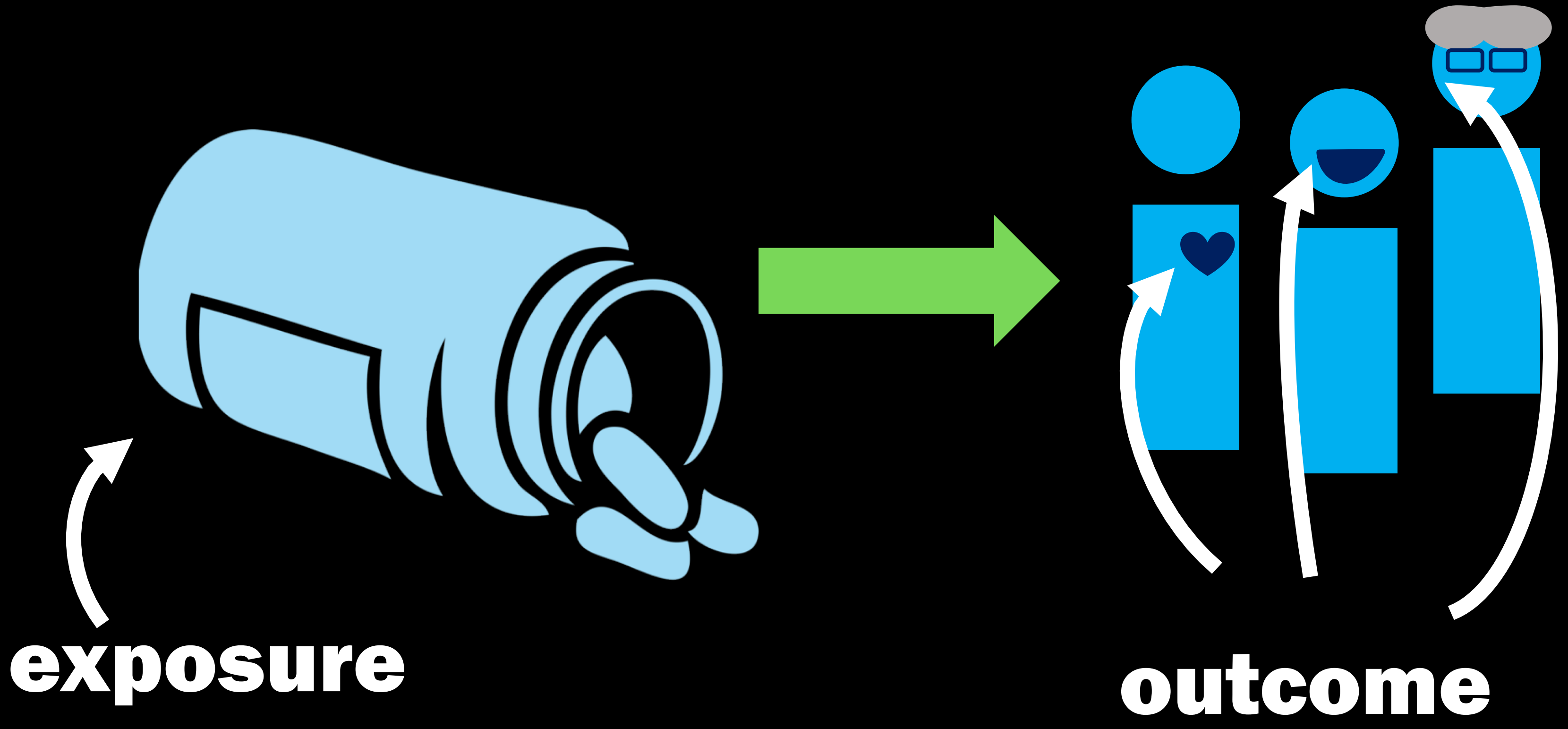
How should I estimate my effect and variance?

What if I've missed something?

Exploring finite-sample bias in propensity score weighting choices

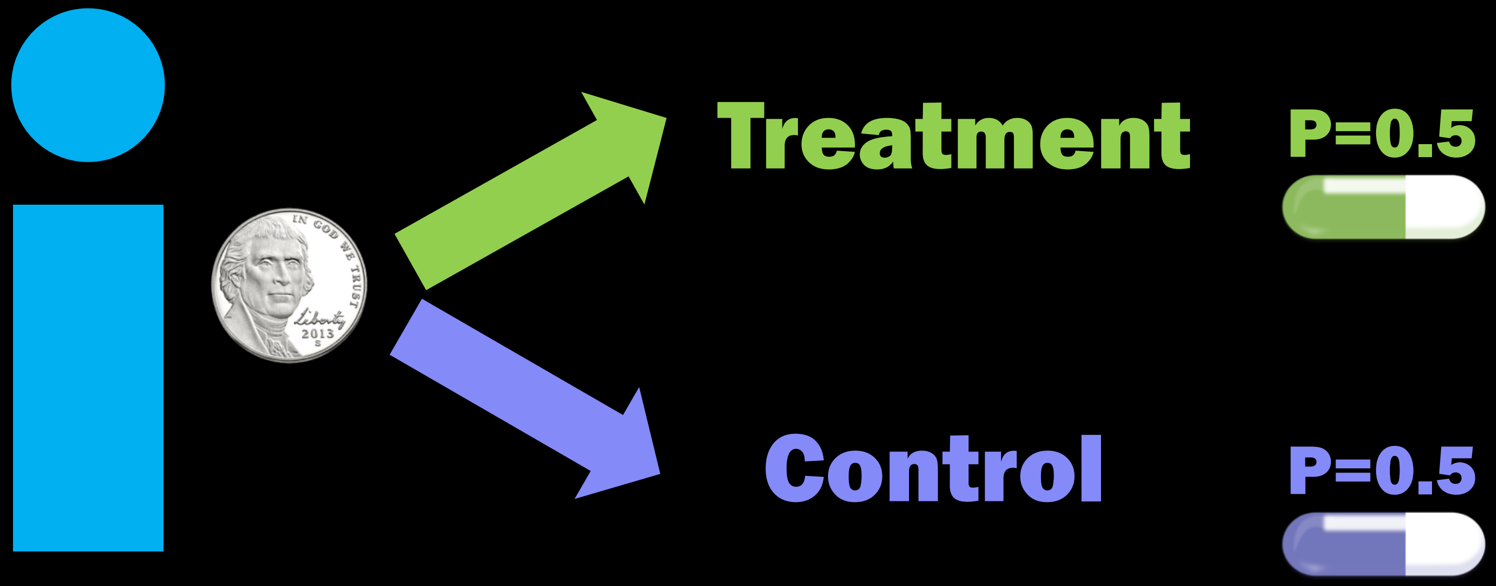
DISSERTATION DEFENSE

Observational Studies

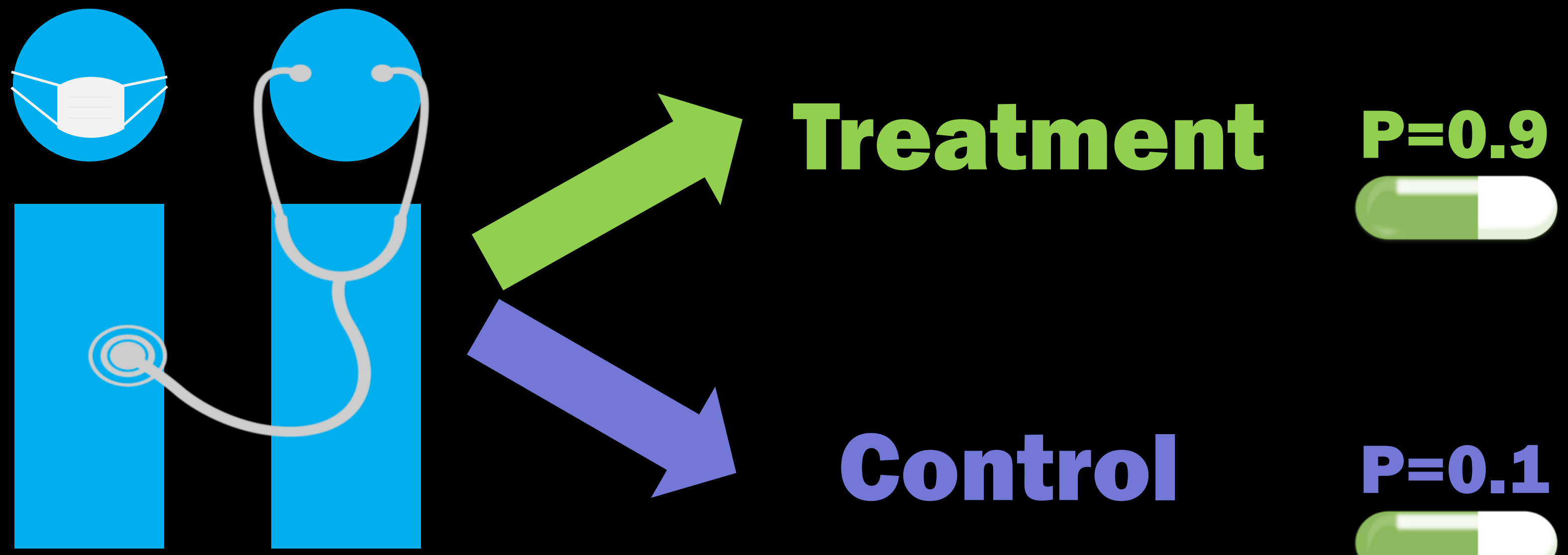


~~Observational Studies~~

Randomized Controlled Trials



Observational Studies



Potential outcomes

$$Y_i(0) \quad - \quad Y_i(1)$$

Potential outcomes

$Y_i(0)$ $Y_i(1)$

Potential outcomes

$$Y_i(0) \qquad Y_i(1)$$

Potential outcomes

$$Y_i(0) \quad - \quad Y_i(1)$$

Causal Quantities

Average treatment effect

$$E[Y(1) - Y(0)]$$

Average treatment effect

$$\frac{1}{n_Z}\sum_{i=1}^{n} Y_i Z_i - \frac{1}{n_{1-Z}}\sum_{i=1}^{n} Y_i(1 - Z_i)$$

No unmeasured confounders

$$\big(Y(0), Y(1)\big) \perp Z|\mathbf{X}$$

Propensity score

$$e_i = P(Z_i = 1|\mathbf{X})$$

Average treatment effect

$$w_{ATE} = \frac{Z_i}{e_i} + \frac{1 - Z_i}{1 - e_i}$$

Average treatment effect

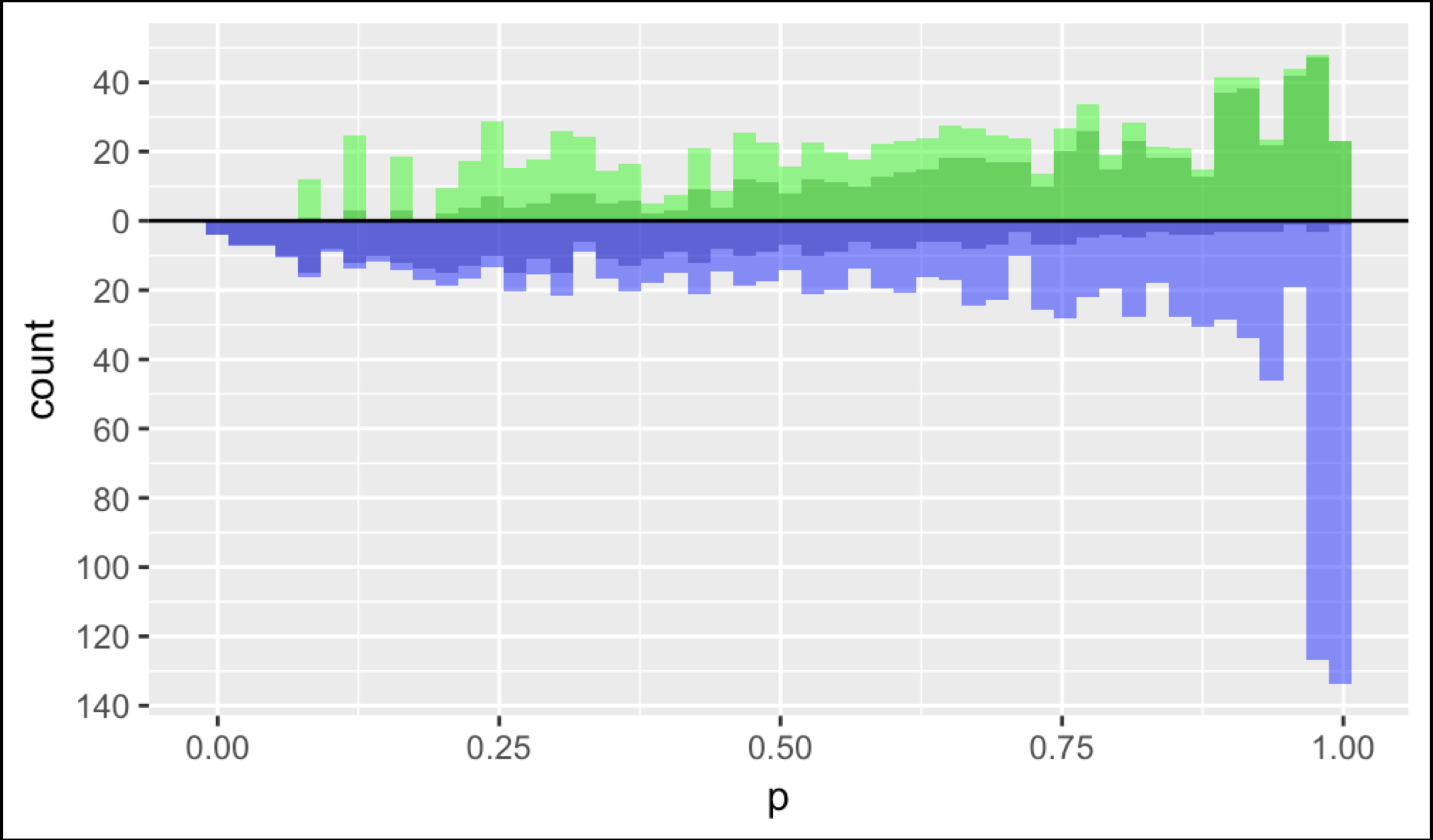
$$w_{ATE} = \frac{Z_i}{e_i} + \frac{1 - Z_i}{1 - e_i}$$

Average treatment effect

$$w_{ATE} = \frac{Z_i}{e_i} + \frac{1 - Z_i}{1 - e_i}$$

Average treatment effect

$$\frac{\sum_{i=1}^{n} Y_i Z_i w_i}{\sum_{i=1}^{n} Z_i w_i} - \frac{\sum_{i=1}^{n} Y_i (1 - Z_i) w_i}{\sum_{i=1}^{n} (1 - Z_i) w_i}$$



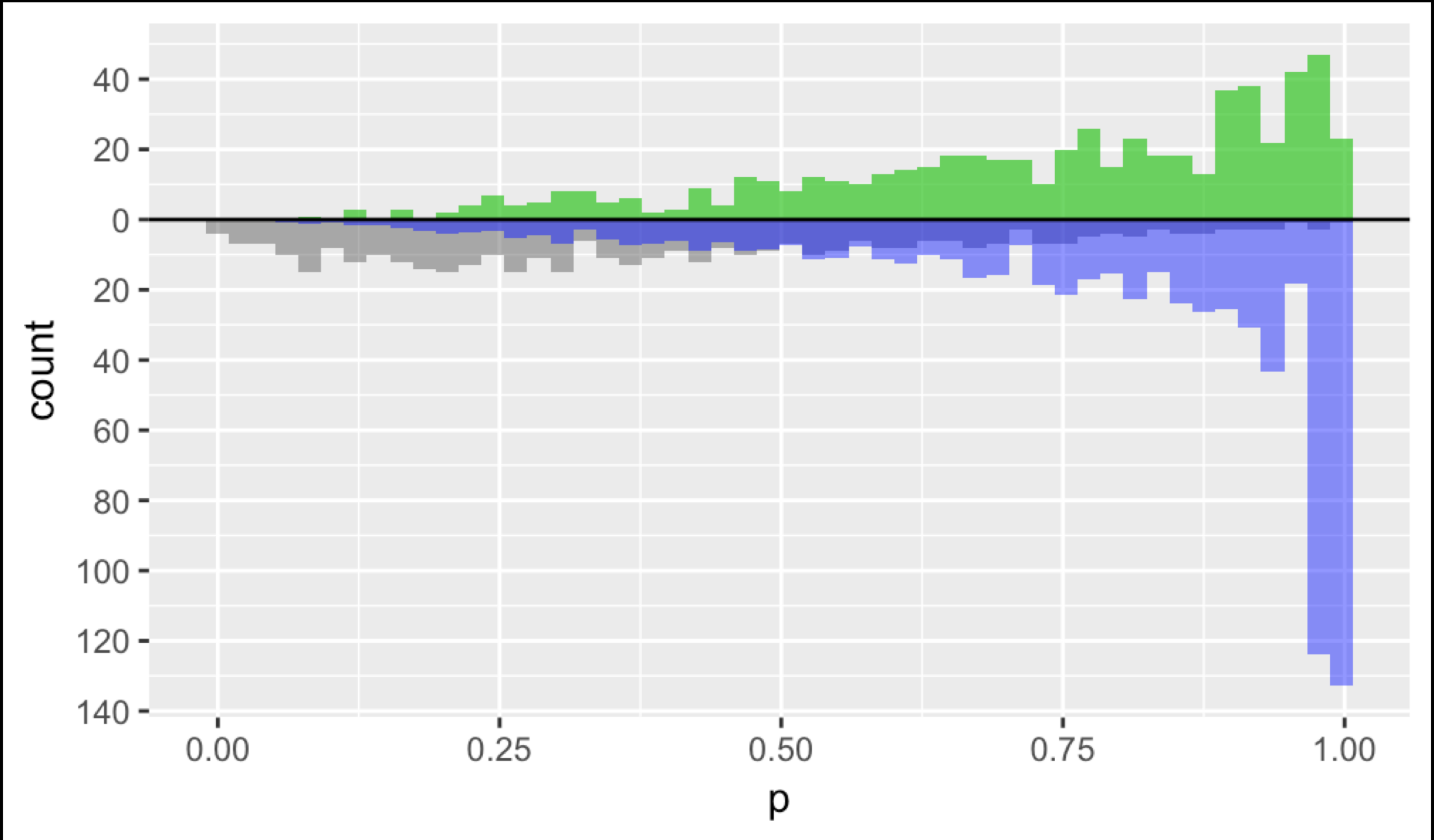
count
40
20
0
20
40
60
80
100
120
140
0.00
0.25
0.50
0.75
1.00
p

Average treatment effect among the treated

$$E[Y(1) - Y(0)|Z = 1]$$

Average treatment effect among the treated

$$w_{ATT} = \frac{e_i Z_i}{e_i} + \frac{e_i(1 - Z_i)}{1 - e_i}$$



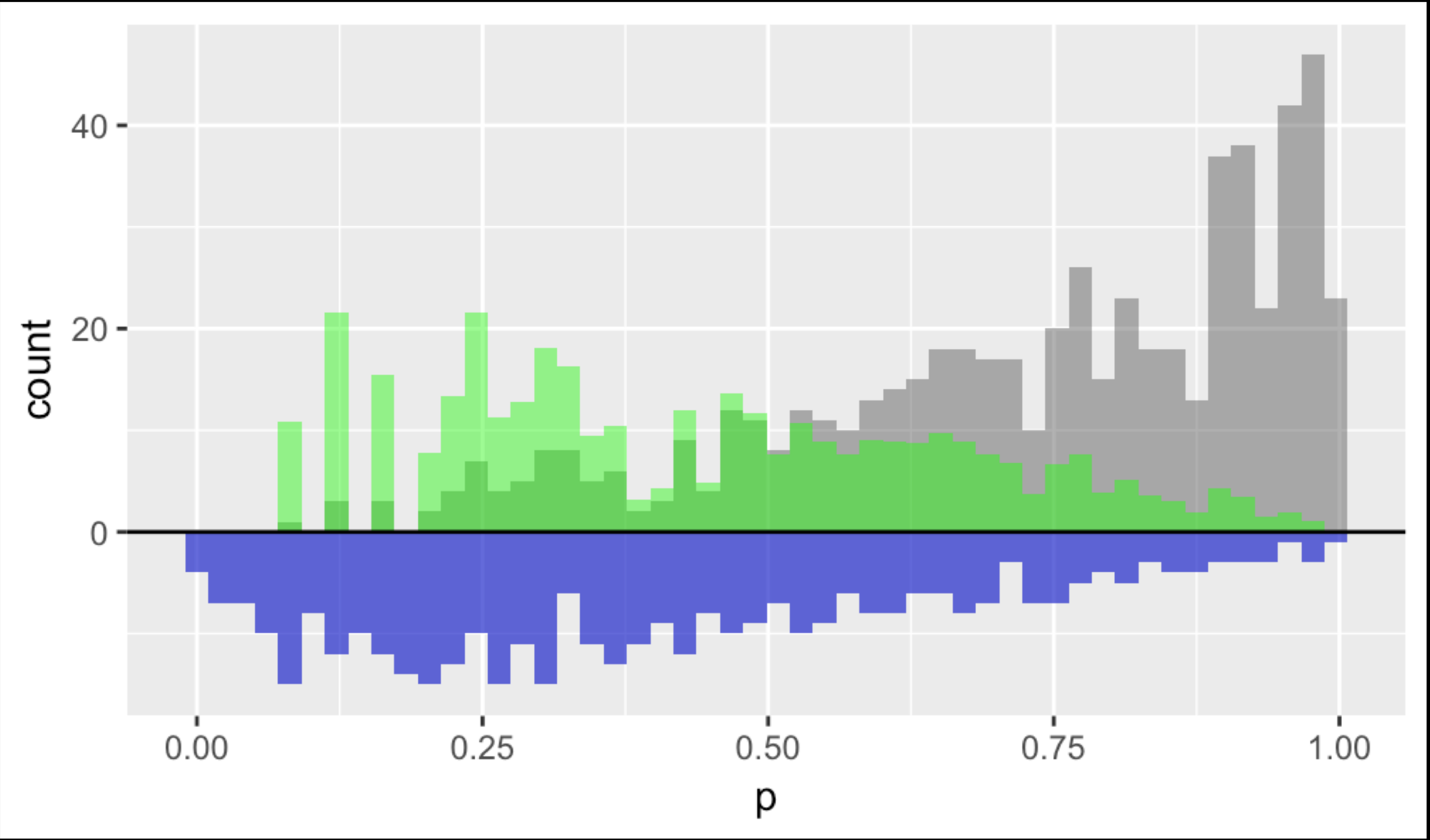
40
20
0
20
40
60
80
100
120
140
count
0.00
0.25
0.50
0.75
1.00
p

Average treatment effect among the controls

$$E[Y(1) - Y(0)|Z = 0]$$

Average treatment effect among the controls

$$w_{ATC} = \frac{(1 - e_i)Z_i}{e_i} + \frac{(1 - e_i)(1 - Z_i)}{1 - e_i}$$


count
40
20
0
0.00
0.25
0.50
0.75
1.00
p

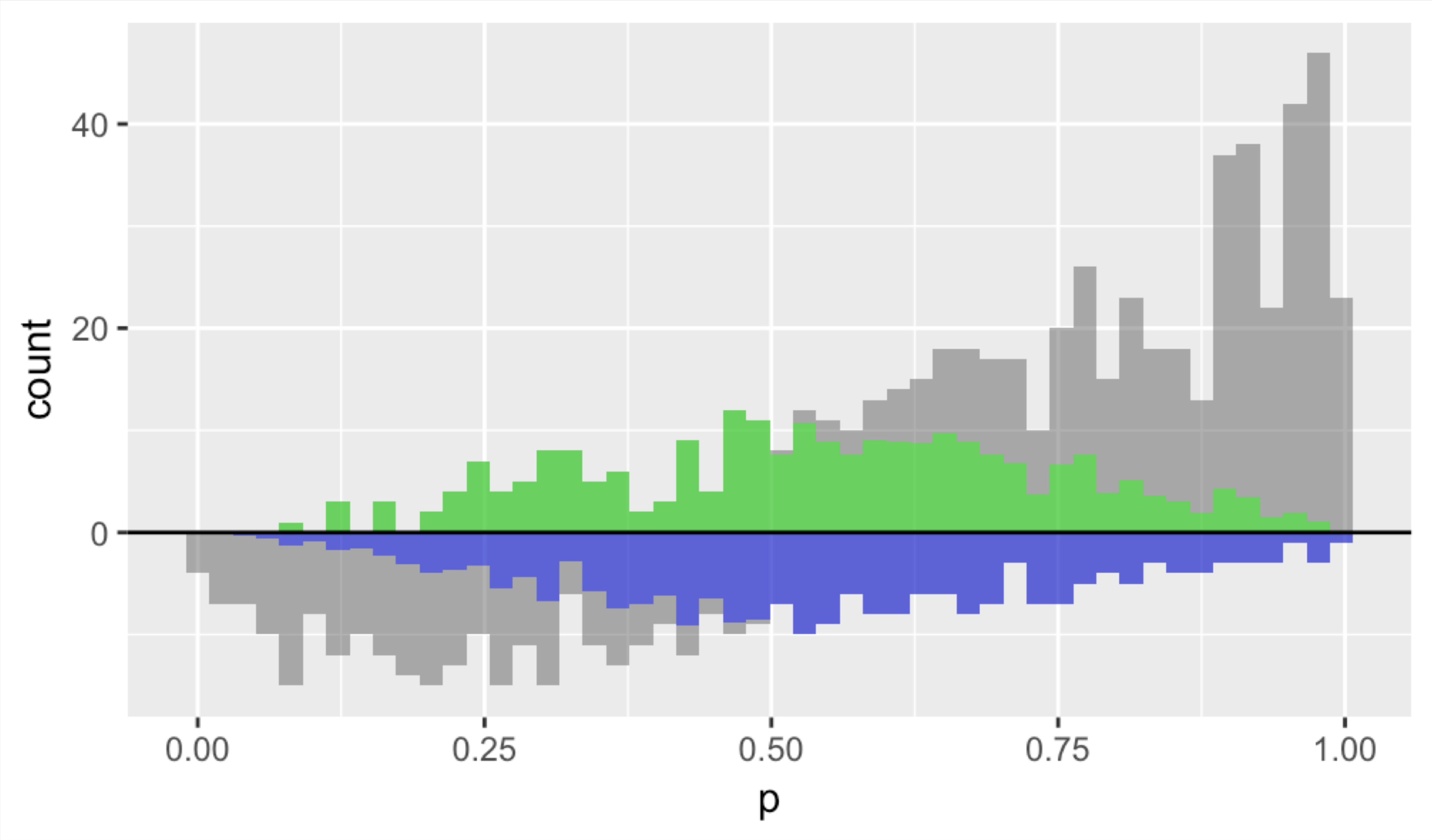
Average treatment effect among the evenly matchable

$$E[Y(1) - Y(0) | M_d = 1]$$

Average treatment effect among the evenly matchable

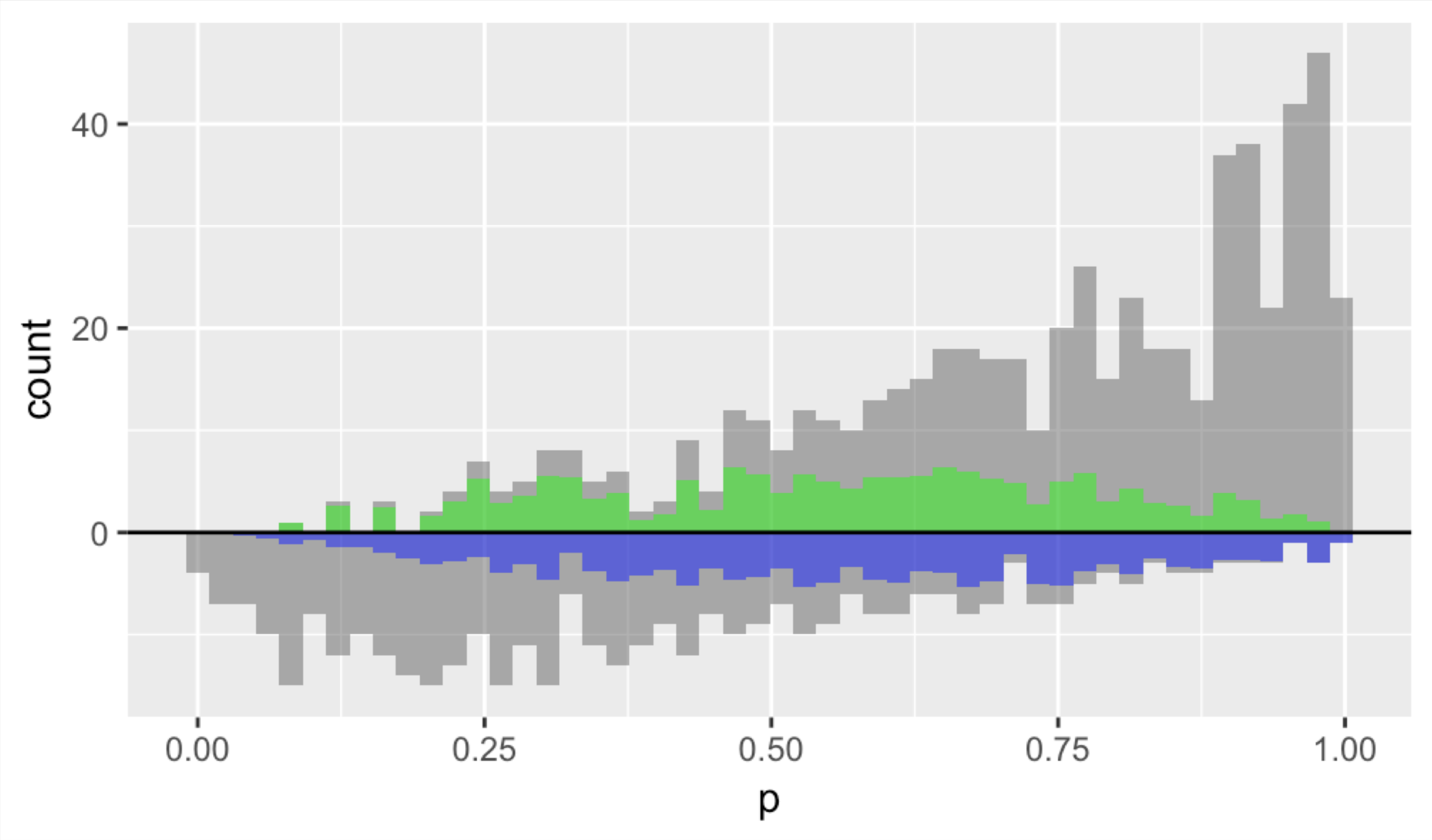
$$w_{ATM} = \frac{\min\{e_i, 1 - e_i\}}{e_i Z_i + (1 - e_i)(1 - Z_i)}$$

Minimum of ATT and ATC weights (Samuels 2017)


count
40
20
0
0.00
0.25
0.50
0.75
1.00
p

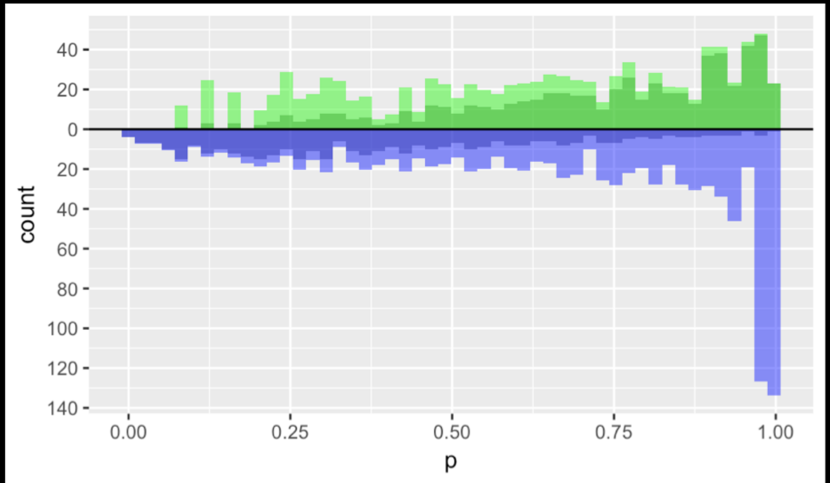
Average treatment effect among the overlap population

$$w_{ATO} = (1 - e_i)Z_i + e_i(1 - Z_i)$$

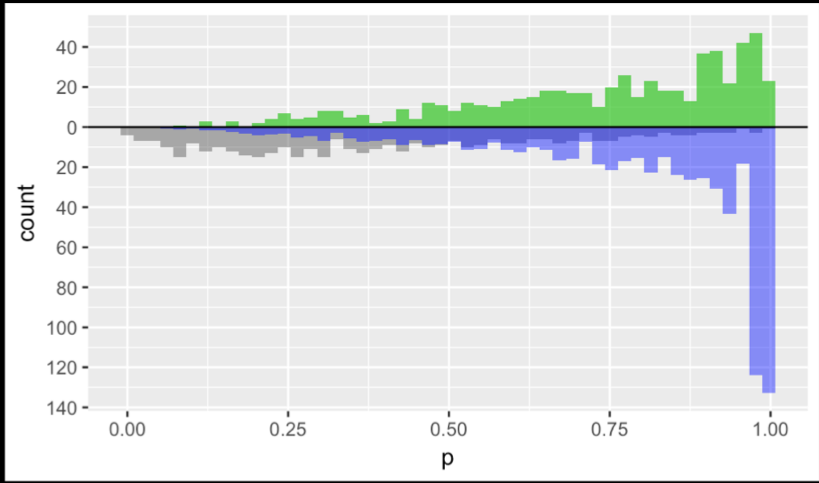


count
40
20
0
0.00
0.25
0.50
0.75
1.00
p

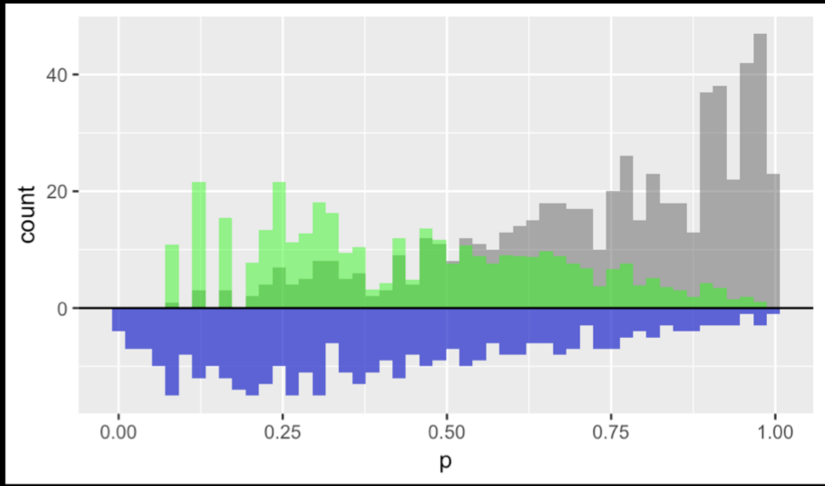
ATE

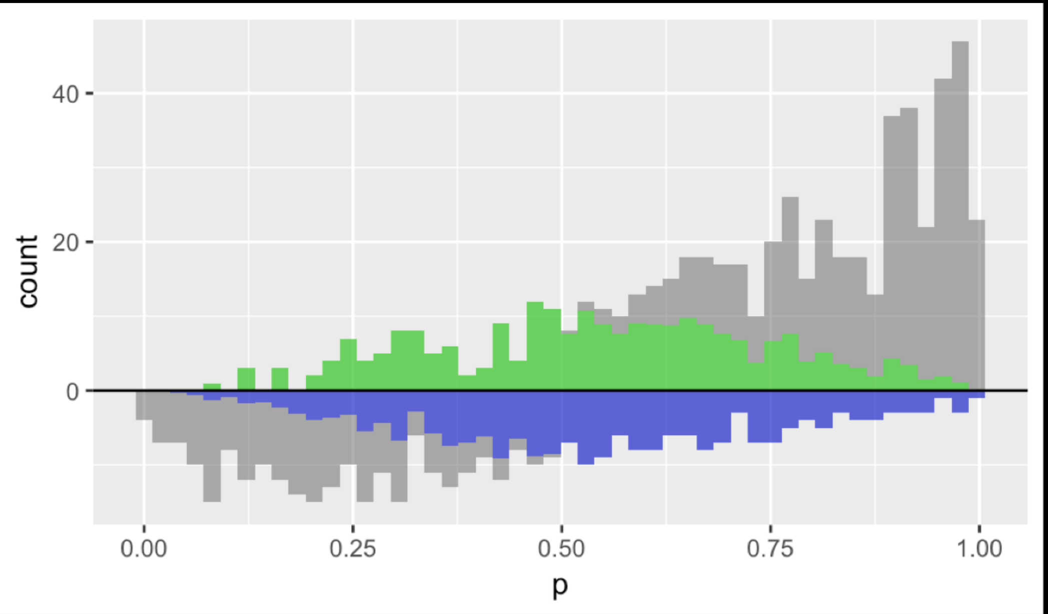


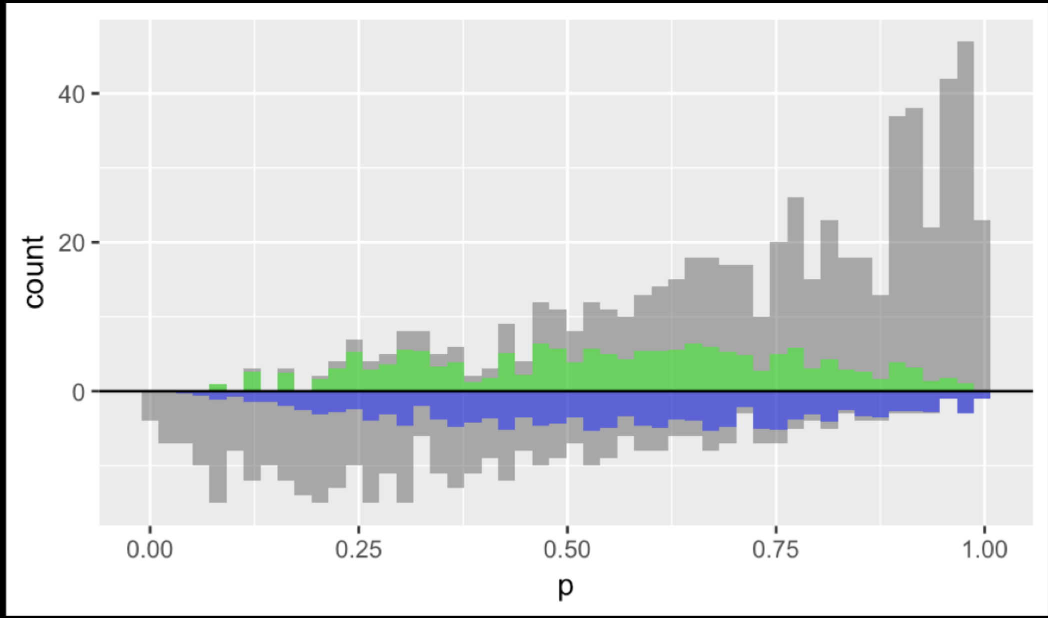
ATT



ATC



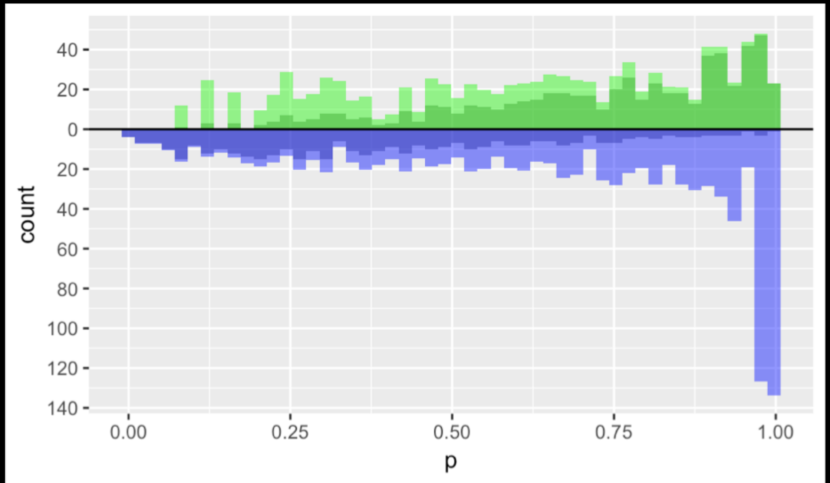




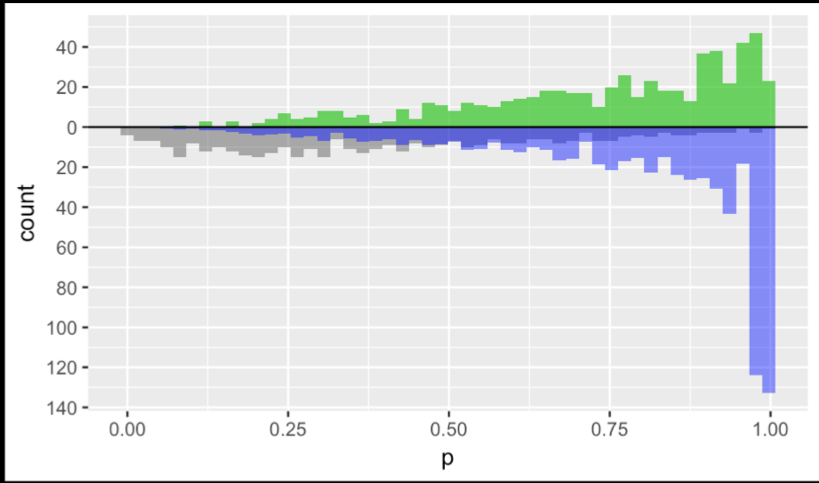
ATM

ATO

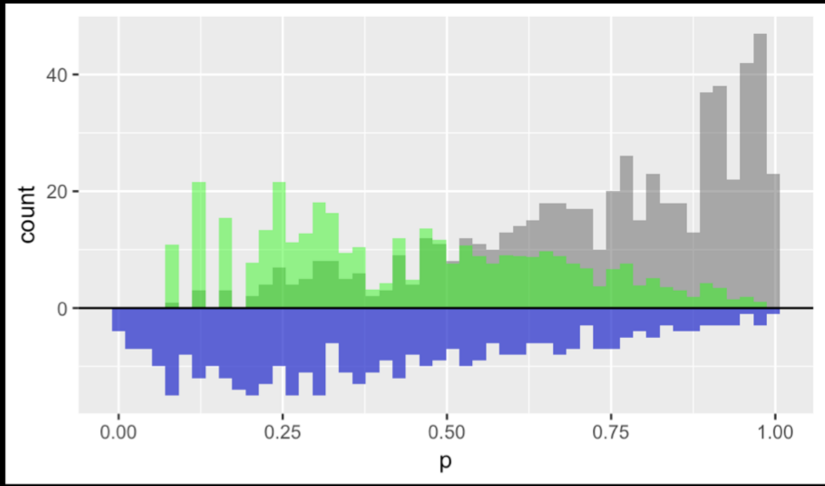
ATE

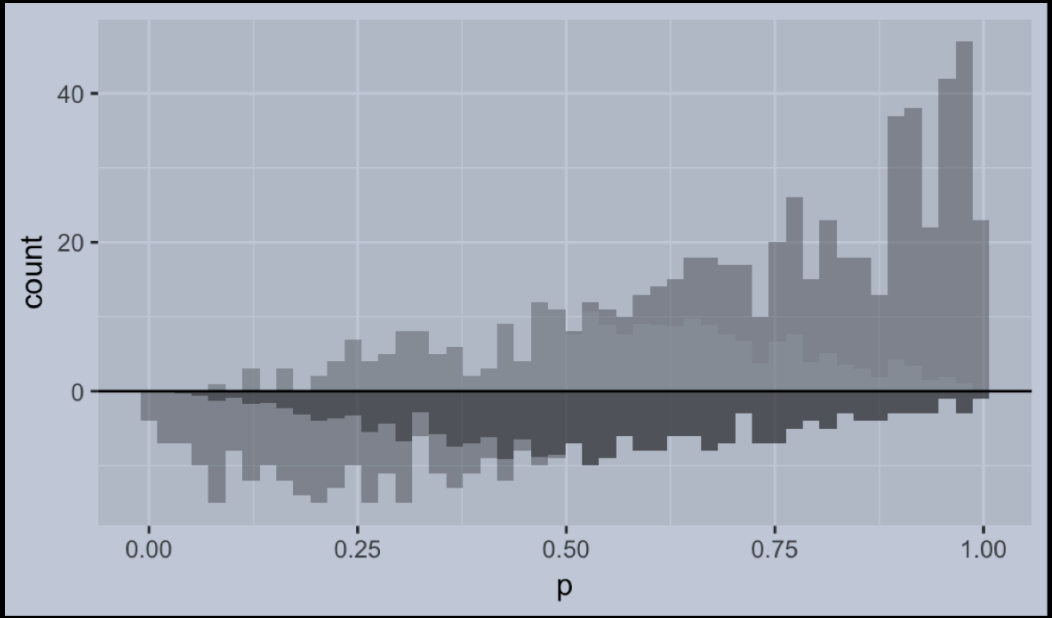


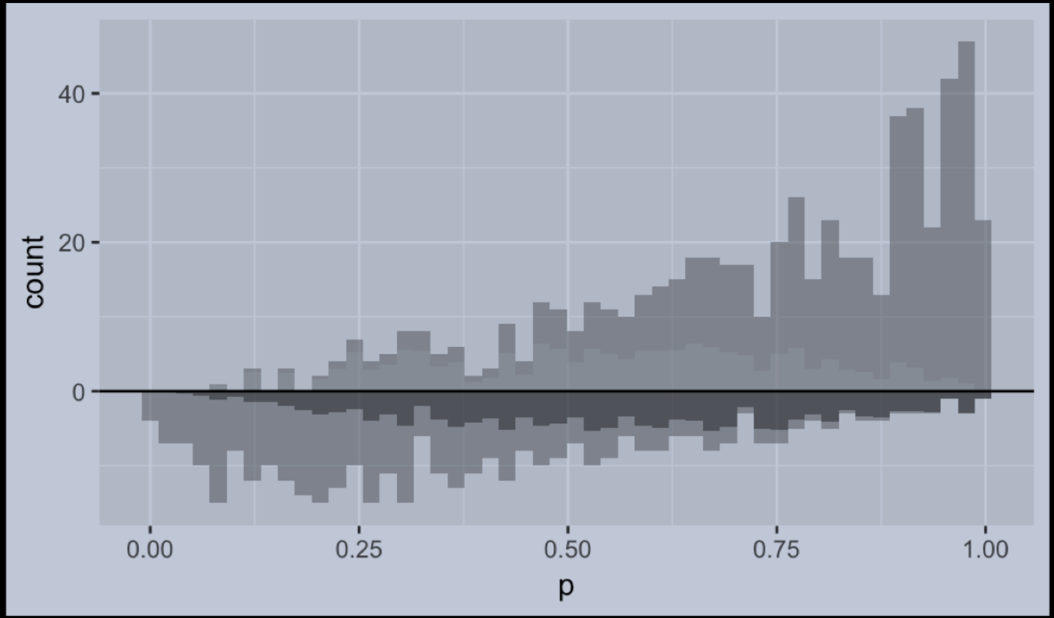
ATT



ATC



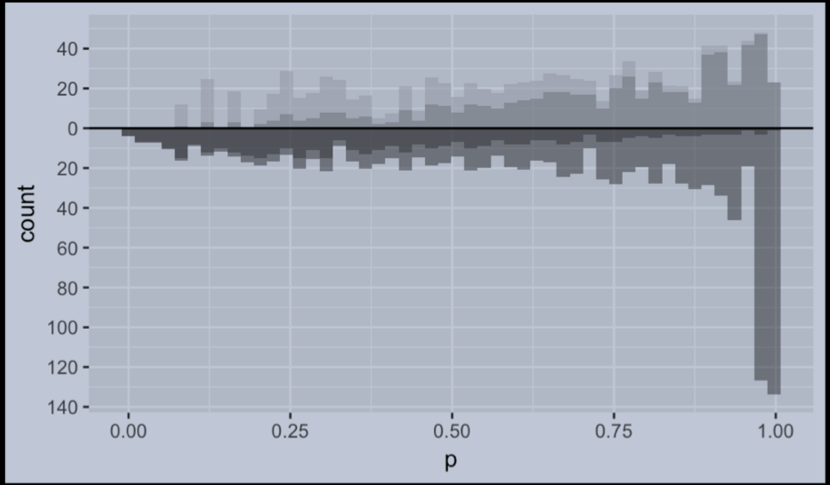




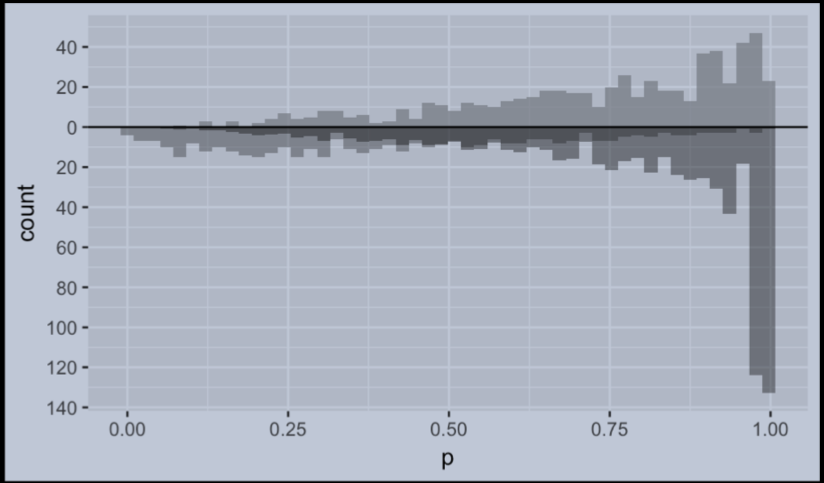
ATM

ATO

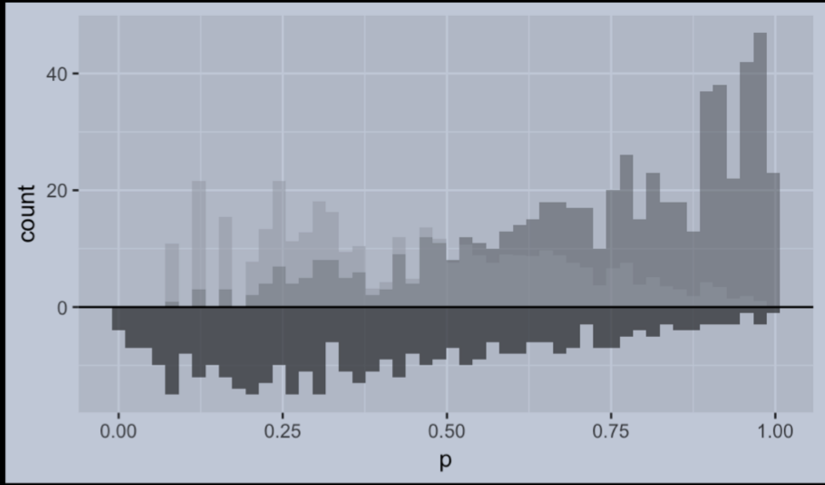
ATE

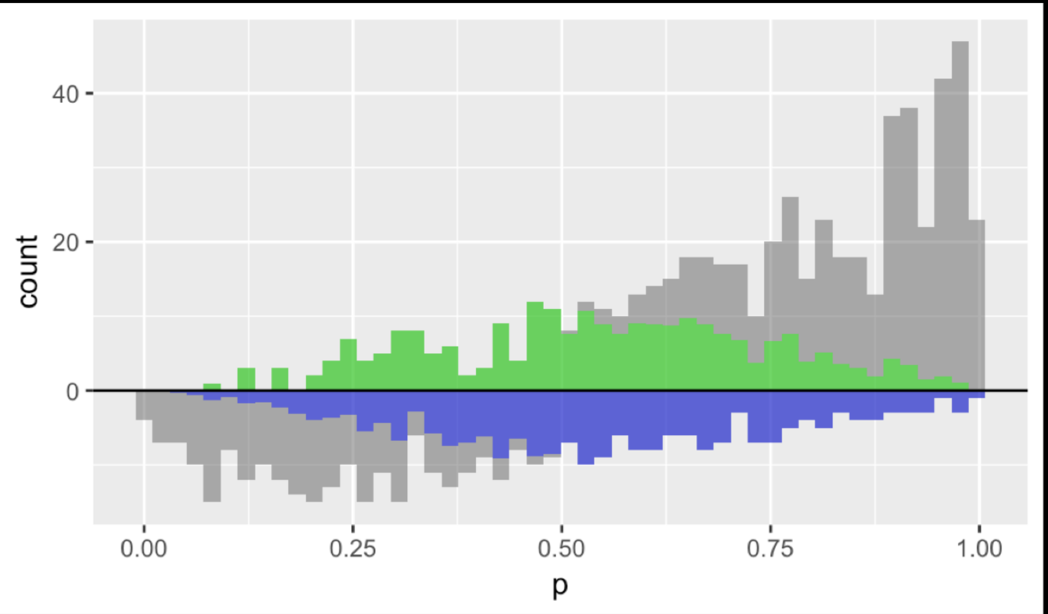


ATT

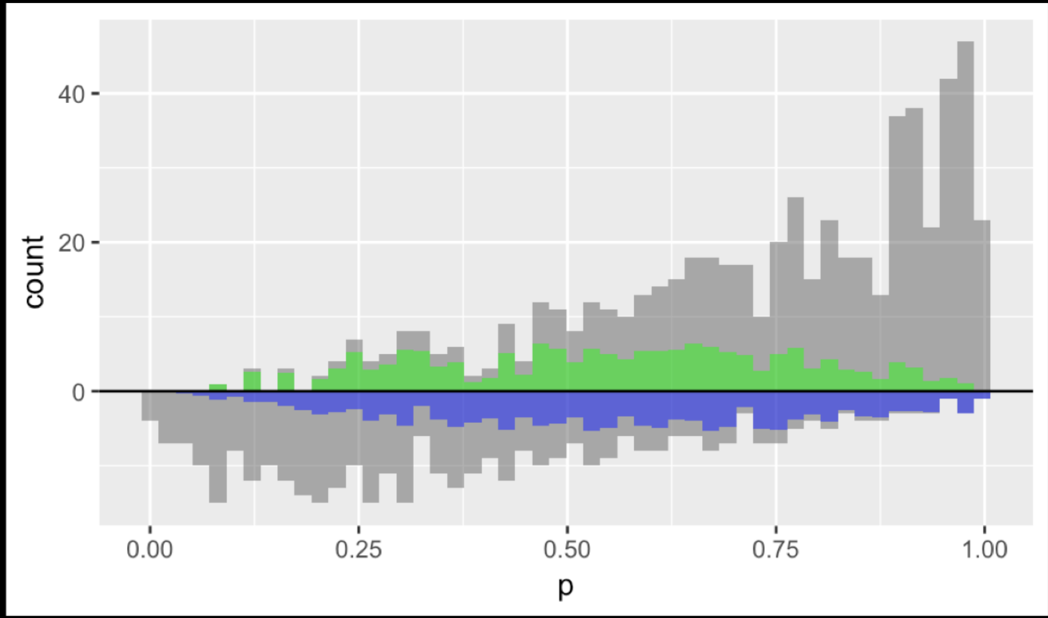


ATC





ATM



ATO

Freedman & Berk Simulation

Evaluation Revie
Volume 32 Number 4
August 2008 392-409

10.1177/0193841X08317586
http://erx.sagepub.com
hosted at
http://online.sagepub.com

Weighting Regressions by Propensity Scores

David A. Freedman
University of California, Berkeley
Richard A. Berk
University of Pennsylvania

Regressions can be weighted by propensity scores in order to reduce bias. However, weighting is likely to increase random error in the estimates, and to bias the estimated standard errors downward, even when selection mechanisms are well understood. Moreover, in some cases, weighting will increase the bias in estimated causal parameters. If investigators have a good causal model, it seems better just to fit the model without weights. If the causal model is improperly specified, there can be significant problems in retrieving the situation by weighting, although weighting may help under some circumstances.

Keywords: *causation; selection; models; experiments; observational studies; regression; propensity scores*

Estimating causal effects is often the key to evaluating social programs, but the interventions of interest are seldom assigned at random. Observational data are therefore frequently encountered. In order to

Fit a misspecified model

- Allow the propensity score to be **correct**
- The outcome model is **missing** X_2
- Fit the following
  - **Unweighted**
  - **Doubly robust ATE**
  - **Doubly robust ATM**
  - **Doubly robust ATO**

Doubly Robust Estimator

- Can misspecify **either** the propensity score model *or* the outcome model
- Lunceford and Davidian (2004) proved an **unbiased estimator for the ATE**
- Li and Greene proved an unbiased estimator for the **ATM**
- We proved an unbiased estimator for the **ATO**

Doubly robust estimator

$$\hat{\Delta}_{DR,w} = \frac{\sum_{i=1}^{n} w_i\big(m_1(\mathbf{X}_i, \hat{\boldsymbol{\alpha}}_1) - m_0(\mathbf{X}_i, \hat{\boldsymbol{\alpha}}_0)\big)}{\sum_{i=1}^{n} w_i} + \frac{\sum_{i=1}^{n} w_i Z_i\big(Y_i - m_1(\mathbf{X}_i, \hat{\boldsymbol{\alpha}}_1)\big)}{\sum_{i=1}^{n} w_i Z_i} - \frac{\sum_{i=1}^{n} w_i (1 - Z_i)\big(Y_i - m_0(\mathbf{X}_i, \hat{\boldsymbol{\alpha}}_0)\big)}{\sum_{i=1}^{n} w_i (1 - Z_i)}$$

Doubly robust estimator

$$\hat{\Delta}_{DR,w} = \frac{\sum_{i=1}^{n} w_i \big(m_1(\mathbf{X}_i, \hat{\boldsymbol{\alpha}}_1) - m_0(\mathbf{X}_i, \hat{\boldsymbol{\alpha}}_0)\big)}{\sum_{i=1}^{n} w_i} + \frac{\sum_{i=1}^{n} w_i Z_i \big(Y_i - m_1(\mathbf{X}_i, \hat{\boldsymbol{\alpha}}_1)\big)}{\sum_{i=1}^{n} w_i\, Z_i} - \frac{\sum_{i=1}^{n} w_i (1 - Z_i)\big(Y_i - m_0(\mathbf{X}_i, \hat{\boldsymbol{\alpha}}_0)\big)}{\sum_{i=1}^{n} w_i\, (1 - Z_i)}$$

Doubly robust estimator

$$\hat{\Delta}_{DR,w} = \frac{\sum_{i=1}^{n} w_i \left(m_1(\mathbf{X}_i, \hat{\boldsymbol{\alpha}}_1) - m_0(\mathbf{X}_i, \hat{\boldsymbol{\alpha}}_0) \right)}{\sum_{i=1}^{n} w_i} + \frac{\sum_{i=1}^{n} w_i Z_i \left(Y_i - m_1(\mathbf{X}_i, \hat{\boldsymbol{\alpha}}_1) \right)}{\sum_{i=1}^{n} w_i Z_i} - \frac{\sum_{i=1}^{n} w_i (1 - Z_i) \left(Y_i - m_0(\mathbf{X}_i, \hat{\boldsymbol{\alpha}}_0) \right)}{\sum_{i=1}^{n} w_i (1 - Z_i)}$$

Doubly robust estimator

$$\hat{\Delta}_{DR,w} = \frac{\sum_{i=1}^{n} w_i \big(m_1(\mathbf{X}_i, \hat{\boldsymbol{\alpha}}_1) - m_0(\mathbf{X}_i, \hat{\boldsymbol{\alpha}}_0)\big)}{\sum_{i=1}^{n} w_i} + \frac{\sum_{i=1}^{n} w_i Z_i \big(Y_i - m_1(\mathbf{X}_i, \hat{\boldsymbol{\alpha}}_1)\big)}{\sum_{i=1}^{n} w_i Z_i} - \frac{\sum_{i=1}^{n} w_i (1 - Z_i) \big(Y_i - m_0(\mathbf{X}_i, \hat{\boldsymbol{\alpha}}_0)\big)}{\sum_{i=1}^{n} w_i (1 - Z_i)}$$

Doubly robust estimator

$$\hat{\Delta}_{DR,w} = \frac{\sum_{i=1}^{n} w_i \left(m_1(\mathbf{X}_i, \hat{\boldsymbol{\alpha}}_1) - m_0(\mathbf{X}_i, \hat{\boldsymbol{\alpha}}_0)\right)}{\sum_{i=1}^{n} w_i} + \frac{\sum_{i=1}^{n} w_i Z_i \left(Y_i - m_1(\mathbf{X}_i, \hat{\boldsymbol{\alpha}}_1)\right)}{\sum_{i=1}^{n} w_i Z_i} - \frac{\sum_{i=1}^{n} w_i (1 - Z_i)\left(Y_i - m_0(\mathbf{X}_i, \hat{\boldsymbol{\alpha}}_0)\right)}{\sum_{i=1}^{n} w_i (1 - Z_i)}$$

Doubly robust estimator

$$\hat{\Delta}_{DR,w} = \frac{\sum_{i=1}^{n} w_i\big(m_1(\mathbf{X}_i, \hat{\boldsymbol{\alpha}}_1) - m_0(\mathbf{X}_i, \hat{\boldsymbol{\alpha}}_0)\big)}{\sum_{i=1}^{n} w_i} + \frac{\sum_{i=1}^{n} w_i Z_i\big(Y_i - m_1(\mathbf{X}_i, \hat{\boldsymbol{\alpha}}_1)\big)}{\sum_{i=1}^{n} w_i Z_i} - \frac{\sum_{i=1}^{n} w_i (1 - Z_i)\big(Y_i - m_0(\mathbf{X}_i, \hat{\boldsymbol{\alpha}}_0)\big)}{\sum_{i=1}^{n} w_i (1 - Z_i)}$$

Doubly robust estimator

$$\hat{\Delta}_{DR,w} = \frac{\sum_{i=1}^{n} w_i \left(m_1(\mathbf{X}_i, \hat{\boldsymbol{\alpha}}_1) - m_0(\mathbf{X}_i, \hat{\boldsymbol{\alpha}}_0) \right)}{\sum_{i=1}^{n} w_i} + \frac{\sum_{i=1}^{n} w_i Z_i \left(Y_i - m_1(\mathbf{X}_i, \hat{\boldsymbol{\alpha}}_1) \right)}{\sum_{i=1}^{n} w_i Z_i} - \frac{\sum_{i=1}^{n} w_i (1 - Z_i) \left(Y_i - m_0(\mathbf{X}_i, \hat{\boldsymbol{\alpha}}_0) \right)}{\sum_{i=1}^{n} w_i (1 - Z_i)}$$

Simulation

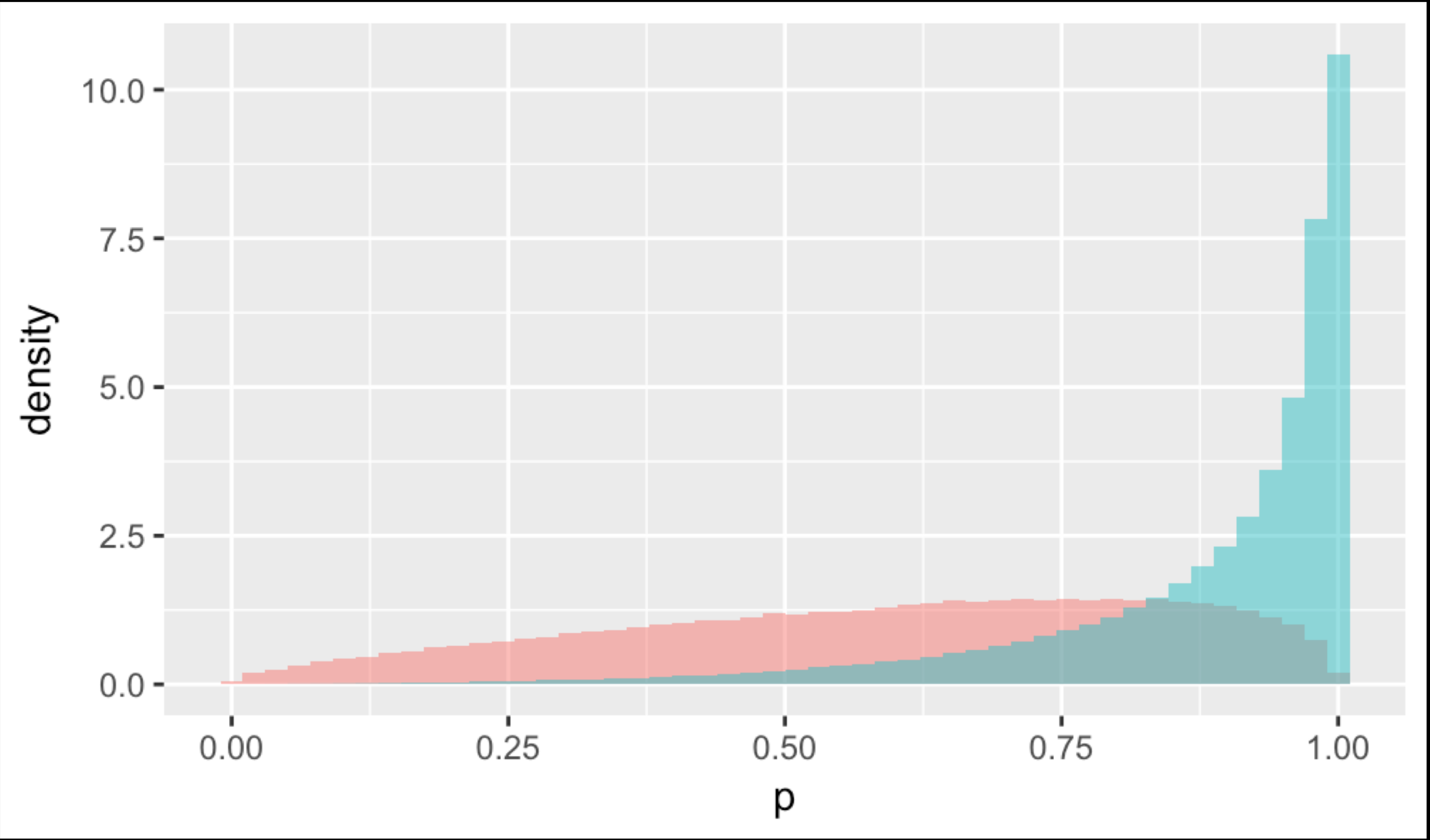
$$\mathbf{X} \sim MVN\left(\begin{pmatrix}0.5\\1\end{pmatrix}, \begin{pmatrix}2 & 1\\1 & 1\end{pmatrix}\right)$$

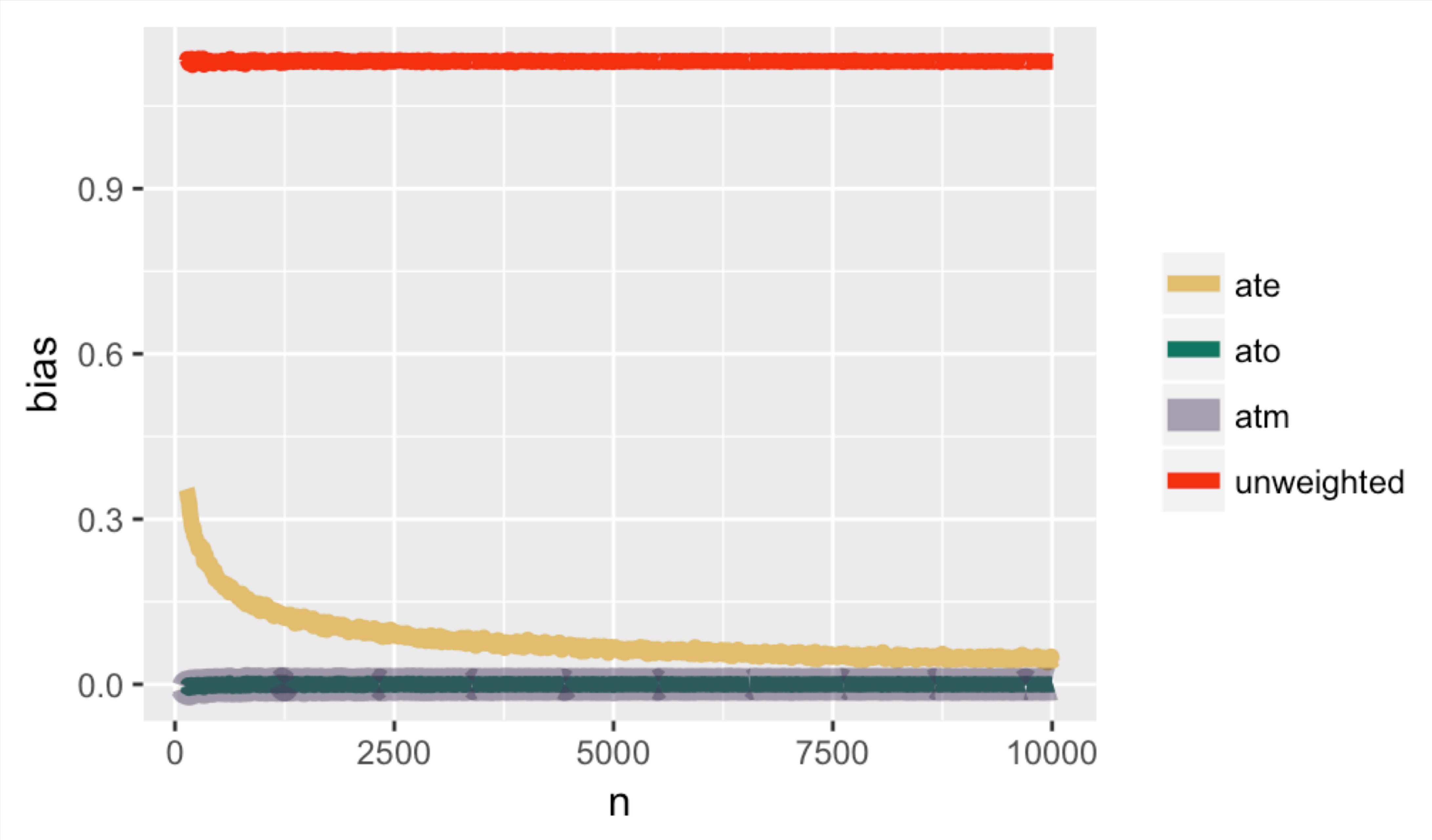
$$V \sim N\,(0,1)$$

$$Z = \begin{cases}1 & 0.5 + 0.25X_1 + 0.75X_2 + V > 0\\0 & o.w.\end{cases}$$

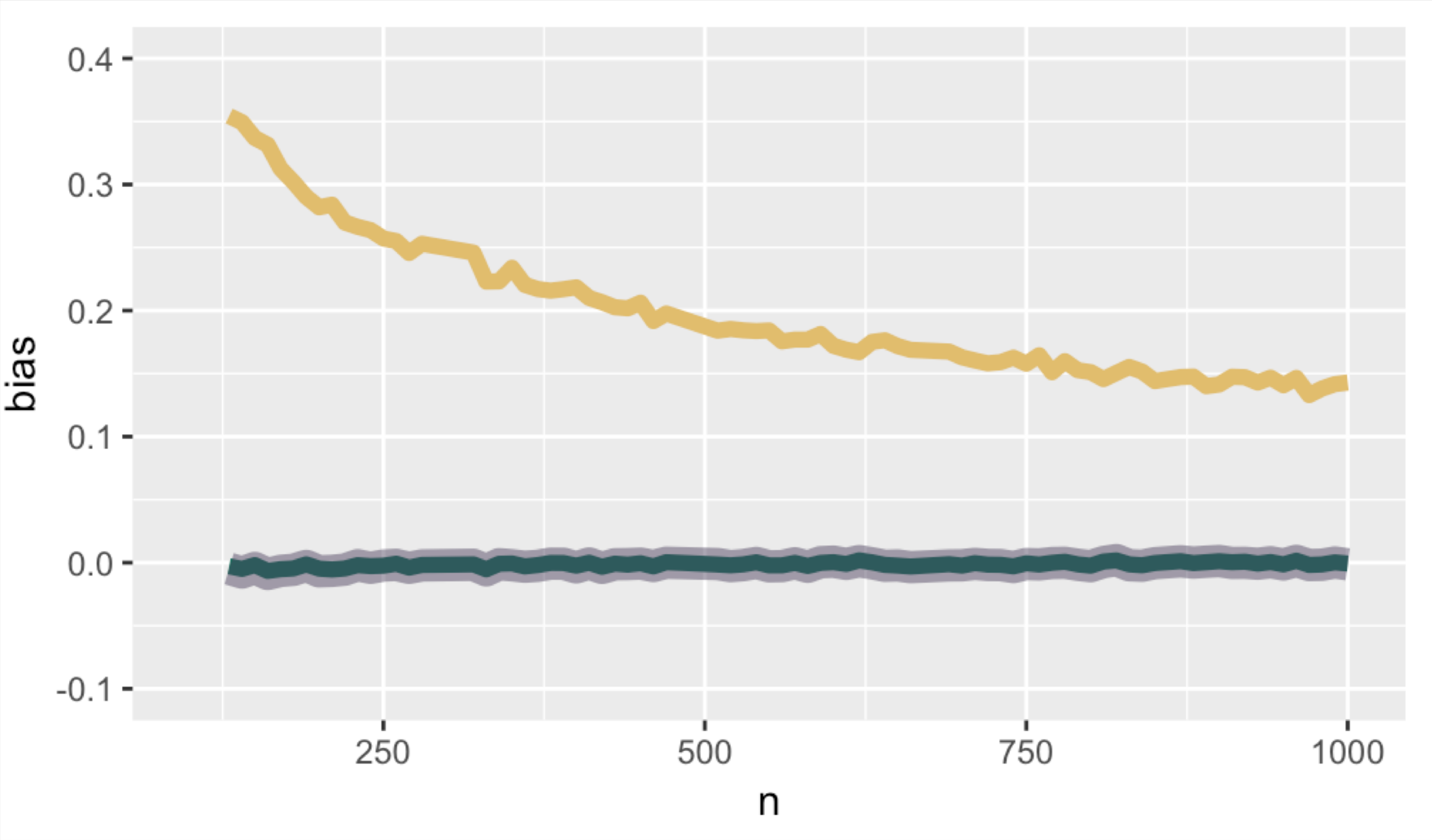
$$U \sim N(0,1)$$

$$Y = 1 + Z + X_1 + 2X_2 + U$$

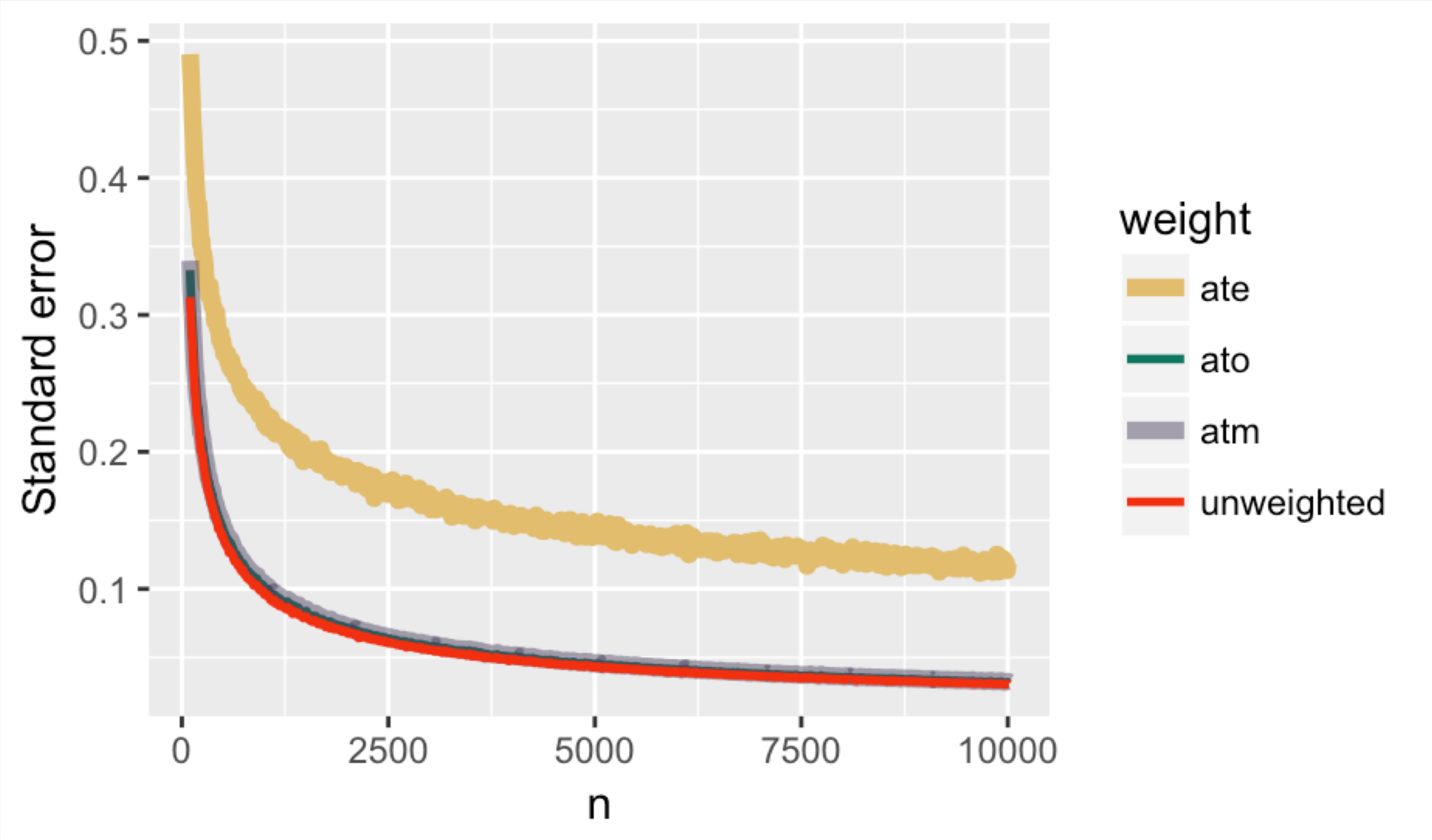

10.0
7.5
5.0
2.5
0.0
density
0.00
0.25
0.50
0.75
1.00
p

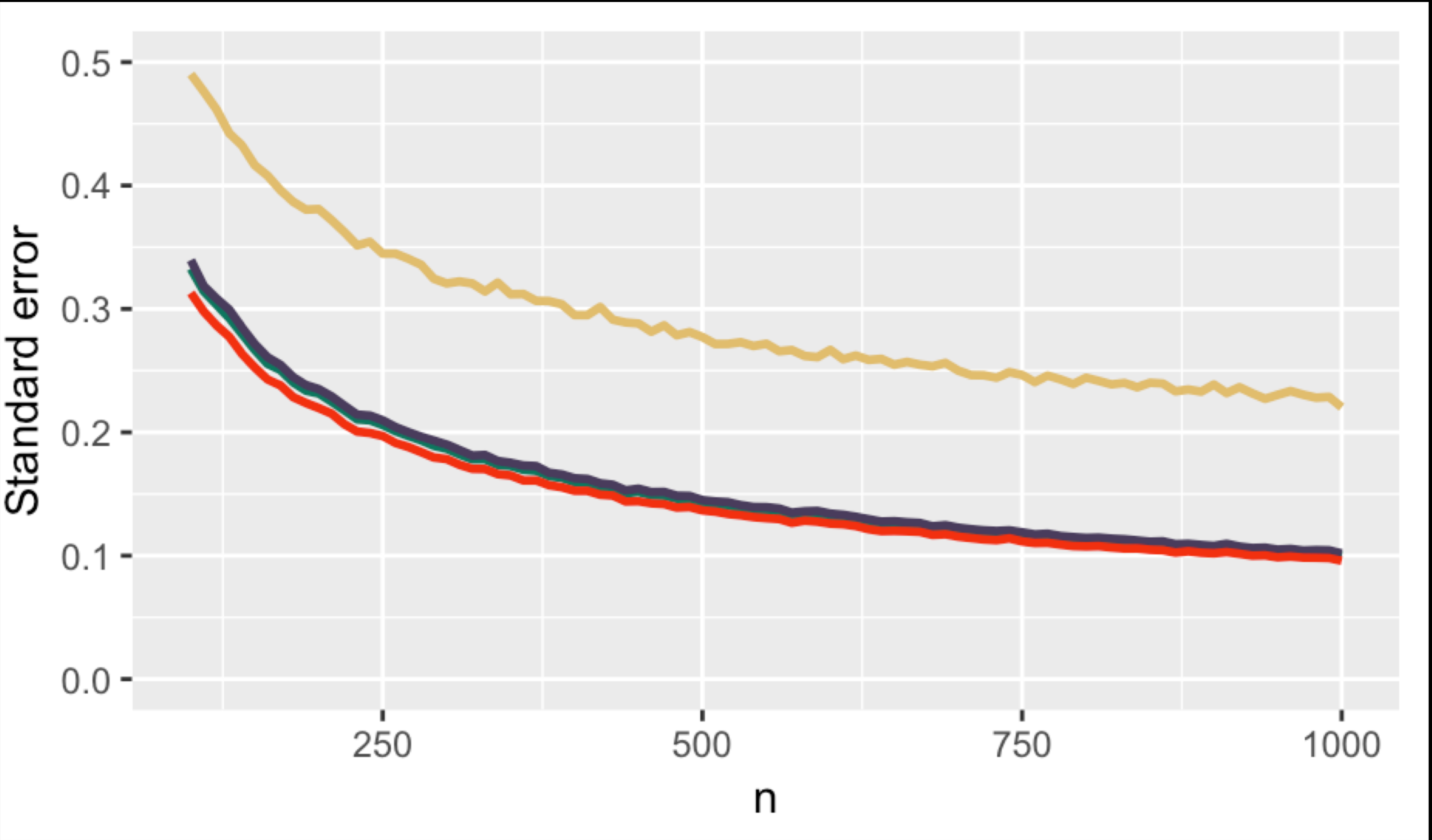


bias
0.9
0.6
0.3
0.0
0
2500
5000
7500
10000
n
ate
ato
atm
unweighted


0.4
0.3
0.2
0.1
0.0
-0.1
bias
250
500
750
1000
n

Standard errors


0.5
0.4
0.3
0.2
0.1
Standard error
0
2500
5000
7500
10000
n
weight
ate
ato
atm
unweighted


0.5
0.4
0.3
0.2
0.1
0.0
Standard error
250
500
750
1000
n

Recap

- **Replicated** the finite-sample bias seen by Freedman and Berk using the **ATE** weights
- **ATM** and **ATO** weights had **improved** finite-sample properties
- The **variance** for the **ATO** and **ATM** is **preferable** to that of the ATE

Q&A

Doubly robust and large sample variance estimator for overlap weights

ATO estimator

$$\frac{\sum_{i=1}^{n}\left(1-e\left(\mathbf{X}_i,\widehat{\boldsymbol{\beta}}\right)\right)Z_iY_i}{\sum_{i=1}^{n}\left(1-e\left(\mathbf{X}_i,\widehat{\boldsymbol{\beta}}\right)\right)Z_i}-\frac{\sum_{i=1}^{n}e\left(\mathbf{X}_i,\widehat{\boldsymbol{\beta}}\right)(1-Z_i)Y_i}{\sum_{i=1}^{n}e\left(\mathbf{X}_i,\widehat{\boldsymbol{\beta}}\right)(1-Z_i)}$$

ATO estimator

$$\frac{\sum_{i=1}^{n}\left(1-e\left(\mathbf{X}_i,\widehat{\boldsymbol{\beta}}\right)\right)Z_iY_i}{\sum_{i=1}^{n}\left(1-e\left(\mathbf{X}_i,\widehat{\boldsymbol{\beta}}\right)\right)Z_i}-\frac{\sum_{i=1}^{n}e\left(\mathbf{X}_i,\widehat{\boldsymbol{\beta}}\right)(1-Z_i)Y_i}{\sum_{i=1}^{n}e\left(\mathbf{X}_i,\widehat{\boldsymbol{\beta}}\right)(1-Z_i)}$$

ATO estimator

$$\frac{\sum_{i=1}^{n}\left(1-e\left(\mathbf{X}_i,\widehat{\boldsymbol{\beta}}\right)\right)Z_iY_i}{\sum_{i=1}^{n}\left(1-e\left(\mathbf{X}_i,\widehat{\boldsymbol{\beta}}\right)\right)Z_i}-\frac{\sum_{i=1}^{n}e\left(\mathbf{X}_i,\widehat{\boldsymbol{\beta}}\right)(1-Z_i)Y_i}{\sum_{i=1}^{n}e\left(\mathbf{X}_i,\widehat{\boldsymbol{\beta}}\right)(1-Z_i)}$$

ATO estimator

$$\underbrace{\frac{\sum_{i=1}^{n}\left(1-e\left(\mathbf{X}_i,\widehat{\boldsymbol{\beta}}\right)\right)Z_iY_i}{\sum_{i=1}^{n}\left(1-e\left(\mathbf{X}_i,\widehat{\boldsymbol{\beta}}\right)\right)Z_i}}_{\hat{\mu}_1} - \frac{\sum_{i=1}^{n} e\left(\mathbf{X}_i,\widehat{\boldsymbol{\beta}}\right)(1-Z_i)Y_i}{\sum_{i=1}^{n} e\left(\mathbf{X}_i,\widehat{\boldsymbol{\beta}}\right)(1-Z_i)}$$

ATO estimator

$$\underbrace{\frac{\sum_{i=1}^{n}\left(1-e\left(\mathbf{X}_i,\widehat{\boldsymbol{\beta}}\right)\right)Z_iY_i}{\sum_{i=1}^{n}\left(1-e\left(\mathbf{X}_i,\widehat{\boldsymbol{\beta}}\right)\right)Z_i}}_{\hat{\mu}_1}-\underbrace{\frac{\sum_{i=1}^{n}e\left(\mathbf{X}_i,\widehat{\boldsymbol{\beta}}\right)(1-Z_i)Y_i}{\sum_{i=1}^{n}e\left(\mathbf{X}_i,\widehat{\boldsymbol{\beta}}\right)(1-Z_i)}}_{\hat{\mu}_0}$$

ATO estimator

$$\frac{\sum_{i=1}^{n}\left(1-e(\mathbf{X}_i,\widehat{\boldsymbol{\beta}})\right)Z_iY_i}{\sum_{i=1}^{n}\left(1-e(\mathbf{X}_i,\widehat{\boldsymbol{\beta}})\right)Z_i} - \frac{\sum_{i=1}^{n}e(\mathbf{X}_i,\widehat{\boldsymbol{\beta}})(1-Z_i)Y_i}{\sum_{i=1}^{n}e(\mathbf{X}_i,\widehat{\boldsymbol{\beta}})(1-Z_i)}$$

$$\hat{\mu}_1 \qquad - \qquad \hat{\mu}_0$$

How to

- Plug in
- Fit a weighted generalized linear model Y ~ Z

How to

- Plug in
- Fit a weighted generalized linear model Y ~ Z

But what about the variance?

- Pull from the GLM? 🙅‍♀️
- Robust standard errors? (sandwich estimator) 🤷‍♀️

Williamson 2013

" …ignoring the estimation of the propensity score in the calculation of the variance leads to the **erroneous conclusion** that the IPTW treatment effect estimator has the same variance as an unadjusted estimator; thus, **it is important to use a variance estimator that correctly takes into account the estimation of the propensity score.** "

ATO estimator

ATO

weight

$$\boldsymbol{u}(\boldsymbol{\theta}) = \begin{pmatrix} (Y-\mu_1)Z\big(1-e(\mathbf{X},\boldsymbol{\beta})\big) \\ (Y-\mu_0)(1-Z)e(\mathbf{X},\boldsymbol{\beta}) \\ \mathbf{X}\big(Z-e(\mathbf{X},\boldsymbol{\beta})\big) \end{pmatrix}$$

$$\boldsymbol{\theta} = (\mu_1, \mu_0, \boldsymbol{\beta}^T)^T$$

ATO estimator

$$u(\boldsymbol{\theta}) = \begin{pmatrix} (Y - \mu_1)Z\big(1 - e(\mathbf{X}, \boldsymbol{\beta})\big) \\ (Y - \mu_0)(1 - Z)e(\mathbf{X}, \boldsymbol{\beta}) \\ \mathbf{X}\big(Z - e(\mathbf{X}, \boldsymbol{\beta})\big) \end{pmatrix}$$

$$\boldsymbol{\theta} = (\mu_1, \mu_0, \boldsymbol{\beta}^T)^T$$

M-Estimation

$$\sqrt{n}\left(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}\right) \xrightarrow{d} N(0, \mathbf{A}^{-1}\mathbf{B}\mathbf{A}^{-T})$$

Sandwich estimator

$$\mathrm{var}(\boldsymbol{\theta}) = \mathbf{A}^{-1}\mathbf{B}\mathbf{A}^{-T}$$

Sandwich estimator

$$\mathrm{var}(\boldsymbol{\theta}) = \mathbf{A}^{-1}\mathbf{B}\mathbf{A}^{-T}$$

$$\mathbf{A} = -E\left[\frac{\partial \boldsymbol{u}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\right]$$

$$\boldsymbol{\theta} = (\mu_1, \mu_0, \boldsymbol{\beta}^T)^T$$

$$\boldsymbol{u}(\boldsymbol{\theta}) = \begin{pmatrix} (Y - \mu_1)Z\big(1 - e(\mathbf{X}, \boldsymbol{\beta})\big) \\ (Y - \mu_0)(1 - Z)e(\mathbf{X}, \boldsymbol{\beta}) \\ \mathbf{X}\big(Z - e(\mathbf{X}, \boldsymbol{\beta})\big) \end{pmatrix}$$

$$\mathbf{A} = -E\left[\frac{\partial \boldsymbol{u}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\right]$$

$$\boldsymbol{\theta} = (\mu_1, \mu_0, \boldsymbol{\beta}^T)^T$$

$$\boldsymbol{u}(\boldsymbol{\theta}) = \begin{pmatrix} (Y - \mu_1)Z\big(1 - e(\mathbf{X}, \boldsymbol{\beta})\big) \\ (Y - \mu_0)(1 - Z)e(\mathbf{X}, \boldsymbol{\beta}) \\ \mathbf{X}\big(Z - e(\mathbf{X}, \boldsymbol{\beta})\big) \end{pmatrix}$$

$$\begin{pmatrix} E[Z(1-e)] & 0 & E[\mathbf{X}^T(Y-\mu_1)Ze(1-e)] \\ 0 & E[(1-Z)e] & -E[\mathbf{X}^T(Y-\mu_0)(1-Z)e(1-e)] \\ 0 & 0 & E[\mathbf{X}\mathbf{X}^T e(1-e)] \end{pmatrix}$$

$$\boldsymbol{\theta} = (\mu_1, \mu_0, \boldsymbol{\beta}^T)^T$$

$$\boldsymbol{u}(\boldsymbol{\theta}) = \begin{pmatrix} (Y - \mu_1)Z\big(1 - e(\mathbf{X}, \boldsymbol{\beta})\big) \\ (Y - \mu_0)(1 - Z)e(\mathbf{X}, \boldsymbol{\beta}) \\ \mathbf{X}\big(Z - e(\mathbf{X}, \boldsymbol{\beta})\big) \end{pmatrix}$$

$$\begin{pmatrix} E[Z(1-e)] & 0 & E[\mathbf{X}^T(Y-\mu_1)Ze(1-e)] \\ 0 & E[(1-Z)e] & -E[\mathbf{X}^T(Y-\mu_0)(1-Z)e(1-e)] \\ 0 & 0 & E[\mathbf{X}\mathbf{X}^T e(1-e)] \end{pmatrix}$$

$$\boldsymbol{\theta} = (\mu_1, \mu_0, \boldsymbol{\beta}^T)^T$$

$$\boldsymbol{u}(\boldsymbol{\theta}) = \begin{pmatrix} (Y - \mu_1)Z\big(1 - e(\mathbf{X}, \boldsymbol{\beta})\big) \\ (Y - \mu_0)(1 - Z)e(\mathbf{X}, \boldsymbol{\beta}) \\ \mathbf{X}\big(Z - e(\mathbf{X}, \boldsymbol{\beta})\big) \end{pmatrix}$$

$$\begin{pmatrix} E[Z(1-e)] & 0 & E[\mathbf{X}^T(Y-\mu_1)Ze(1-e)] \\ 0 & E[(1-Z)e] & -E[\mathbf{X}^T(Y-\mu_0)(1-Z)e(1-e)] \\ 0 & 0 & E[\mathbf{X}\mathbf{X}^T e(1-e)] \end{pmatrix}$$

$$\boldsymbol{\theta} = (\mu_1, \mu_0, \boldsymbol{\beta}^T)^T$$

$$\boldsymbol{u}(\boldsymbol{\theta}) = \begin{pmatrix} (Y-\mu_1)Z\big(1-e(\mathbf{X},\boldsymbol{\beta})\big) \\ (Y-\mu_0)(1-Z)e(\mathbf{X},\boldsymbol{\beta}) \\ \mathbf{X}\big(Z-e(\mathbf{X},\boldsymbol{\beta})\big) \end{pmatrix}$$

$$\frac{\partial e}{\partial \boldsymbol{\beta}^T} = \mathbf{X}^T e(1-e)$$

$$\frac{\partial (1-e)}{\partial \boldsymbol{\beta}^T} = -\mathbf{X}^T e(1-e)$$

$$\begin{pmatrix} E & & E[\mathbf{X}^T(Y-\mu_1)Ze(1-e)] \\ 0 & E[(1-Z)e] & -E[\mathbf{X}^T(Y-\mu_0)(1-Z)e(1-e)] \\ & & E[\mathbf{X}\mathbf{X}^T e(1-e)] \end{pmatrix}$$

Sandwich estimator

$$\mathrm{var}(\boldsymbol{\theta}) = \mathbf{A}^{-1}\mathbf{B}\mathbf{A}^{-T}$$

$$\mathbf{B} = E[\boldsymbol{u}(\boldsymbol{\theta})\boldsymbol{u}(\boldsymbol{\theta})^{T}]$$

$$\mathbf{B} = E[\boldsymbol{u}(\boldsymbol{\theta})\boldsymbol{u}(\boldsymbol{\theta})^T]$$

$$\mathbf{B} = E\left[\begin{pmatrix} (Y-\mu_1)Z\big(1-e(\mathbf{X},\boldsymbol{\beta})\big) \\ (Y-\mu_0)(1-Z)e(\mathbf{X},\boldsymbol{\beta}) \\ \mathbf{X}\big(Z-e(\mathbf{X},\boldsymbol{\beta})\big) \end{pmatrix} \times \begin{pmatrix} (Y-\mu_1)Z\big(1-e(\mathbf{X},\boldsymbol{\beta})\big) \\ (Y-\mu_0)(1-Z)e(\mathbf{X},\boldsymbol{\beta}) \\ \mathbf{X}\big(Z-e(\mathbf{X},\boldsymbol{\beta})\big) \end{pmatrix}^T\right]$$

Sandwich estimator

$$\widehat{\mathrm{var}}(\hat{\theta}) = \frac{1}{\mathrm{n}} \hat{\mathbf{A}}_n{}^{-1} \hat{\mathbf{B}}_n \hat{\mathbf{A}}_n{}^{-T}$$

Sandwich estimator

$$\widehat{\text{var}}(\hat{\mu}_1) = \frac{1}{\text{n}}\left(\widehat{\mathbf{A}}_n{}^{-1}\widehat{\mathbf{B}}_n\widehat{\mathbf{A}}_n{}^{-T}\right)_{1,1}$$

$$\widehat{\text{var}}(\hat{\mu}_0) = \frac{1}{\text{n}}\left(\widehat{\mathbf{A}}_n{}^{-1}\widehat{\mathbf{B}}_n\widehat{\mathbf{A}}_n{}^{-T}\right)_{2,2}$$

$$\widehat{\text{cov}}(\hat{\mu}_1, \hat{\mu}_0) = \frac{1}{\text{n}}\left(\widehat{\mathbf{A}}_n{}^{-1}\widehat{\mathbf{B}}_n\widehat{\mathbf{A}}_n{}^{-T}\right)_{1,2}$$

Sandwich estimator

$$\widehat{\mathrm{var}}(\hat{\mu}_1) = \frac{1}{\mathrm{n}}\left(\widehat{\mathbf{A}}_n{}^{-1}\widehat{\mathbf{B}}_n\widehat{\mathbf{A}}_n{}^{-T}\right)_{1,1}$$

$$\widehat{\mathrm{var}}(\hat{\mu}_0) = \frac{1}{\mathrm{n}}\left(\widehat{\mathbf{A}}_n{}^{-1}\widehat{\mathbf{B}}_n\widehat{\mathbf{A}}_n{}^{-T}\right)_{2,2}$$

$$\widehat{\mathrm{cov}}(\hat{\mu}_1, \hat{\mu}_0) = \frac{1}{\mathrm{n}}\left(\widehat{\mathbf{A}}_n{}^{-1}\widehat{\mathbf{B}}_n\widehat{\mathbf{A}}_n{}^{-T}\right)_{1,2}$$

Sandwich estimator

$$\widehat{\text{var}}(\hat{\mu}_1) = \frac{1}{\text{n}}\left(\widehat{\mathbf{A}}_n^{-1}\widehat{\mathbf{B}}_n\widehat{\mathbf{A}}_n^{-T}\right)_{1,1}$$

$$\widehat{\text{var}}(\hat{\mu}_0) = \frac{1}{\text{n}}\left(\widehat{\mathbf{A}}_n^{-1}\widehat{\mathbf{B}}_n\widehat{\mathbf{A}}_n^{-T}\right)_{2,2}$$

$$\widehat{\text{cov}}(\hat{\mu}_1, \hat{\mu}_0) = \frac{1}{\text{n}}\left(\widehat{\mathbf{A}}_n^{-1}\widehat{\mathbf{B}}_n\widehat{\mathbf{A}}_n^{-T}\right)_{1,2}$$

Sandwich estimator

$$\widehat{\mathrm{var}}(\hat{\mu}_1) = \frac{1}{\mathrm{n}}\left(\widehat{\mathbf{A}}_n{}^{-1}\widehat{\mathbf{B}}_n\widehat{\mathbf{A}}_n{}^{-T}\right)_{1,1}$$

$$\widehat{\mathrm{var}}(\hat{\mu}_0) = \frac{1}{\mathrm{n}}\left(\widehat{\mathbf{A}}_n{}^{-1}\widehat{\mathbf{B}}_n\widehat{\mathbf{A}}_n{}^{-T}\right)_{2,2}$$

$$\widehat{\mathrm{cov}}(\hat{\mu}_1, \hat{\mu}_0) = \frac{1}{\mathrm{n}}\left(\widehat{\mathbf{A}}_n{}^{-1}\widehat{\mathbf{B}}_n\widehat{\mathbf{A}}_n{}^{-T}\right)_{1,2}$$

What if we want to *also* adjust the outcome model?

How to

- Fit a weighted generalized linear model Y ~ Z + X
- Doubly robust estimator

How to

- Fit a weighted generalized linear model Y ~ Z + X
- Doubly robust estimator

But what about the variance?

- Pull from the GLM? 🙅‍♀️
- Robust standard errors? (sandwich estimator) 🤷‍♀️

ATO DR estimator

$$\widehat{\Delta}_{DR,ATO} = \hat{\delta}_1 + \hat{\delta}_2 - \hat{\delta}_3$$

ATO DR estimator

$$\widehat{\Delta}_{DR,ATO} = \hat{\delta}_1 + \hat{\delta}_2 - \hat{\delta}_3$$

$$\hat{\delta}_1 = \frac{\sum_{i=1}^{n} \Big(Z_i \big(1 - e(\mathbf{X}_i, \boldsymbol{\beta})\big) + (1 - Z_i) e(\mathbf{X}_\mathrm{i}, \boldsymbol{\beta}) \Big) \big(m_1(\mathbf{X}_i, \widehat{\boldsymbol{\alpha}}_1) - m_0(\mathbf{X}_i, \widehat{\boldsymbol{\alpha}}_0) \big)}{\sum_{i=1}^{n} \Big(Z_i \big(1 - e(\mathbf{X}_i, \boldsymbol{\beta})\big) + (1 - Z_i) e(\mathbf{X}_\mathrm{i}, \boldsymbol{\beta}) \Big)}$$

ATO DR estimator

$$\widehat{\Delta}_{DR,ATO} = \hat{\delta}_1 + \hat{\delta}_2 - \hat{\delta}_3$$

$$\hat{\delta}_1 = \frac{\sum_{i=1}^{n} \left(Z_i \big(1 - e(\mathbf{X}_i, \boldsymbol{\beta})\big) + (1 - Z_i) e(\mathbf{X}_\mathrm{i}, \boldsymbol{\beta}) \right) \left(m_1(\mathbf{X}_i, \widehat{\boldsymbol{\alpha}}_1) - m_0(\mathbf{X}_i, \widehat{\boldsymbol{\alpha}}_0) \right)}{\sum_{i=1}^{n} \left(Z_i \big(1 - e(\mathbf{X}_i, \boldsymbol{\beta})\big) + (1 - Z_i) e(\mathbf{X}_\mathrm{i}, \boldsymbol{\beta}) \right)}$$

ATO DR estimator

$$\widehat{\Delta}_{DR,ATO} = \hat{\delta}_1 + \hat{\delta}_2 - \hat{\delta}_3$$

$$\hat{\delta}_1 = \frac{\sum_{i=1}^{n}\Big(Z_i\big(1 - e(\mathbf{X}_i, \boldsymbol{\beta})\big) + (1 - Z_i)e(\mathbf{X}_i, \boldsymbol{\beta})\Big)\big(m_1(\mathbf{X}_i, \widehat{\boldsymbol{\alpha}}_1) - m_0(\mathbf{X}_i, \widehat{\boldsymbol{\alpha}}_0)\big)}{\sum_{i=1}^{n}\Big(Z_i\big(1 - e(\mathbf{X}_i, \boldsymbol{\beta})\big) + (1 - Z_i)e(\mathbf{X}_i, \boldsymbol{\beta})\Big)}$$

ATO DR estimator

$$\widehat{\Delta}_{DR,ATO} = \hat{\delta}_1 + \hat{\delta}_2 - \hat{\delta}_3$$

$$\hat{\delta}_1 = \frac{\sum_{i=1}^{n}\Big(Z_i\big(1 - e(\mathbf{X}_i, \boldsymbol{\beta})\big) + (1 - Z_i)e(\mathbf{X}_i, \boldsymbol{\beta})\Big)\big(m_1(\mathbf{X}_i, \widehat{\boldsymbol{\alpha}}_1) - m_0(\mathbf{X}_i, \widehat{\boldsymbol{\alpha}}_0)\big)}{\sum_{i=1}^{n}\Big(Z_i\big(1 - e(\mathbf{X}_i, \boldsymbol{\beta})\big) + (1 - Z_i)e(\mathbf{X}_i, \boldsymbol{\beta})\Big)}$$

ATO DR estimator

$$\widehat{\Delta}_{DR,ATO} = \hat{\delta}_1 + \hat{\delta}_2 - \hat{\delta}_3$$

$$\hat{\delta}_2 = \frac{\sum_{i=1}^{n} Z_i\big(1 - e(\mathbf{X}_i, \boldsymbol{\beta})\big)\big(Y_i - m_1(\mathbf{X}_i, \widehat{\boldsymbol{\alpha}}_1)\big)}{\sum_{i=1}^{n} Z_i\big(1 - e(\mathbf{X}_i, \boldsymbol{\beta})\big)}$$

ATO DR estimator

$$\widehat{\Delta}_{DR,ATO} = \hat{\delta}_1 + \hat{\delta}_2 - \hat{\delta}_3$$

$$\hat{\delta}_2 = \frac{\sum_{i=1}^{n} Z_i\left(1 - e(\mathbf{X}_i, \boldsymbol{\beta})\right)\left(Y_i - m_1(\mathbf{X}_i, \widehat{\boldsymbol{\alpha}}_1)\right)}{\sum_{i=1}^{n} Z_i\left(1 - e(\mathbf{X}_i, \boldsymbol{\beta})\right)}$$

ATO DR estimator

$$\widehat{\Delta}_{DR,ATO} = \hat{\delta}_1 + \hat{\delta}_2 - \hat{\delta}_3$$

$$\hat{\delta}_2 = \frac{\sum_{i=1}^{n} Z_i\big(1 - e(\mathbf{X}_i, \boldsymbol{\beta})\big)\big(Y_i - m_1(\mathbf{X}_i, \widehat{\boldsymbol{\alpha}}_1)\big)}{\sum_{i=1}^{n} Z_i\big(1 - e(\mathbf{X}_i, \boldsymbol{\beta})\big)}$$

ATO DR estimator

$$\widehat{\Delta}_{DR,ATO} = \hat{\delta}_1 + \hat{\delta}_2 - \hat{\delta}_3$$

$$\hat{\delta}_3 = \frac{\sum_{i=1}^{n}(1 - Z_i)e(\mathbf{X}_i, \boldsymbol{\beta})\left(Y_i - m_0(\mathbf{X}_i, \widehat{\boldsymbol{\alpha}}_0)\right)}{\sum_{i=1}^{n}(1 - Z_i)e(\mathbf{X}_i, \boldsymbol{\beta})}$$

ATO DR estimator

$$\widehat{\Delta}_{DR,ATO} = \hat{\delta}_1 + \hat{\delta}_2 - \hat{\delta}_3$$

$$\hat{\delta}_3 = \frac{\sum_{i=1}^{n}(1 - Z_i)e(\mathbf{X}_i, \boldsymbol{\beta})\left(Y_i - m_0(\mathbf{X}_i, \widehat{\boldsymbol{\alpha}}_0)\right)}{\sum_{i=1}^{n}(1 - Z_i)e(\mathbf{X}_i, \boldsymbol{\beta})}$$

ATO DR estimator

$$\widehat{\Delta}_{DR,ATO} = \hat{\delta}_1 + \hat{\delta}_2 - \hat{\delta}_3$$

$$\hat{\delta}_3 = \frac{\sum_{i=1}^{n}(1 - Z_i)e(\mathbf{X}_i, \boldsymbol{\beta})\left(Y_i - m_0(\mathbf{X}_i, \widehat{\boldsymbol{\alpha}}_0)\right)}{\sum_{i=1}^{n}(1 - Z_i)e(\mathbf{X}_i, \boldsymbol{\beta})}$$

ATO DR estimator

$$\boldsymbol{u}(\boldsymbol{\theta}) = \begin{pmatrix} (m_1(\mathbf{X},\boldsymbol{\alpha}_1) - m_0(\mathbf{X},\boldsymbol{\alpha}_0) - \delta_1)(Z\big(1 - e(\mathbf{X},\boldsymbol{\beta})\big) + (1-Z)e(\mathbf{X},\boldsymbol{\beta})) \\ (Y - m_1(\mathbf{X},\boldsymbol{\alpha}_1) - \delta_2)Z(1 - e(\mathbf{X},\boldsymbol{\beta})) \\ (Y - m_0(\mathbf{X},\boldsymbol{\alpha}_0) - \delta_3)(1-Z)e(\mathbf{X},\boldsymbol{\beta}) \\ \big(Y - m_1(\mathbf{X},\boldsymbol{\alpha}_0)\big)Z\mathbf{X} \\ \big(Y - m_0(\mathbf{X},\boldsymbol{\alpha}_0)\big)(1-Z)\mathbf{X} \\ \mathbf{X}\big(Z - e(\mathbf{X},\boldsymbol{\beta})\big) \end{pmatrix}$$

$$\boldsymbol{\theta} = (\delta_1, \delta_2, \delta_3, \boldsymbol{\alpha}_1^T, \boldsymbol{\alpha}_0^T, \boldsymbol{\beta}^T)^T$$

ATO DR estimator

$$\boldsymbol{u}(\boldsymbol{\theta}) = \begin{pmatrix} (m_1(\mathbf{X}, \boldsymbol{\alpha}_1) - m_0(\mathbf{X}, \boldsymbol{\alpha}_0) - \delta_1)(Z\left(1 - e(\mathbf{X}, \boldsymbol{\beta})\right) + (1 - Z)e(\mathbf{X}, \boldsymbol{\beta})) \\ (Y - m_1(\mathbf{X}, \boldsymbol{\alpha}_1) - \delta_2)Z(1 - e(\mathbf{X}, \boldsymbol{\beta})) \\ (Y - m_0(\mathbf{X}, \boldsymbol{\alpha}_0) - \delta_3)(1 - Z)e(\mathbf{X}, \boldsymbol{\beta}) \\ \left(Y - m_1(\mathbf{X}, \boldsymbol{\alpha}_0)\right)Z\mathbf{X} \\ \left(Y - m_0(\mathbf{X}, \boldsymbol{\alpha}_0)\right)(1 - Z)\mathbf{X} \\ \mathbf{X}\left(Z - e(\mathbf{X}, \boldsymbol{\beta})\right) \end{pmatrix}$$

$$\boldsymbol{\theta} = (\delta_1, \delta_2, \delta_3, \boldsymbol{\alpha}_1^T, \boldsymbol{\alpha}_0^T, \boldsymbol{\beta}^T)^T$$

ATO DR estimator

$$
\boldsymbol{u}(\boldsymbol{\theta}) = \begin{pmatrix}
(m_1(\mathbf{X}, \boldsymbol{\alpha}_1) - m_0(\mathbf{X}, \boldsymbol{\alpha}_0) - \delta_1)\big(Z\big(1 - e(\mathbf{X}, \boldsymbol{\beta})\big) + (1 - Z)e(\mathbf{X}, \boldsymbol{\beta})\big) \\
(Y - m_1(\mathbf{X}, \boldsymbol{\alpha}_1) - \delta_2)Z(1 - e(\mathbf{X}, \boldsymbol{\beta})) \\
(Y - m_0(\mathbf{X}, \boldsymbol{\alpha}_0) - \delta_3)(1 - Z)e(\mathbf{X}, \boldsymbol{\beta}) \\
\big(Y - m_1(\mathbf{X}, \boldsymbol{\alpha}_0)\big)Z\mathbf{X} \\
\big(Y - m_0(\mathbf{X}, \boldsymbol{\alpha}_0)\big)(1 - Z)\mathbf{X} \\
\mathbf{X}\big(Z - e(\mathbf{X}, \boldsymbol{\beta})\big)
\end{pmatrix}
$$

$$
\boldsymbol{\theta} = (\delta_1, \delta_2, \delta_3, \boldsymbol{\alpha}_1^T, \boldsymbol{\alpha}_0^T, \boldsymbol{\beta}^T)^T
$$

ATO DR estimator

$$\boldsymbol{u}(\boldsymbol{\theta}) = \begin{pmatrix} (m_1(\mathbf{X}, \boldsymbol{\alpha}_1) - m_0(\mathbf{X}, \boldsymbol{\alpha}_0) - \delta_1)\left(Z\left(1 - e(\mathbf{X}, \boldsymbol{\beta})\right) + (1 - Z)e(\mathbf{X}, \boldsymbol{\beta})\right) \\ (Y - m_1(\mathbf{X}, \boldsymbol{\alpha}_1) - \delta_2)Z(1 - e(\mathbf{X}, \boldsymbol{\beta})) \\ (Y - m_0(\mathbf{X}, \boldsymbol{\alpha}_0) - \delta_3)(1 - Z)e(\mathbf{X}, \boldsymbol{\beta}) \\ \left(Y - m_1(\mathbf{X}, \boldsymbol{\alpha}_0)\right)Z\mathbf{X} \\ \left(Y - m_0(\mathbf{X}, \boldsymbol{\alpha}_0)\right)(1 - Z)\mathbf{X} \\ \mathbf{X}\left(Z - e(\mathbf{X}, \boldsymbol{\beta})\right) \end{pmatrix}$$

$$\boldsymbol{\theta} = (\delta_1, \delta_2, \delta_3, \boldsymbol{\alpha}_1^T, \boldsymbol{\alpha}_0^T, \boldsymbol{\beta}^T)^T$$

ATO DR estimator

$$\boldsymbol{u}(\boldsymbol{\theta}) = \begin{pmatrix} (m_1(\mathbf{X},\boldsymbol{\alpha}_1) - m_0(\mathbf{X},\boldsymbol{\alpha}_0) - \delta_1)(Z\big(1 - e(\mathbf{X},\boldsymbol{\beta})\big) + (1-Z)e(\mathbf{X},\boldsymbol{\beta})) \\ (Y - m_1(\mathbf{X},\boldsymbol{\alpha}_1) - \delta_2)Z(1 - e(\mathbf{X},\boldsymbol{\beta})) \\ (Y - m_0(\mathbf{X},\boldsymbol{\alpha}_0) - \delta_3)(1-Z)e(\mathbf{X},\boldsymbol{\beta}) \\ \big(Y - m_1(\mathbf{X},\boldsymbol{\alpha}_0)\big)Z\mathbf{X} \\ \big(Y - m_0(\mathbf{X},\boldsymbol{\alpha}_0)\big)(1-Z)\mathbf{X} \\ \mathbf{X}\big(Z - e(\mathbf{X},\boldsymbol{\beta})\big) \end{pmatrix}$$

$$\boldsymbol{\theta} = (\delta_1, \delta_2, \delta_3, \boldsymbol{\alpha}_1^T, \boldsymbol{\alpha}_0^T, \boldsymbol{\beta}^T)^T$$

ATO DR estimator

$$\boldsymbol{u}(\boldsymbol{\theta}) = \begin{pmatrix} (m_1(\mathbf{X}, \boldsymbol{\alpha}_1) - m_0(\mathbf{X}, \boldsymbol{\alpha}_0) - \delta_1)(Z\big(1 - e(\mathbf{X}, \boldsymbol{\beta})\big) + (1 - Z)e(\mathbf{X}, \boldsymbol{\beta})) \\ (Y - m_1(\mathbf{X}, \boldsymbol{\alpha}_1) - \delta_2)Z(1 - e(\mathbf{X}, \boldsymbol{\beta})) \\ (Y - m_0(\mathbf{X}, \boldsymbol{\alpha}_0) - \delta_3)(1 - Z)e(\mathbf{X}, \boldsymbol{\beta}) \\ \big(Y - m_1(\mathbf{X}, \boldsymbol{\alpha}_0)\big)Z\mathbf{X} \\ \big(Y - m_0(\mathbf{X}, \boldsymbol{\alpha}_0)\big)(1 - Z)\mathbf{X} \\ \mathbf{X}\big(Z - e(\mathbf{X}, \boldsymbol{\beta})\big) \end{pmatrix}$$

$$\boldsymbol{\theta} = (\delta_1, \delta_2, \delta_3, \boldsymbol{\alpha}_1^T, \boldsymbol{\alpha}_0^T, \boldsymbol{\beta}^T)^T$$

ATO DR estimator

$$\boldsymbol{u}(\boldsymbol{\theta}) = \begin{pmatrix} (m_1(\mathbf{X},\boldsymbol{\alpha}_1) - m_0(\mathbf{X},\boldsymbol{\alpha}_0) - \delta_1)(Z\big(1 - e(\mathbf{X},\boldsymbol{\beta})\big) + (1-Z)e(\mathbf{X},\boldsymbol{\beta})) \\ (Y - m_1(\mathbf{X},\boldsymbol{\alpha}_1) - \delta_2)Z(1 - e(\mathbf{X},\boldsymbol{\beta})) \\ (Y - m_0(\mathbf{X},\boldsymbol{\alpha}_0) - \delta_3)(1-Z)e(\mathbf{X},\boldsymbol{\beta}) \\ \big(Y - m_1(\mathbf{X},\boldsymbol{\alpha}_0)\big)Z\mathbf{X} \\ \big(Y - m_0(\mathbf{X},\boldsymbol{\alpha}_0)\big)(1-Z)\mathbf{X} \\ \mathbf{X}\big(Z - e(\mathbf{X},\boldsymbol{\beta})\big) \end{pmatrix}$$

$$\boldsymbol{\theta} = (\delta_1, \delta_2, \delta_3, \boldsymbol{\alpha}_1^T, \boldsymbol{\alpha}_0^T, \boldsymbol{\beta}^T)^T$$

ATO DR estimator

$$\boldsymbol{u}(\boldsymbol{\theta}) = \begin{pmatrix} (m_1(\mathbf{X},\boldsymbol{\alpha}_1) - m_0(\mathbf{X},\boldsymbol{\alpha}_0) - \delta_1)(Z\big(1 - e(\mathbf{X},\boldsymbol{\beta})\big) + (1 - Z)e(\mathbf{X},\boldsymbol{\beta})) \\ (Y - m_1(\mathbf{X},\boldsymbol{\alpha}_1) - \delta_2)Z(1 - e(\mathbf{X},\boldsymbol{\beta})) \\ (Y - m_0(\mathbf{X},\boldsymbol{\alpha}_0) - \delta_3)(1 - Z)e(\mathbf{X},\boldsymbol{\beta}) \\ \big(Y - m_1(\mathbf{X},\boldsymbol{\alpha}_0)\big)Z\mathbf{X} \\ \big(Y - m_0(\mathbf{X},\boldsymbol{\alpha}_0)\big)(1 - Z)\mathbf{X} \\ \mathbf{X}\big(Z - e(\mathbf{X},\boldsymbol{\beta})\big) \end{pmatrix}$$

$$\boldsymbol{\theta} = (\delta_1, \delta_2, \delta_3, \boldsymbol{\alpha}_1^T, \boldsymbol{\alpha}_0^T, \boldsymbol{\beta}^T)^T$$

ATO DR estimator

$$\boldsymbol{u}(\boldsymbol{\theta}) = \begin{pmatrix} (m_1(\mathbf{X}, \boldsymbol{\alpha}_1) - m_0(\mathbf{X}, \boldsymbol{\alpha}_0) - \delta_1)(Z\big(1 - e(\mathbf{X}, \boldsymbol{\beta})\big) + (1 - Z)e(\mathbf{X}, \boldsymbol{\beta})) \\ (Y - m_1(\mathbf{X}, \boldsymbol{\alpha}_1) - \delta_2)Z(1 - e(\mathbf{X}, \boldsymbol{\beta})) \\ (Y - m_0(\mathbf{X}, \boldsymbol{\alpha}_0) - \delta_3)(1 - Z)e(\mathbf{X}, \boldsymbol{\beta}) \\ \big(Y - m_1(\mathbf{X}, \boldsymbol{\alpha}_0)\big)Z\mathbf{X} \\ \big(Y - m_0(\mathbf{X}, \boldsymbol{\alpha}_0)\big)(1 - Z)\mathbf{X} \\ \mathbf{X}\big(Z - e(\mathbf{X}, \boldsymbol{\beta})\big) \end{pmatrix}$$

$$\boldsymbol{\theta} = (\delta_1, \delta_2, \delta_3, \boldsymbol{\alpha}_1^T, \boldsymbol{\alpha}_0^T, \boldsymbol{\beta}^T)^T$$

Sandwich estimator

$$\mathrm{var}(\boldsymbol{\theta}) = \mathbf{A}^{-1}\mathbf{B}\mathbf{A}^{-T}$$

ATO DR estimator

$$\widehat{\Delta}_{DR,ATO} = \hat{\delta}_1 + \hat{\delta}_2 - \hat{\delta}_3$$

$$var\left(\widehat{\boldsymbol{\Delta}}_{DR,ATO}\right) = \begin{pmatrix} 1 & 1 & -1 & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{pmatrix} \frac{1}{n} \widehat{\mathbf{A}}_n^{-1} \, \widehat{\mathrm{B}}_n \widehat{\mathrm{A}}_n^{-T} \begin{pmatrix} 1 \\ 1 \\ -1 \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \end{pmatrix}$$

ATO DR estimator

$$\widehat{\Delta}_{DR,ATO} = \hat{\delta}_1 + \hat{\delta}_2 - \hat{\delta}_3$$

$$\widehat{\mathrm{var}}\left(\widehat{\boldsymbol{\Delta}}_{DR,ATO}\right) = \begin{pmatrix} 1 & 1 & -1 & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{pmatrix} \frac{1}{n} \widehat{\mathbf{A}}_n^{-1} \, \widehat{\mathbf{B}}_n \widehat{\mathbf{A}}_n^{-T} \begin{pmatrix} 1 \\ 1 \\ -1 \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \end{pmatrix}$$

ATO DR estimator

$$\widehat{\Delta}_{DR,ATO} = \hat{\delta}_1 + \hat{\delta}_2 - \hat{\delta}_3$$

$$\widehat{\mathrm{var}}\left(\widehat{\Delta}_{DR,ATO}\right) = (1 \;\; 1 \;\; -1 \;\; \mathbf{0} \;\; \mathbf{0} \;\; \mathbf{0}) \frac{1}{n} \widehat{\mathbf{A}}_n^{-1} \, \widehat{\mathbf{B}}_n \widehat{\mathbf{A}}_n^{-T} \begin{pmatrix} 1 \\ 1 \\ -1 \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \end{pmatrix}$$

ATO DR estimator

$$\widehat{\Delta}_{DR,ATO} = \hat{\delta}_1 + \hat{\delta}_2 - \hat{\delta}_3$$

$$\widehat{\mathrm{var}}\left(\widehat{\boldsymbol{\Delta}}_{DR,ATO}\right) = \begin{pmatrix} 1 & 1 & -1 & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{pmatrix} \frac{1}{n} \widehat{\mathbf{A}}_n^{-1} \, \widehat{\mathbf{B}}_n \widehat{\mathbf{A}}_n^{-T} \begin{pmatrix} 1 \\ 1 \\ -1 \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \end{pmatrix}$$

Freedman & Berk*
Simulation

correct-correct
correct-wrong
wrong-correct
wrong-wrong

correct-correct

correct-wrong

wrong-correct

wrong-wrong

correct-correct
correct-wrong
wrong-correct
wrong-wrong

correct-correct
correct-wrong
wrong-correct
wrong-wrong

correct-correct
correct-wrong
wrong-correct
wrong-wrong

extreme
moderate
slight

bias
ratio estimated se / true se
RMSE
95% coverage

bias

ratio estimated se / true se

RMSE

95% coverage

Simulation

$$\mathbf{X} \sim MVN\left(\begin{pmatrix} 0.5 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}\right)$$

$$V \sim N\,(0,1)$$

$$Z = \begin{cases} 1 & 0.5 + 0.25X_1 + 0.75X_2 + V > 0 \\ 0 & o.w. \end{cases}$$

$$U \sim N(0,1)$$

$$Y = 1 + Z + X_1 + 3X_2 + U$$

Simulation

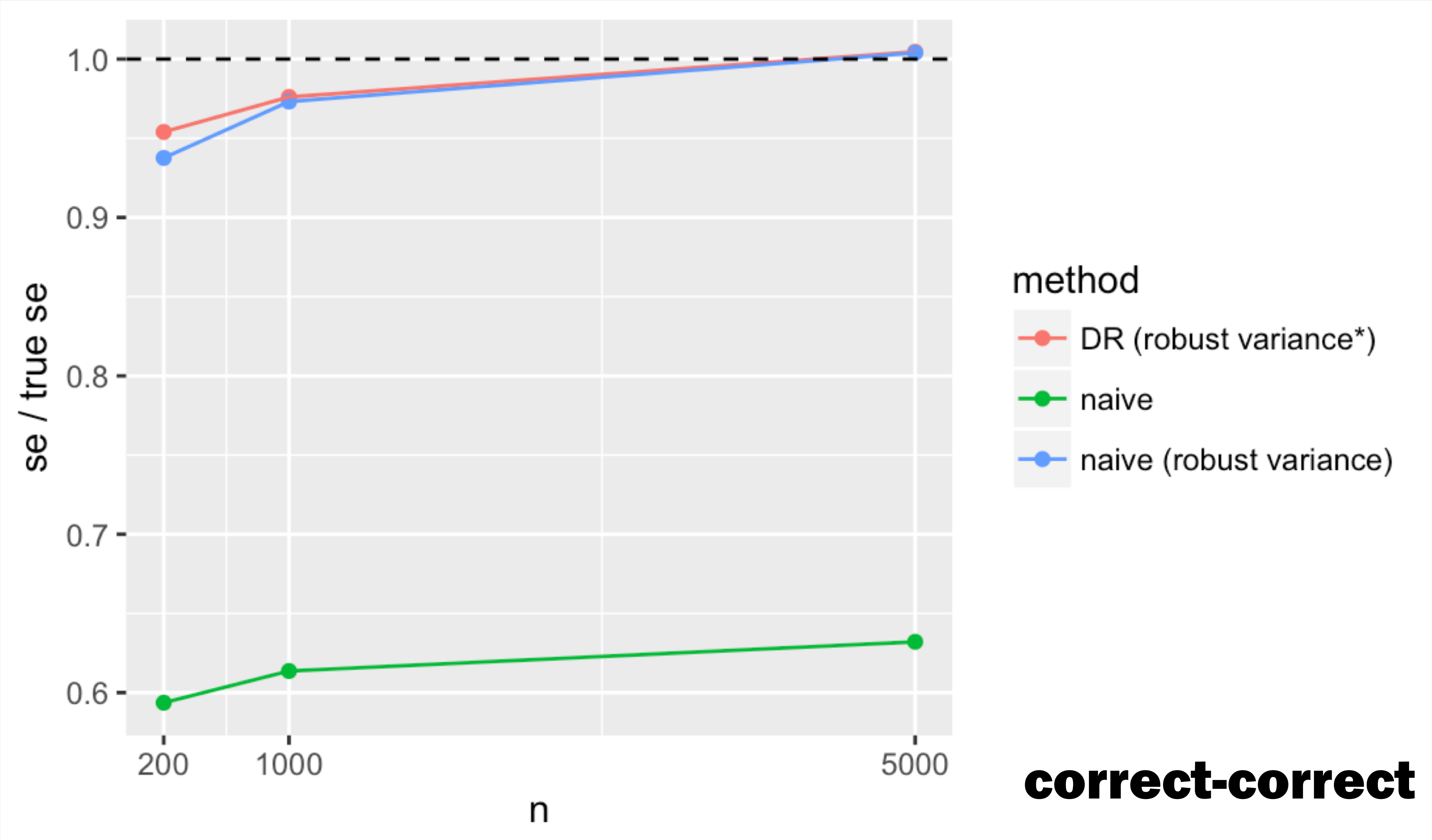
$$\mathbf{X} \sim MVN\left(\begin{pmatrix}0.5\\1\end{pmatrix},\begin{pmatrix}2 & 1\\1 & 1\end{pmatrix}\right)$$

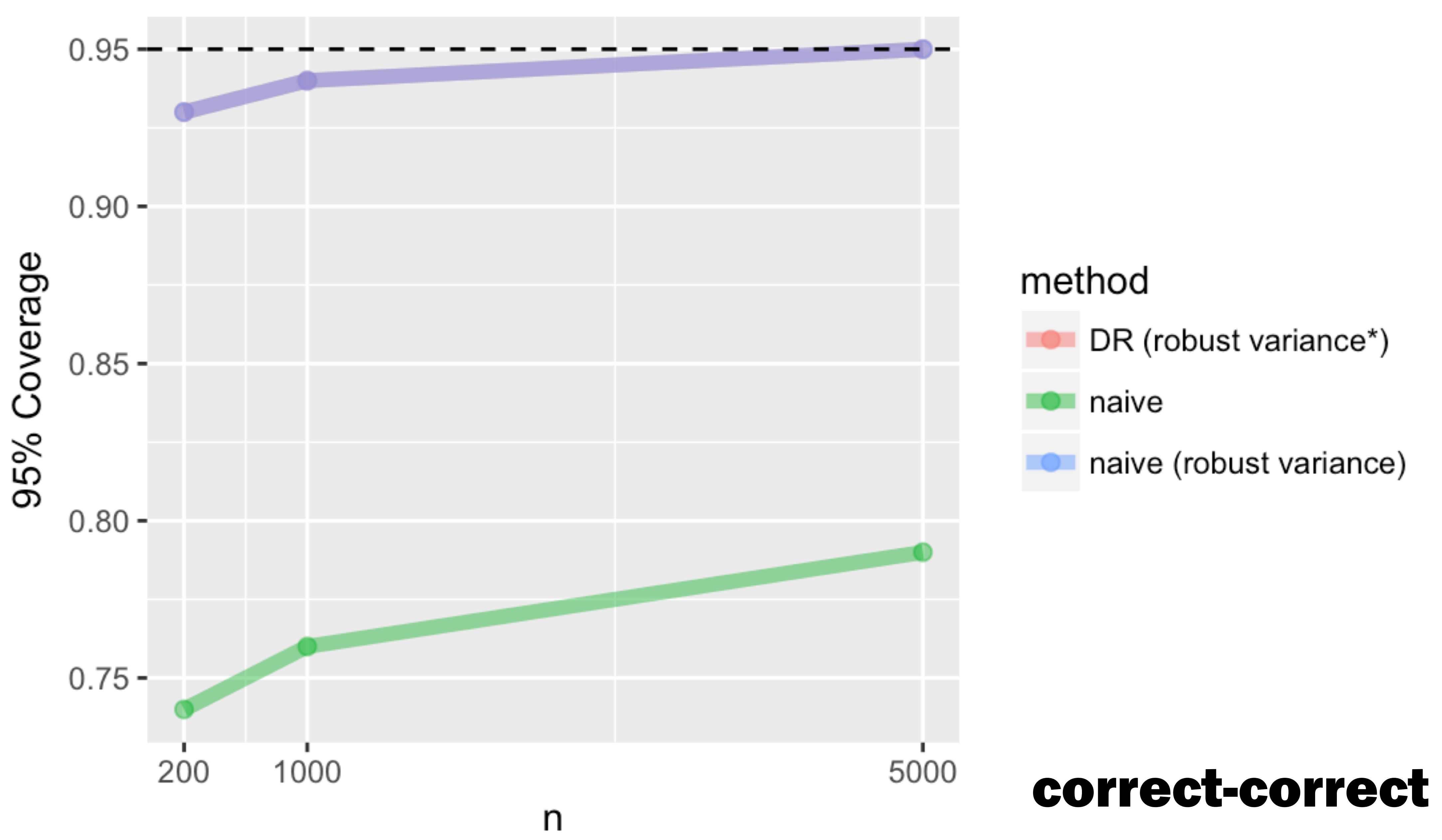
$$V \sim N\,(0,1)$$

$$Z = \begin{cases}1 & 0.5 + 0.25X_1 + 0.75X_2 + V > 0\\0 & o.w.\end{cases}$$

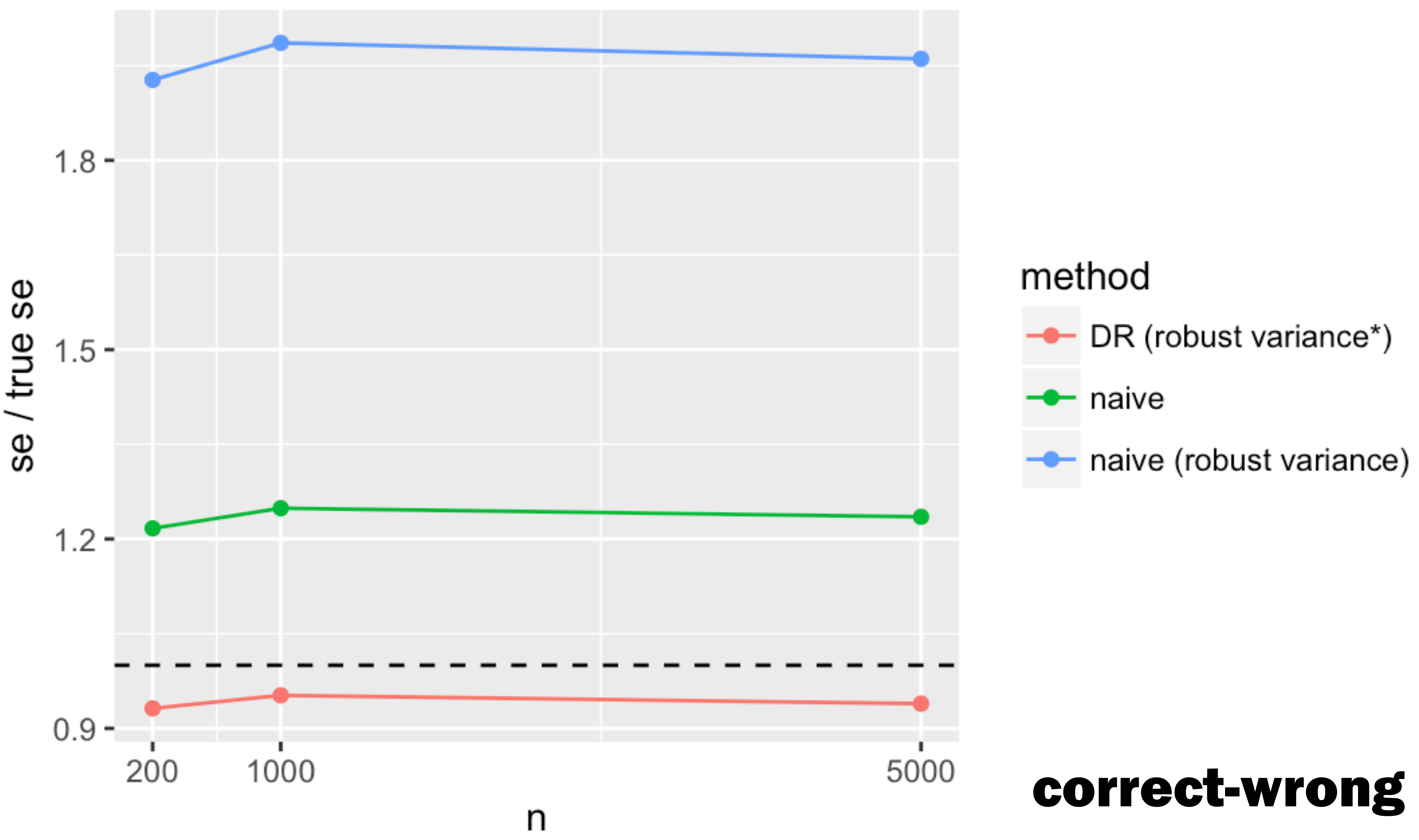
$$U \sim N(0,1)$$

$$Y = 1 + Z + X_1 + 3X_2 + U$$

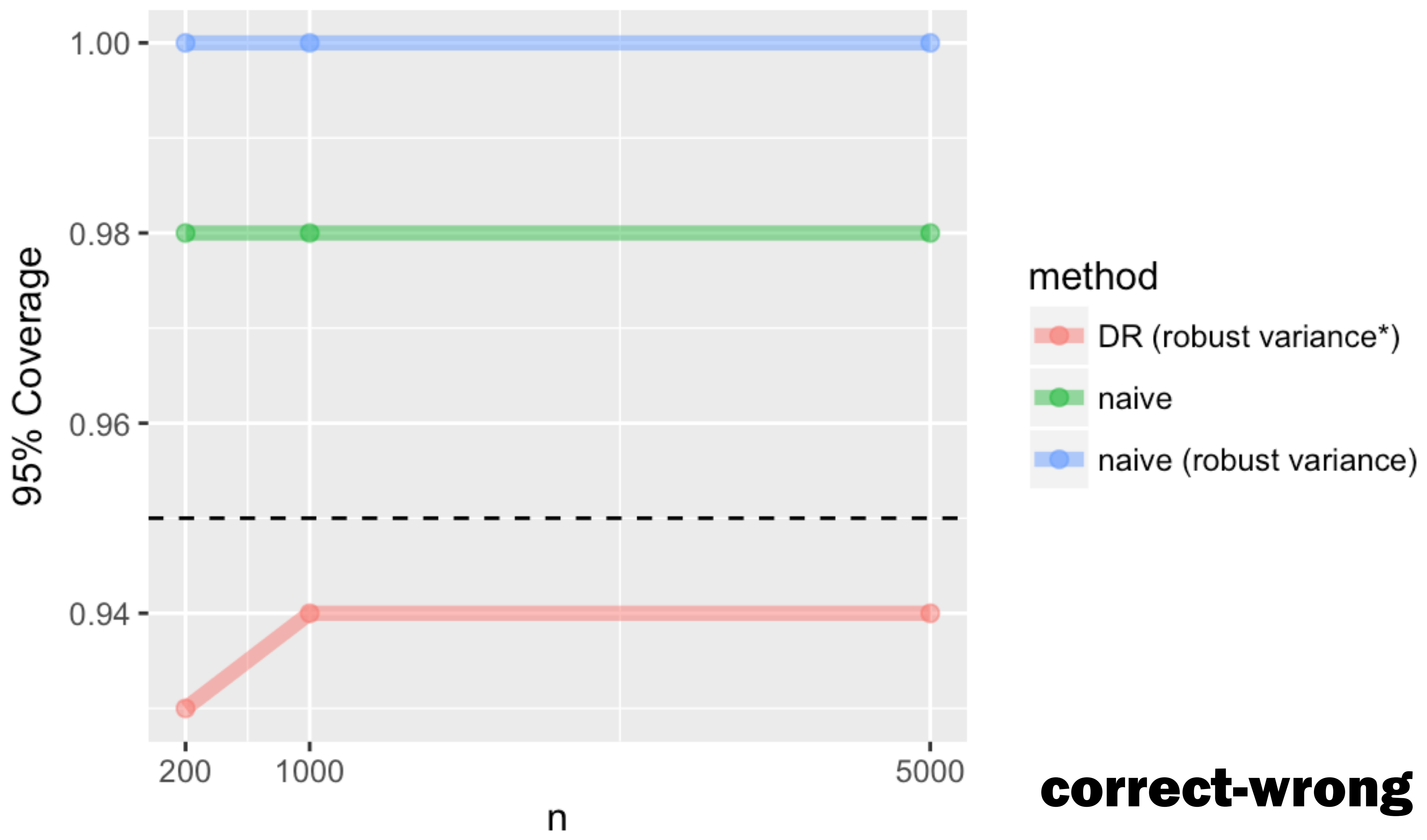

1.0
0.9
0.8
0.7
0.6
se / true se
200
1000
5000
n
method
DR (robust variance*)
naive
naive (robust variance)
correct-correct



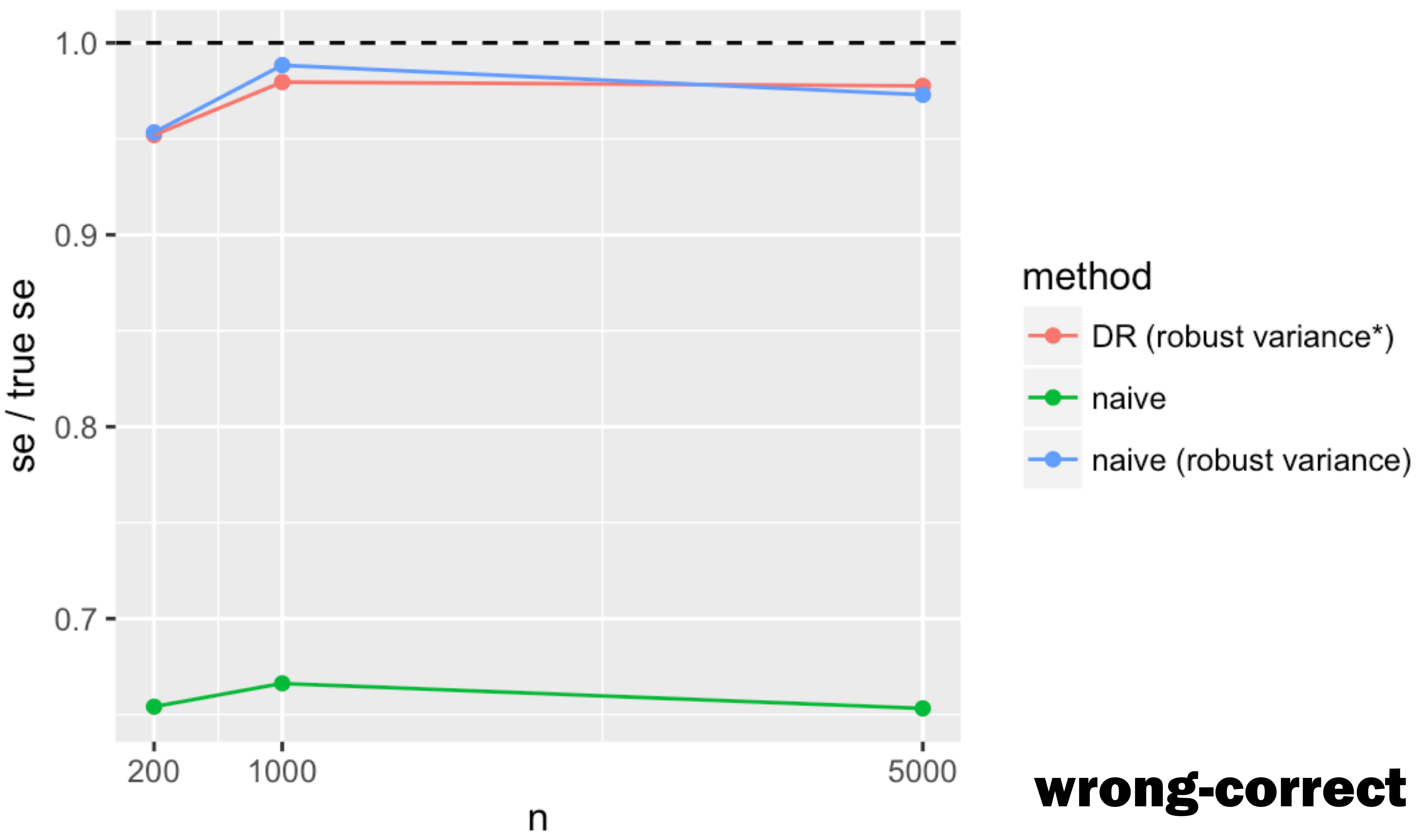
correct-correct



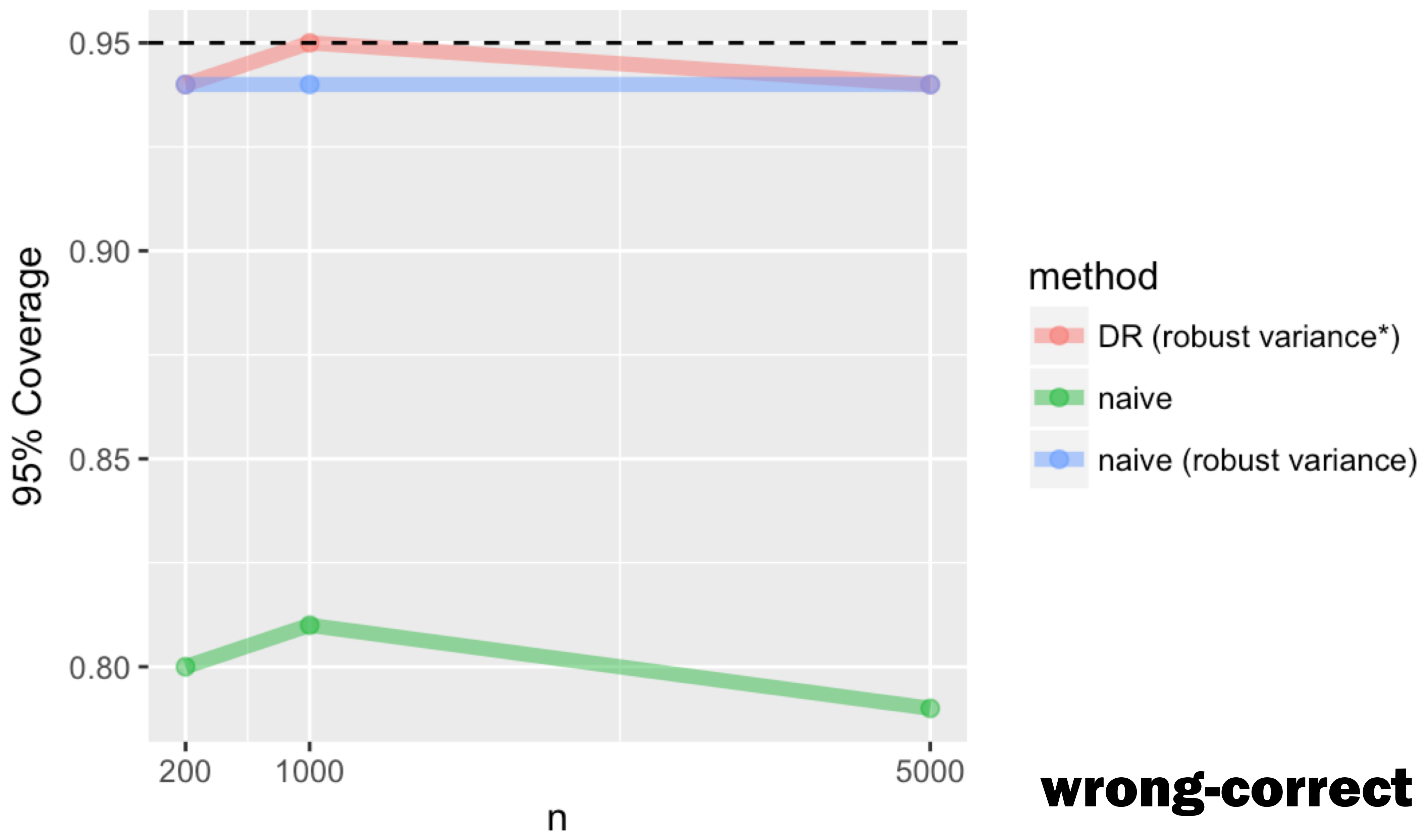
correct-wrong



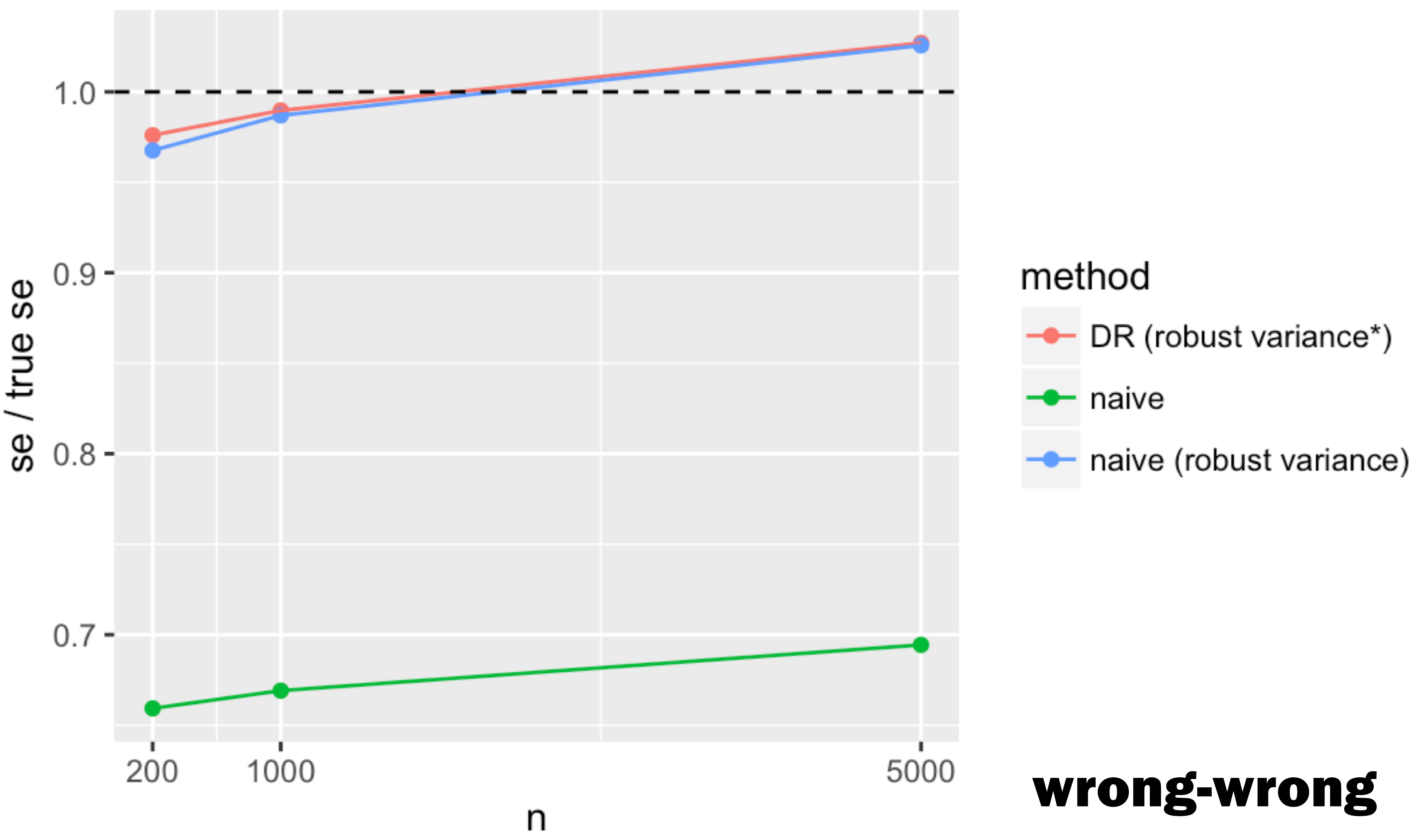
correct-wrong



wrong-correct

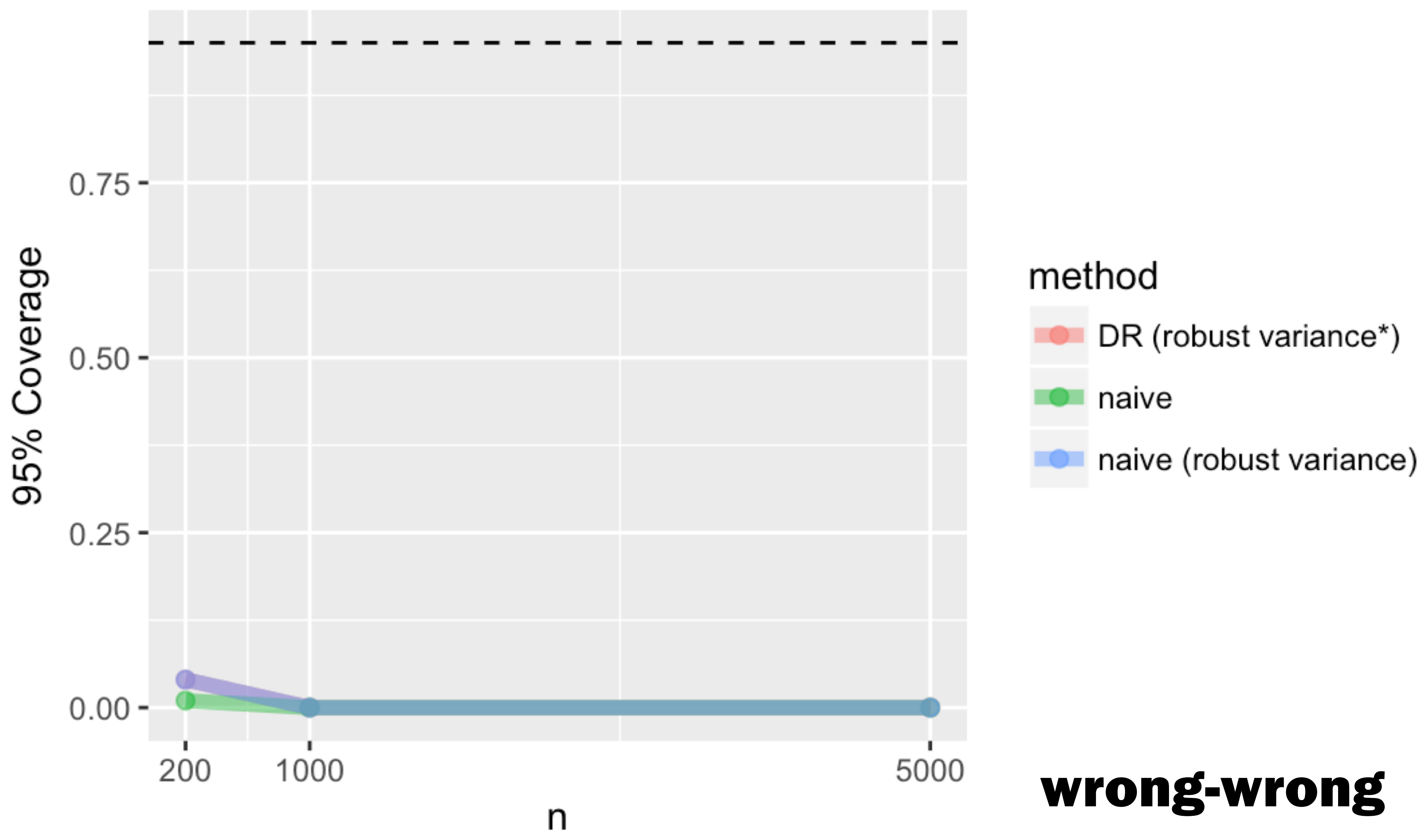


wrong-correct



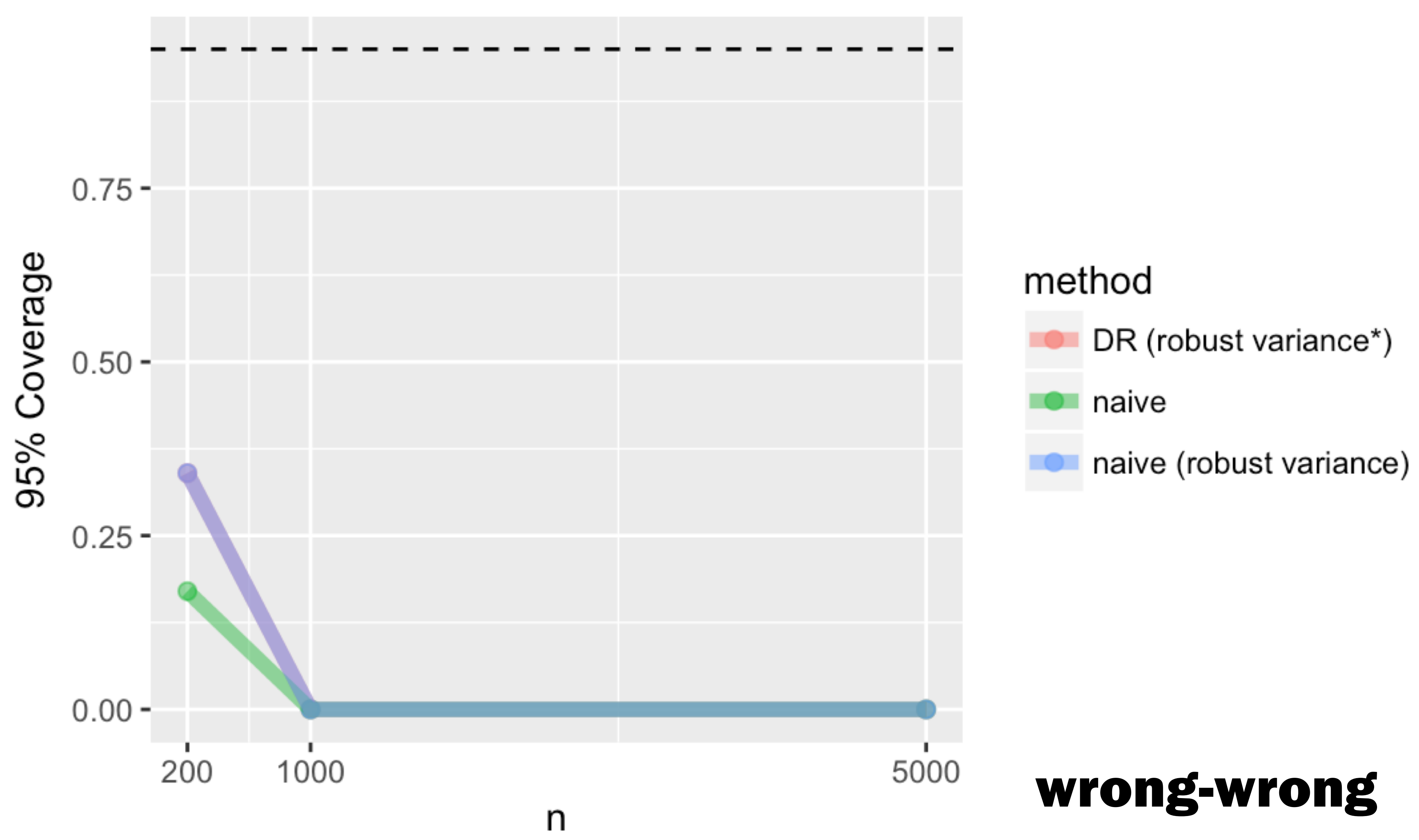
1.0
0.9
0.8
0.7
se / true se
200
1000
5000
n
method
DR (robust variance*)
naive
naive (robust variance)

wrong-wrong

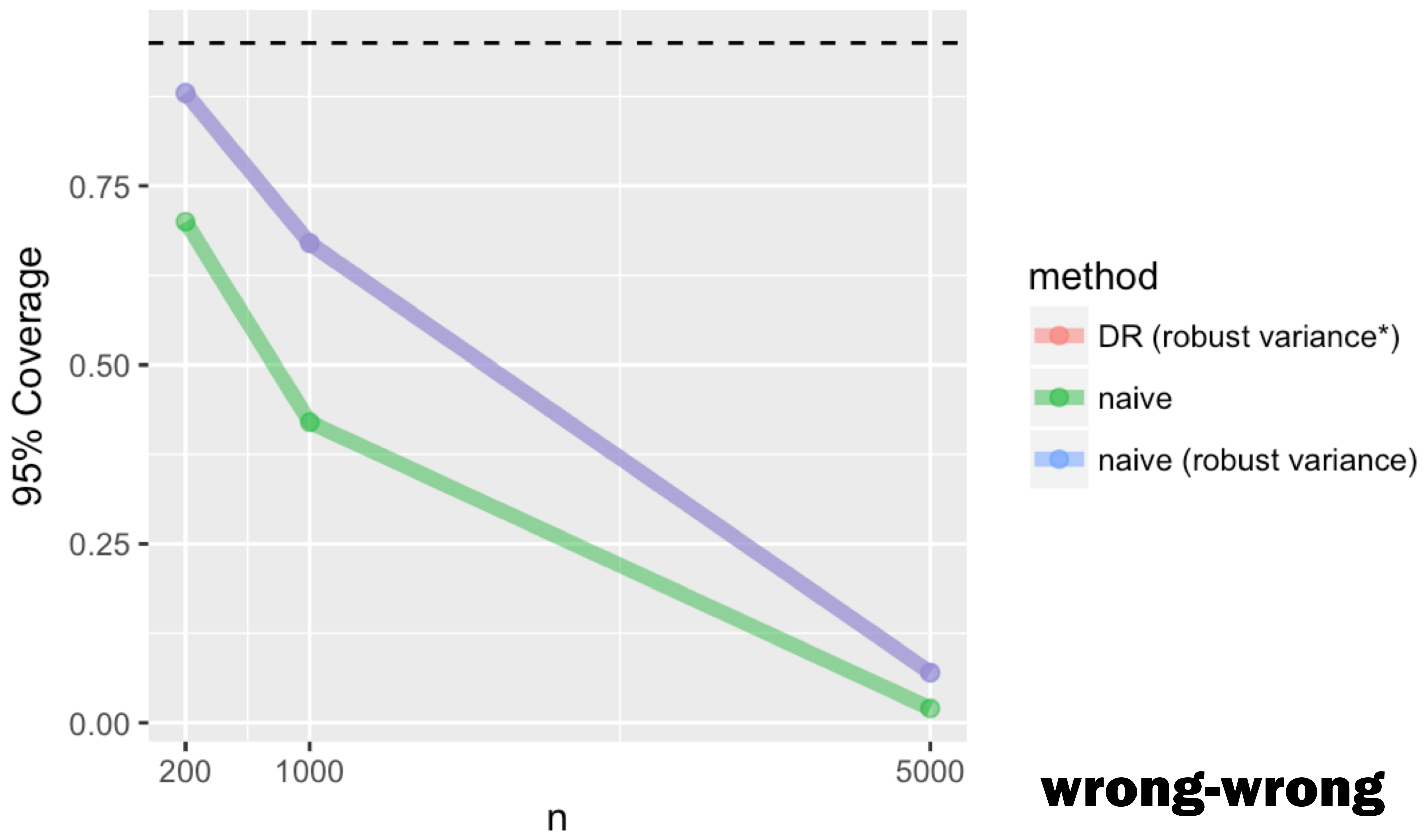


95% Coverage
0.75
0.50
0.25
0.00
200
1000
5000
n
method
DR (robust variance*)
naive
naive (robust variance)

wrong-wrong


95% Coverage
0.75
0.50
0.25
0.00
200
1000
5000
n
method
DR (robust variance*)
naive
naive (robust variance)

wrong-wrong


95% Coverage
0.00
0.25
0.50
0.75
200
1000
5000
n
method
DR (robust variance*)
naive
naive (robust variance)

wrong-wrong

Recap

- **Derived** the DR estimator and large sample variance for ATO
- **Compared** to conventional methods
- When the **outcome model is correctly specified**, robust standard errors obtained from a generalized linear model perform well, **otherwise** our **large-sample variance is preferred**
- If a confounder is completely left out, we need **alternative methods** to adjust for the **bias**

Q&A

Contextualized Tipping Point Sensitivity Analyses for Unmeasured Confounding

Dissertation Defense

Motivation

- Review of 90 observational studies JAMA, the NEJM, and the AJE
- 41 mentioned the issue of unmeasured confounding as a limitation
- 4 performed a quantitative sensitivity analysis

What we want

- Simplified
- Transparent
- Contextualized

What we want

- Simplified
- Transparent
- Contextualized

What we want

- Simplified
- Transparent
- Contextualized

Simplified

Binary unmeasured confounder

$$RR_{adj} = RR_{obs} \frac{RR_{UD}p_0 + (1 - p_0)}{RR_{UD}p_1 + (1 - p_1)}$$

Lin et al (1998)

Binary unmeasured confounder

$$RR_{adj} = RR_{obs} \frac{RR_{UD}p_0 + (1 - p_0)}{RR_{UD}p_1 + (1 - p_1)}$$

Binary unmeasured confounder

$$RR_{adj} = RR_{obs} \frac{RR_{UD} p_0 + (1 - p_0)}{RR_{UD} p_1 + (1 - p_1)}$$

Binary unmeasured confounder

$$RR_{adj} = RR_{obs} \frac{RR_{UD}p_0 + (1 - p_0)}{RR_{UD}p_1 + (1 - p_1)}$$

Binary unmeasured confounder

$$LB_{adj} = LB_{obs} \frac{RR_{UD}p_0 + (1 - p_0)}{RR_{UD}p_1 + (1 - p_1)}$$

Tipping point analyses

what will tip our confidence bound to cross 1

Binary unmeasured confounder

$$1 = LB_{obs} \frac{RR_{UD}p_0 + (1 - p_0)}{RR_{UD}p_1 + (1 - p_1)}$$

Binary unmeasured confounder

$$RR_{UD}(LB_{obs}, p_0, p_1) = \frac{(1 - p_1) + LB_{obs}(p_0 - 1)}{LB_{obs}p_0 - p_1}$$

Continuous unmeasured confounder

$$LB_{adj} = \frac{LB_{obs}}{RR_{UD}^{\mu_1 - \mu_0}}$$

Lin et al (1998)

Continuous unmeasured confounder

$$LB_{adj} = \frac{LB_{obs}}{RR_{UD}^{\mu_1 - \mu_0}}$$

Continuous unmeasured confounder

$$LB_{adj} = \frac{LB_{obs}}{RR_{UD}^{\delta}}$$

Continuous unmeasured confounder

$$\delta(LB_{obs}, RR_{UD}) = \frac{\log(LB_{obs})}{\log(RR_{UD})}$$

Continuous unmeasured confounder

$$n(LB_{obs}, RR_{UD}, \delta) = \frac{\log(LB_{obs})}{\delta \log(RR_{UD})}$$

“Assumption free”

$$1 = LB_{obs} \frac{RR_{UD}/RR_{EU} + (1 - 1/RR_{EU})}{RR_{UD}}$$

Ding and VanderWeele (2016)

"Assumption free"

$$1 = LB_{obs}\frac{RR_{UD}/RR_{EU} + (1 - 1/RR_{EU})}{RR_{UD}}$$

$$\boxed{RR_{EU} = \frac{p_1}{p_0}}$$

"Assumption free"

$$1 = LB_{obs}\frac{RR_{UD}/RR_{EU} + (1 - 1/RR_{EU})}{RR_{UD}}$$

$$\boxed{RR_{EU} = \frac{1}{p_0}}$$

E-value

$$\mathrm{E-value} = LB_{obs} + \sqrt{LB_{obs} \times (LB_{obs} - 1)}$$

Ding and VanderWeele (2017)

Adjusted E-value

$$\mathrm{E-value}_{adj} = \frac{LB_{obs}}{LB_{adj}} + \sqrt{\frac{LB_{obs}}{LB_{adj}} \times \left(\frac{LB_{obs}}{LB_{adj}} - 1\right)}$$

Transparent

Demonstrate balance

- Table 1
- Standardized mean differences / Love plots
- Side-by-side boxplots / ECDFs
- Observed bias plots

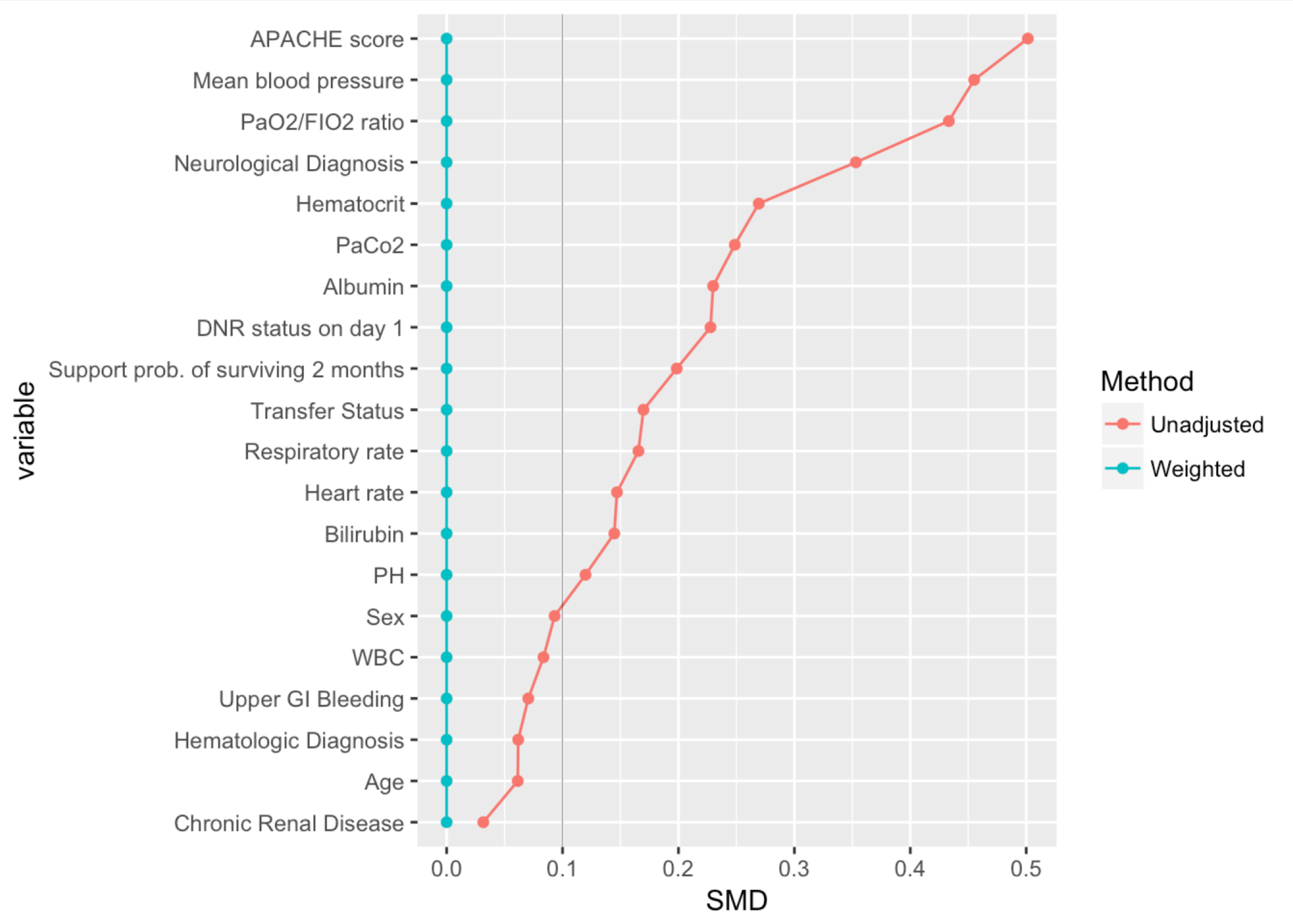
Right Heart Catheterization data

- We chose 20 covariates for demonstration purposes
  - demographics
  - comorbidities
  - physiological measurements
  - diagnosis categories
  - APACHE score
  - SUPPORT (probability of surviving 2 months)
  - DNR status on day 1

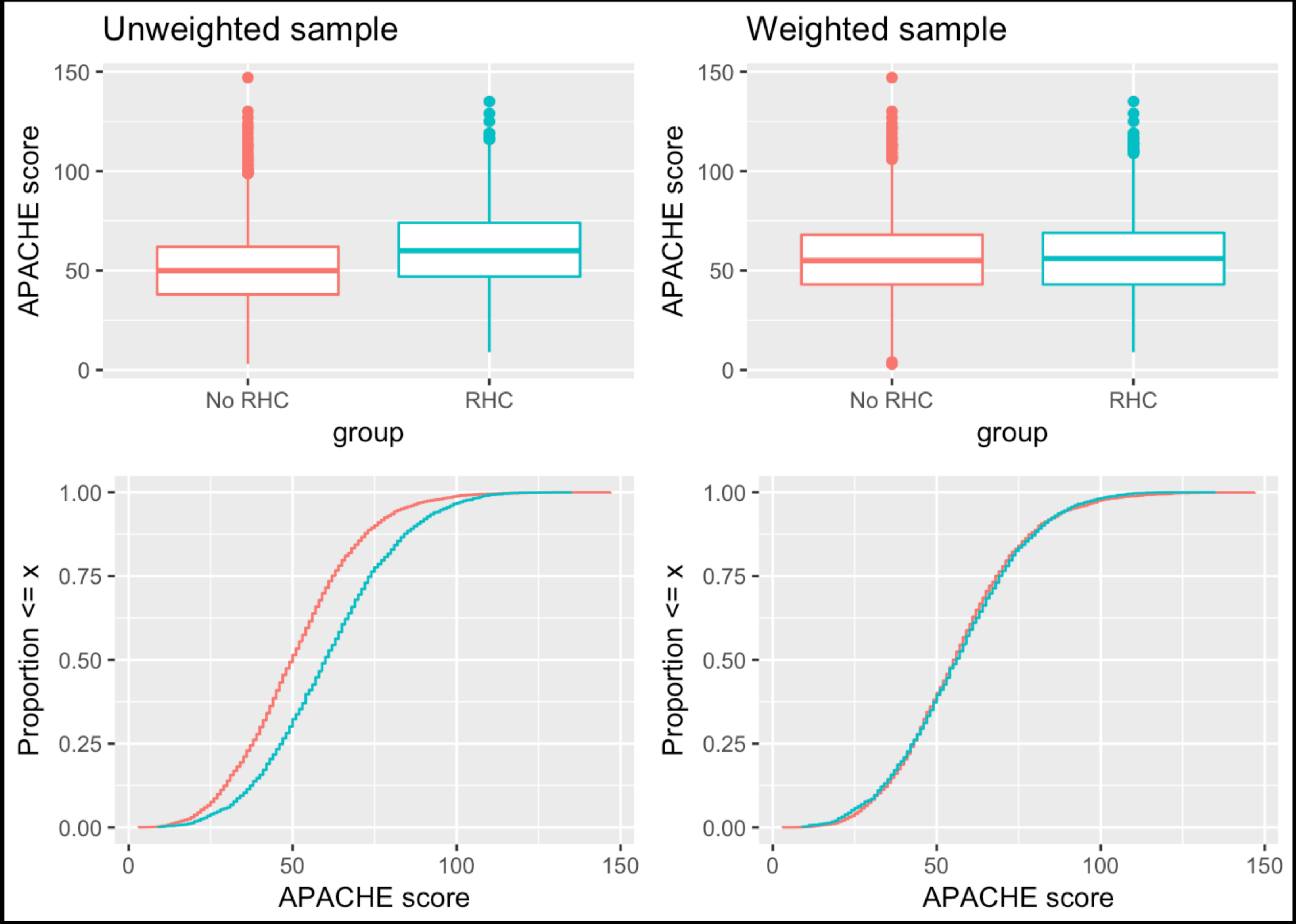
Connors et al (2006)

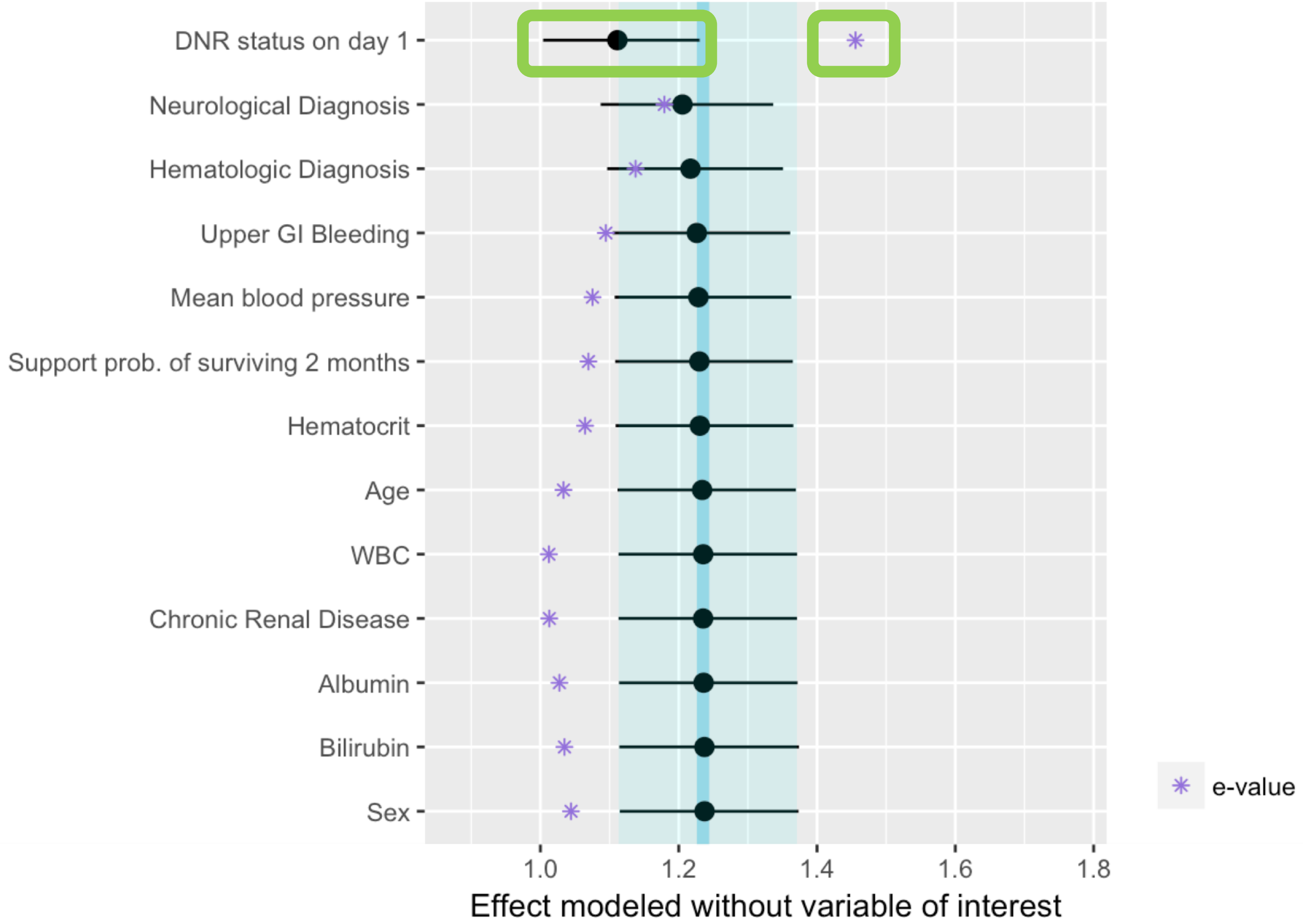
Right Heart Catheterization data

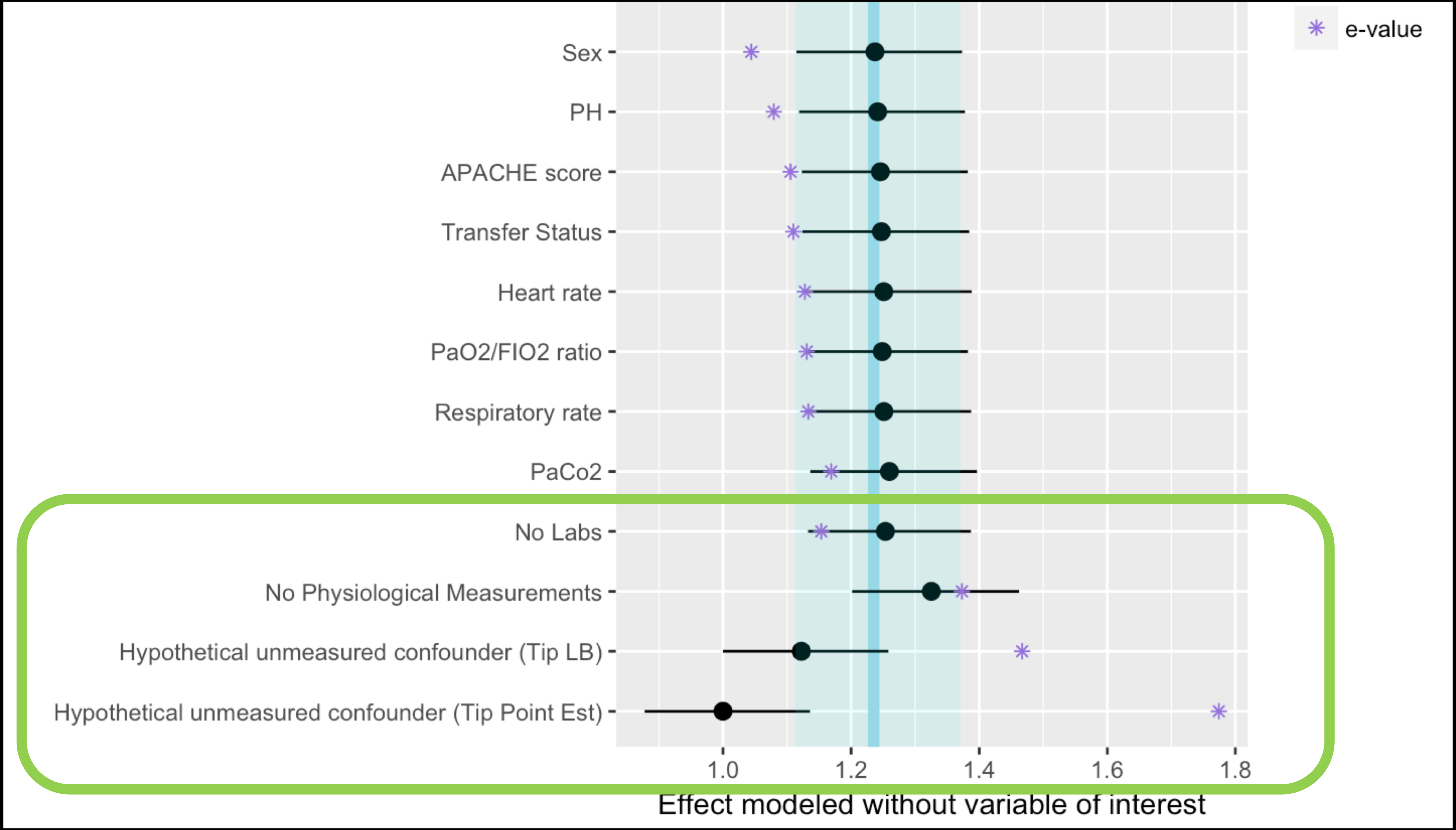
- Fit a propensity score model
- Use ATO weights
- Fit weighted cox model for 30 day survival



APACHE score
Mean blood pressure
PaO2/FIO2 ratio
Neurological Diagnosis
Hematocrit
PaCo2
Albumin
DNR status on day 1
Support prob. of surviving 2 months
Transfer Status
Respiratory rate
Heart rate
Bilirubin
PH
Sex
WBC
Upper GI Bleeding
Hematologic Diagnosis
Age
Chronic Renal Disease
variable
0.0
0.1
0.2
0.3
0.4
0.5
SMD
Method
Unadjusted
Weighted


Unweighted sample
Weighted sample
APACHE score
150
100
50
0
No RHC
RHC
group
Proportion <= x
1.00
0.75
0.50
0.25
0.00
0
50
100
150
APACHE score


DNR status on day 1
Neurological Diagnosis
Hematologic Diagnosis
Upper GI Bleeding
Mean blood pressure
Support prob. of surviving 2 months
Hematocrit
Age
WBC
Chronic Renal Disease
Albumin
Bilirubin
Sex
1.0
1.2
1.4
1.6
1.8
Effect modeled without variable of interest
e-value



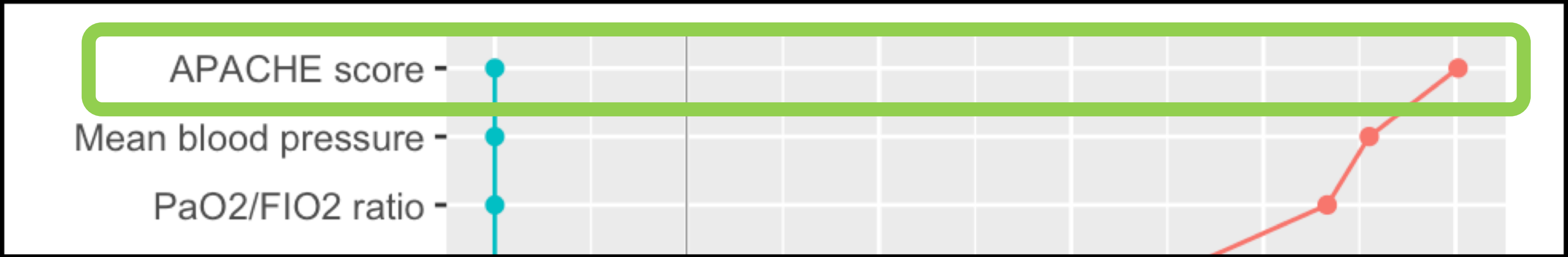
e-value
Sex
PH
APACHE score
Transfer Status
Heart rate
PaO2/FIO2 ratio
Respiratory rate
PaCo2
No Labs
No Physiological Measurements
Hypothetical unmeasured confounder (Tip LB)
Hypothetical unmeasured confounder (Tip Point Est)
1.0
1.2
1.4
1.6
1.8
Effect modeled without variable of interest

Contextualize

R package

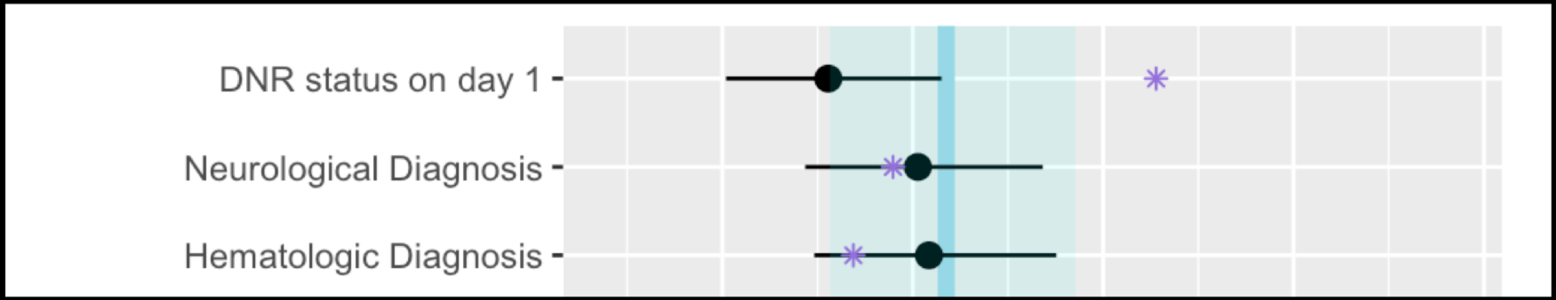
- tipr
- tip_with_binary()
- tip_with_continuous()

Scenario 1



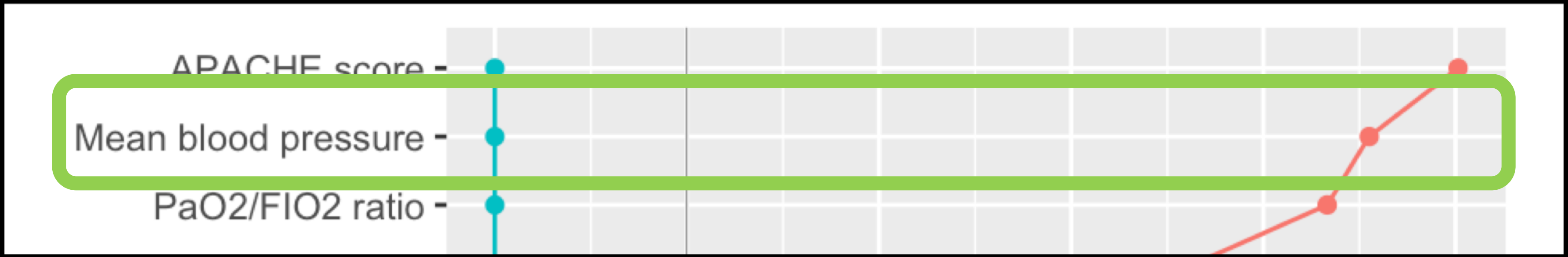

```r
library(tipr)
tip_with_continuous(mean_diff = 0.29,
                    lb = 1.11,
                    ub = 1.37)
#> [1] 1.433132
```

Scenario 2



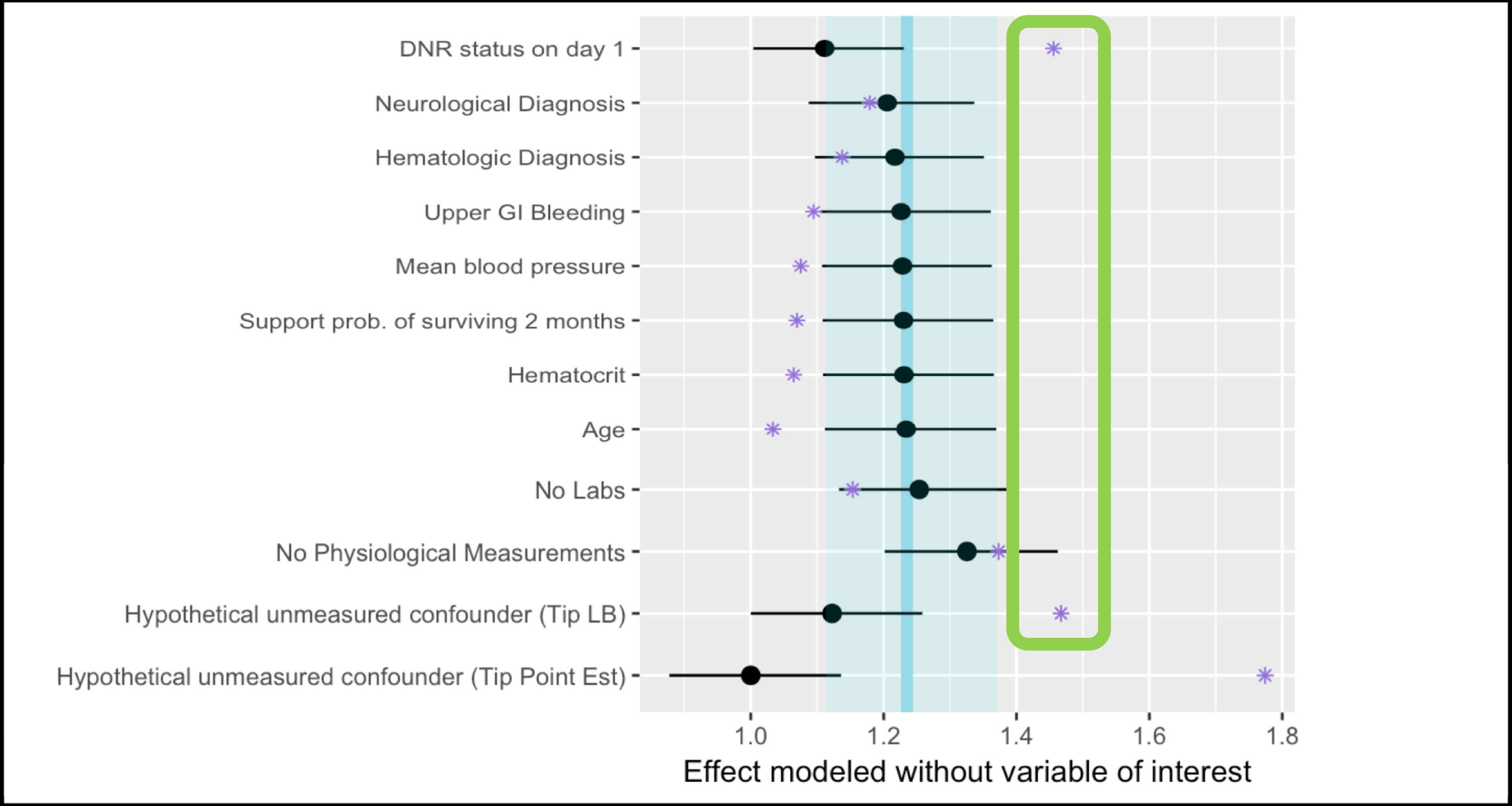

```r
library(tipr)
tip_with_continuous(p_0 = 0.14,
                    gamma = 2.59,
                    lb = 1.11,
                    ub = 1.37)
#> [1] 0.2245824
```

Scenario 3




```r
library(tipr)
tip_with_continuous(mean_diff = -0.19,
                    gamma = 1/1.05,
                    lb = 1.11,
                    ub = 1.37)
#> [1] 11.25766
```

Scenario 4


DNR status on day 1
Neurological Diagnosis
Hematologic Diagnosis
Upper GI Bleeding
Mean blood pressure
Support prob. of surviving 2 months
Hematocrit
Age
No Labs
No Physiological Measurements
Hypothetical unmeasured confounder (Tip LB)
Hypothetical unmeasured confounder (Tip Point Est)
1.0
1.2
1.4
1.6
1.8
Effect modeled without variable of interest

Recap

- **Rearranged** sensitivity analysis formulas to create a tipping point analysis
- **Adjusted** the E-value to allow it to be contextualized
- **Created** tipr, an R package to perform these analysis
- **Demonstrated** scenarios for **contextualizing** these analyses

Q&A