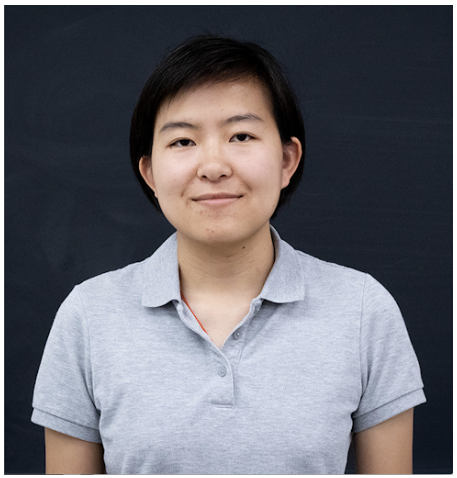
Fighting Noise with Noise

Causal Inference with Many Candidate Instruments

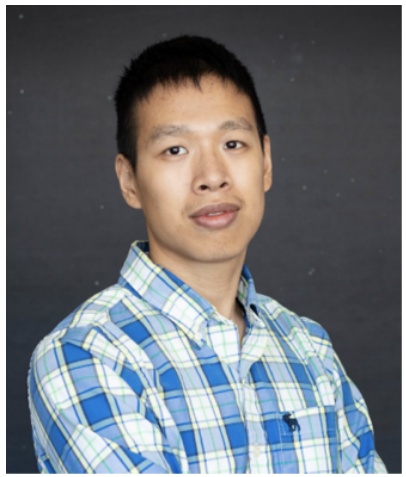
Dehan Kong
University of Toronto

Pacific causal inference conference
Sep 17, 2022

Acknowledgements



Cindy Zhang



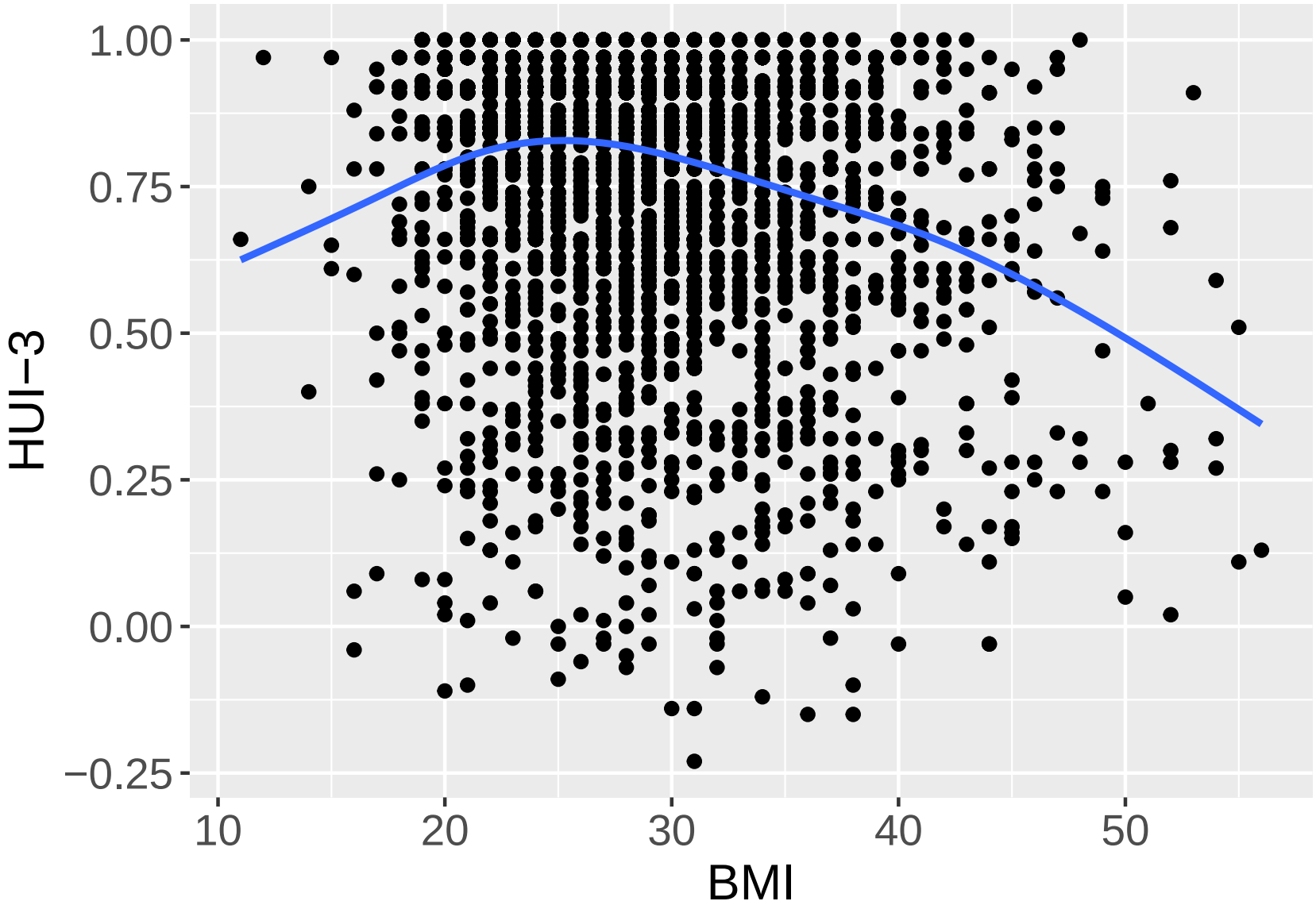
Linbo Wang

Stanislav Volgushev

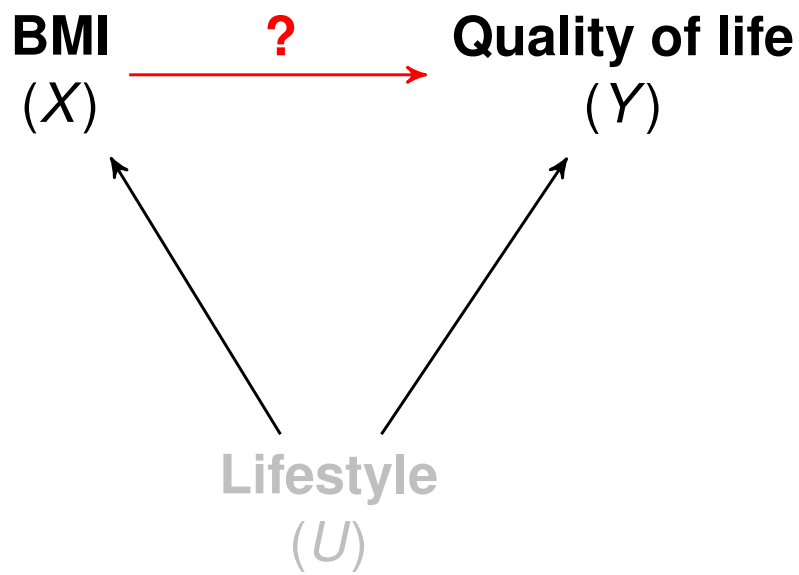
McLAUGHLIN CENTRE

Overweight

Overweight is associated with Quality of Life

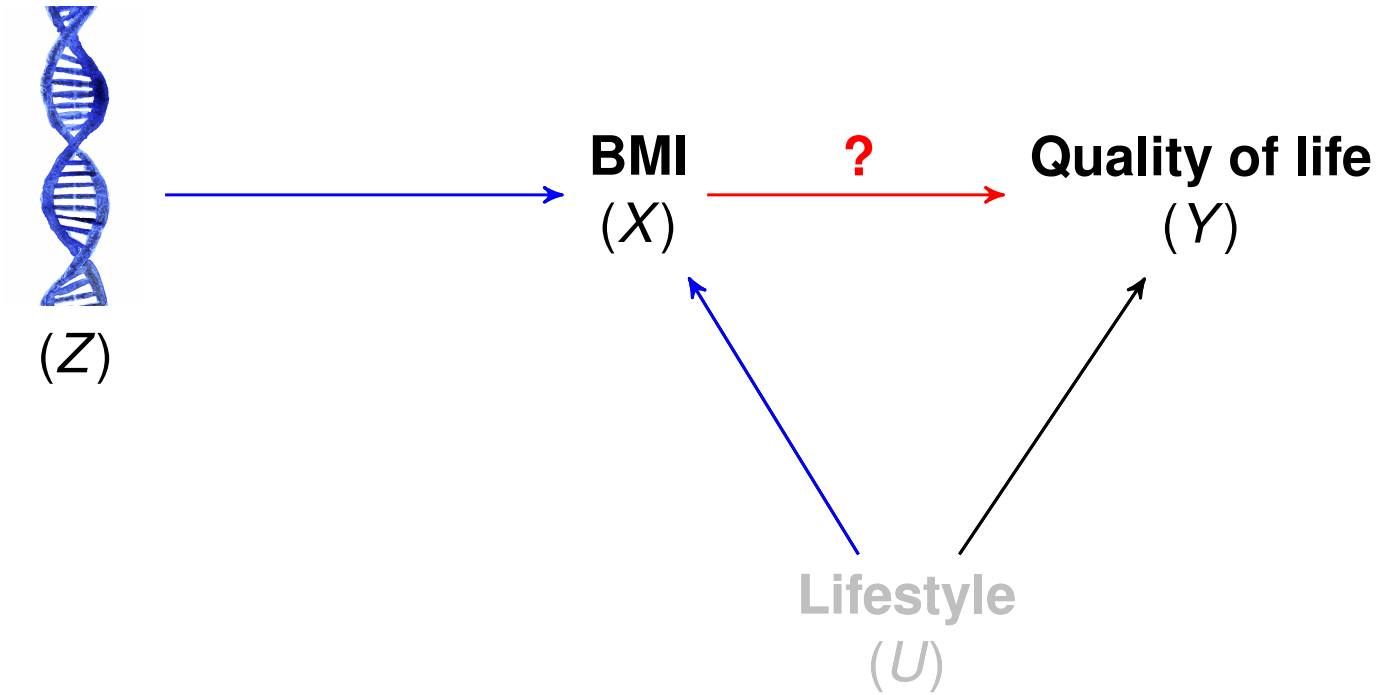


Is this association causal?



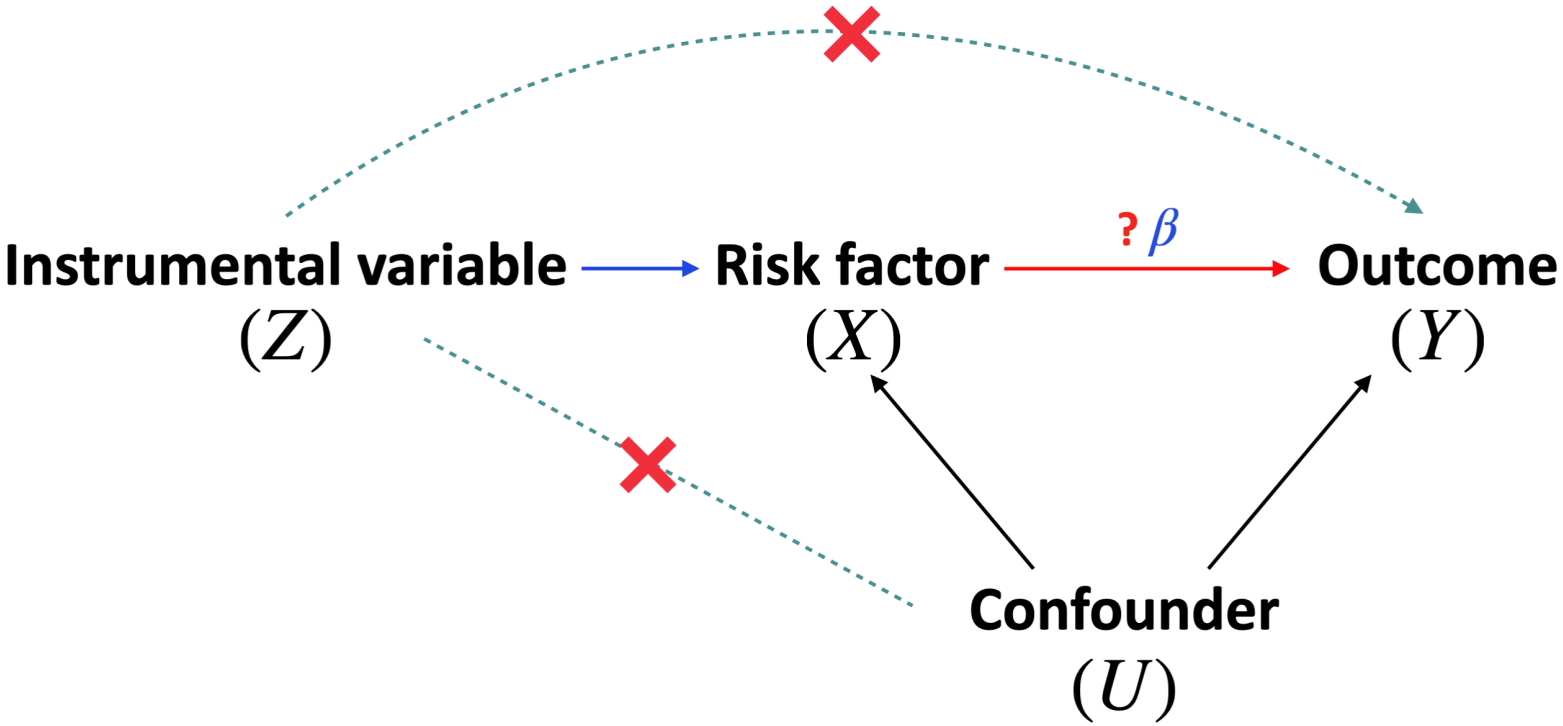
☛ **Unmeasured confounding!**

Mendelian randomization (MR)



Nature rolls the dice!

The math behind



IV1. $Z \to X \neq 0$
IV2. $Z \to Y = 0$
IV3. $\text{cov}(Z, U) = 0$

$\Longrightarrow$

$$Z \to X \to Y = (Z \to X) \times \underbrace{(X \to Y)}_{\beta}$$

$$X \to Y = \frac{Z \to X \to Y}{Z \to X}$$

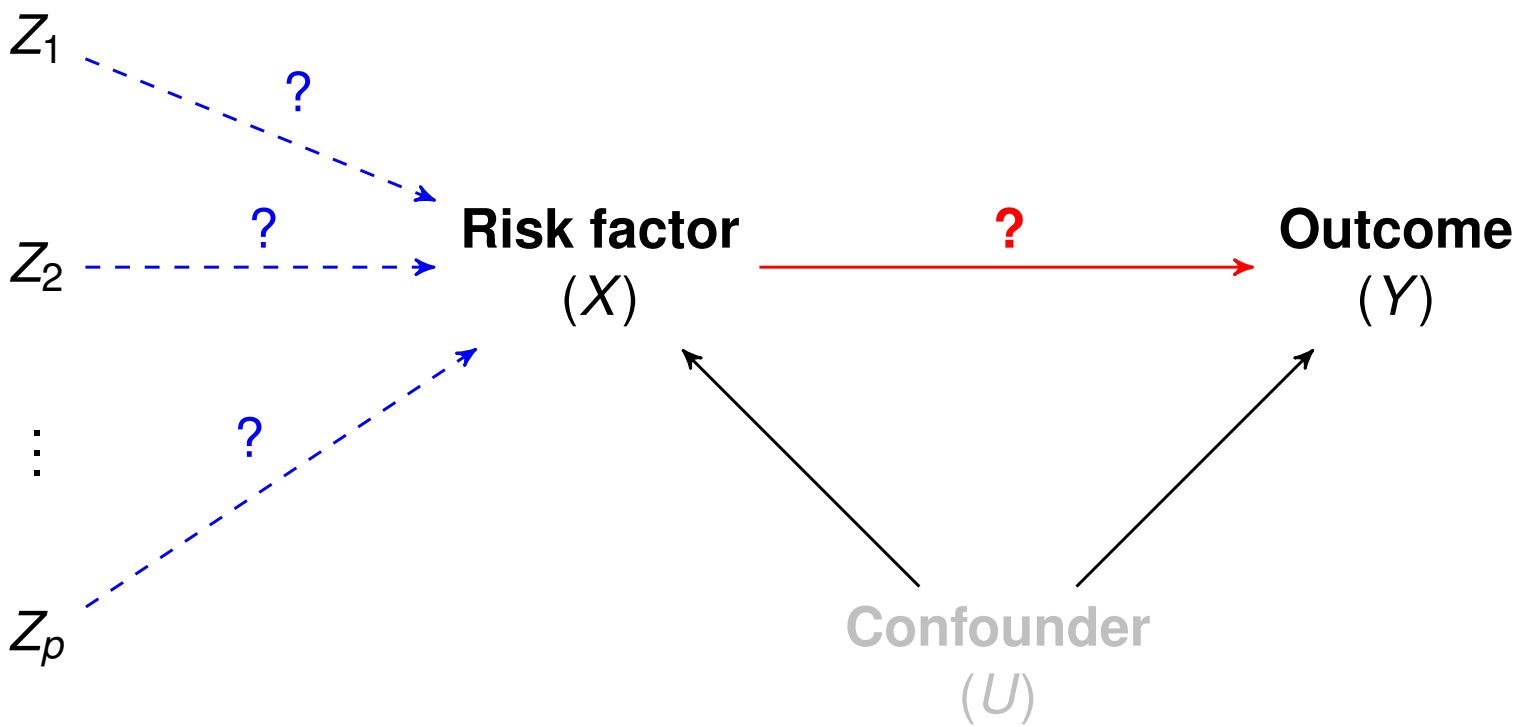
Natural experiments are imperfect!

1. Most genetic variants are irrelevant!
   - In the past: rely on expert knowledge

Natural experiments are imperfect!

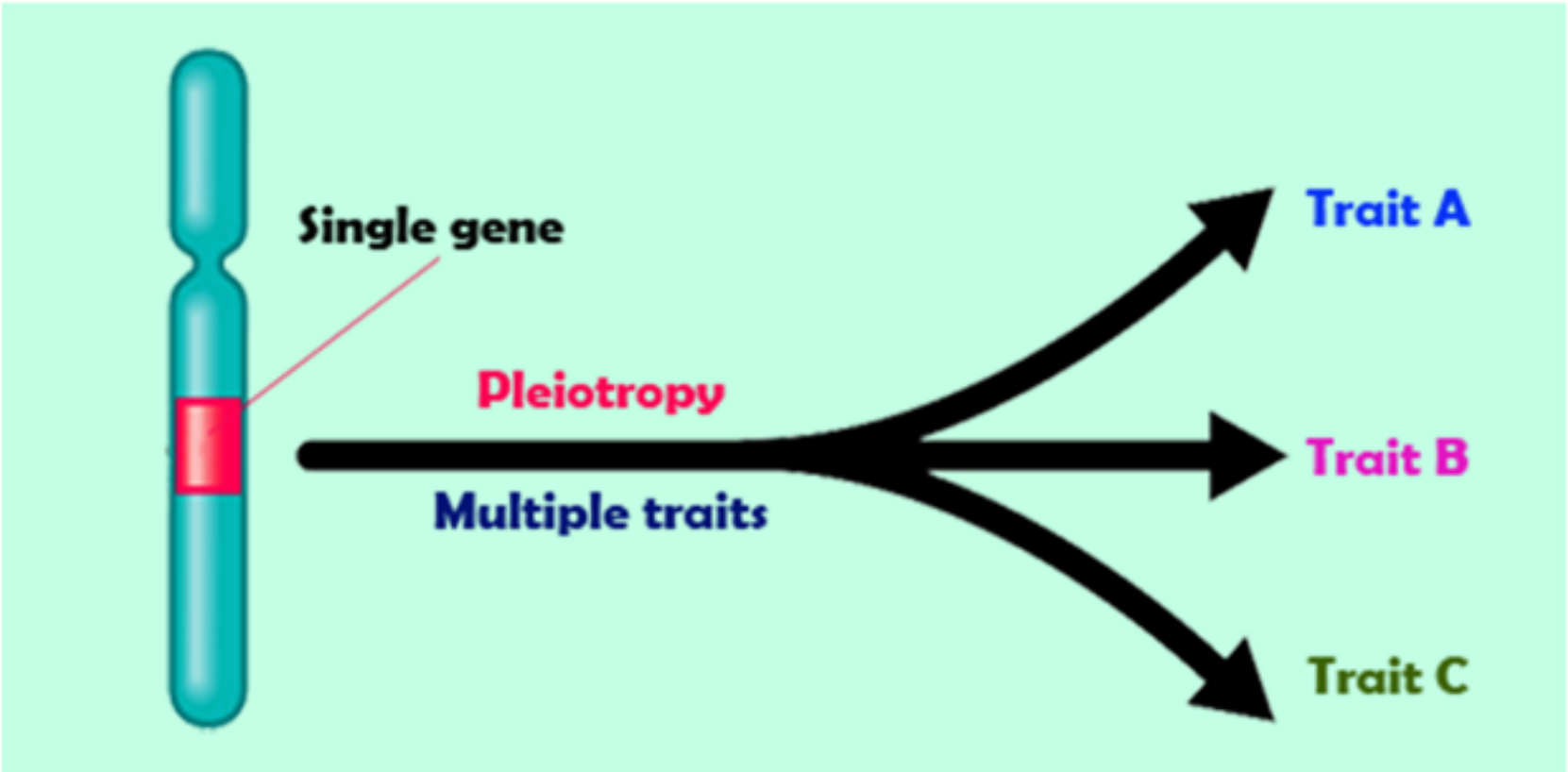
1. Most genetic variants are irrelevant!
   - In the past: rely on expert knowledge

Modern approach: GWAS



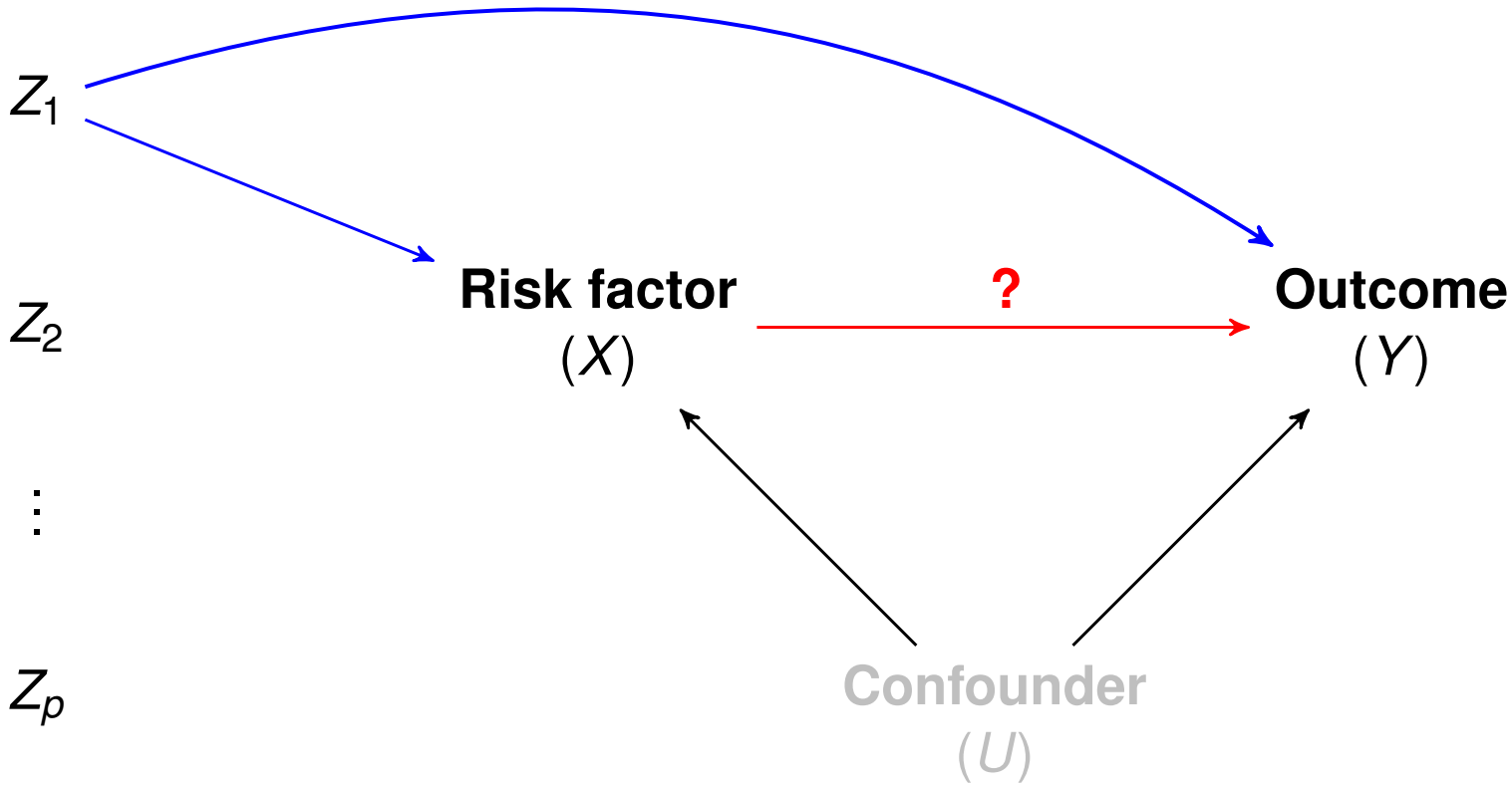
Natural experiments are imperfect!

2 Some of SNPs that pass the GWAS test may still be invalid due to pleiotropy



Natural experiments are imperfect!

2 Some of SNPs that pass the GWAS test may still be invalid due to pleiotropy



Natural experiments are imperfect!

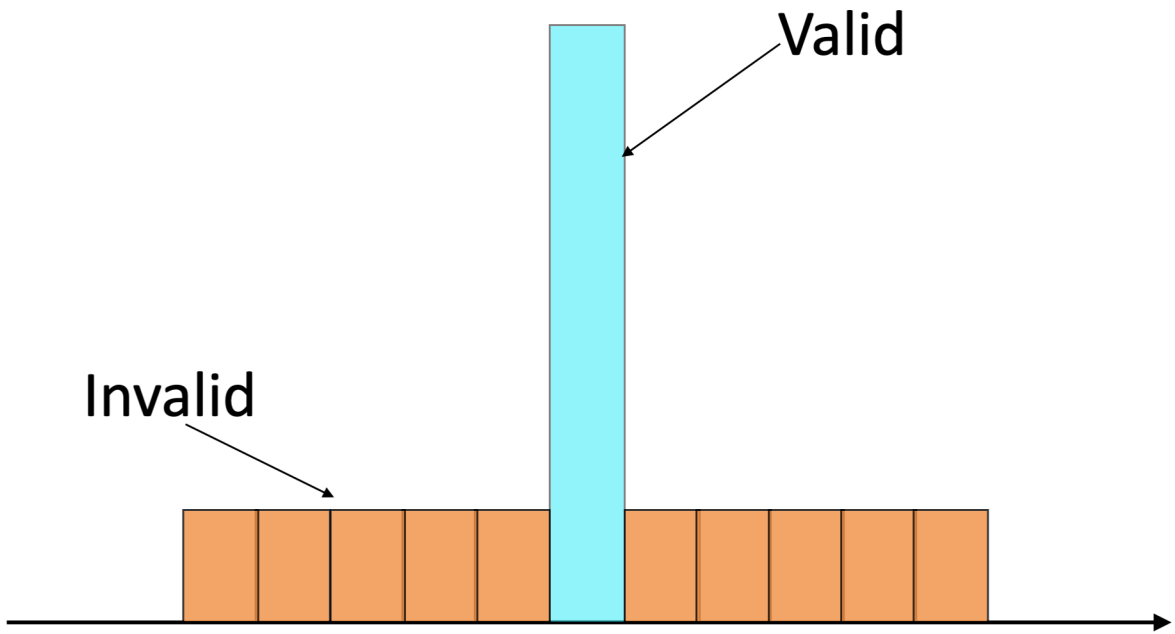
2 Some of SNPs that pass the GWAS test may still be invalid due to pleiotropy

Existing solution: Valid causal inference with some invalid instruments (Bowden et al. [2015], Kolesár et al. [2015], Kang et al. [2016], Hartwig et al. [2017], Guo et al. [2018], Hartford et al. [2020])

Valid causal inference with some invalid instruments

Solution:

All happy instruments are alike;
each unhappy instrument is unhappy in its own way



Plurality rule: Guo et al. , 2018

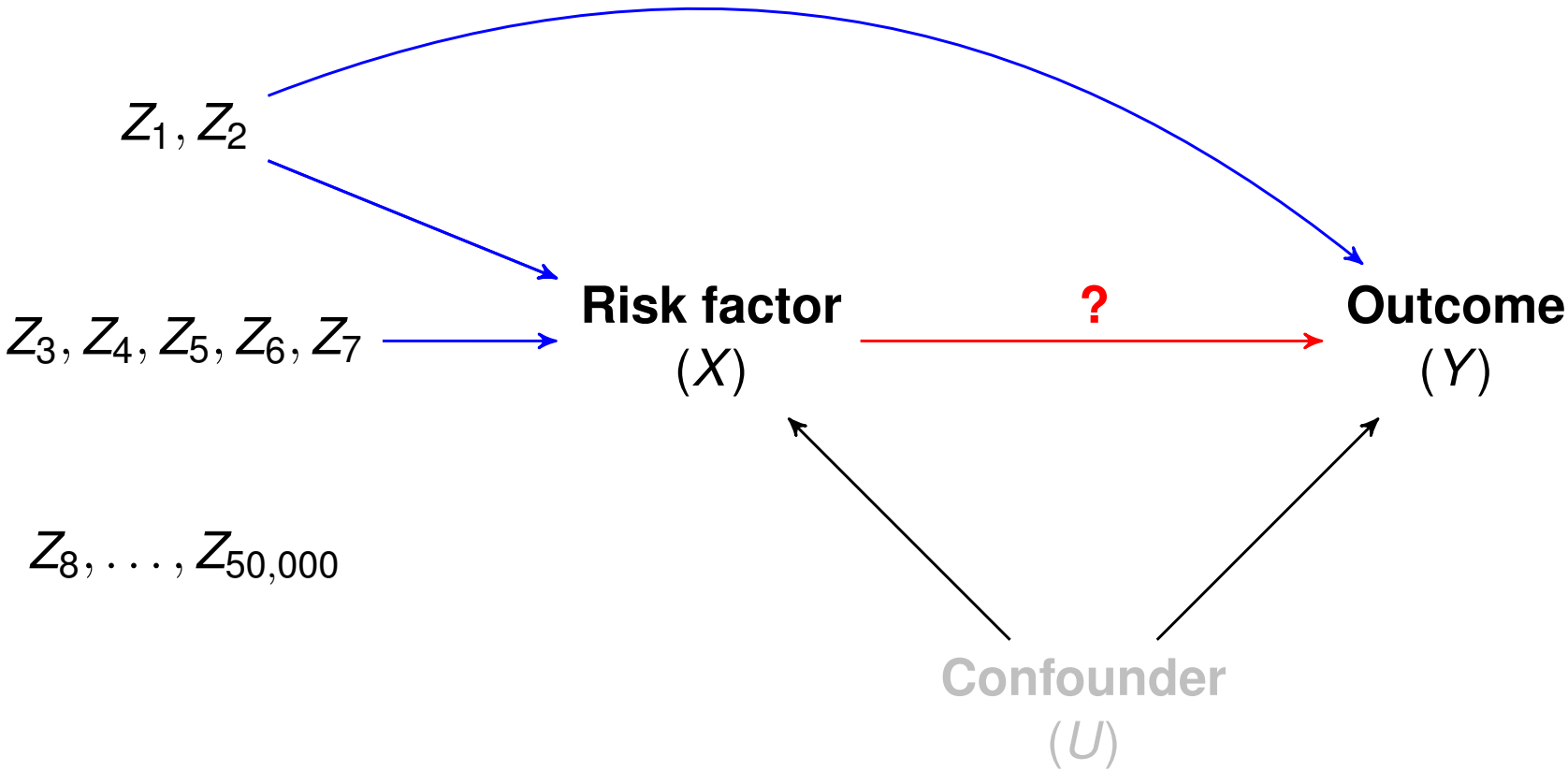
Summary of existing solutions

Step 1 GWAS based on strength of the links $Z_j \to X, j = 1, \ldots, p$

Step 2 Apply a mode-finding algorithm to identify the valid IVs

Step 3 Use the valid IVs to estimate the causal effect

A toy simulation



- Sample size = 500
- $Z_j, j = 1, \ldots, 50,000, U \sim N(0, 1)$
- Linear models
- True causal effect $X \rightarrow Y$ is 2

Step 1: GWAS

Step 1.1 (Marginal screening): Select the top 500 candidate IVs $\boldsymbol{Z}_r$ based on the marginal correlation

$$|Cor(Z_j, X)|, j = 1, \ldots, 50,000$$

Step 1.2 (Joint thresholding): Fit a debiased Lasso model

$$X \sim \boldsymbol{Z}_r$$

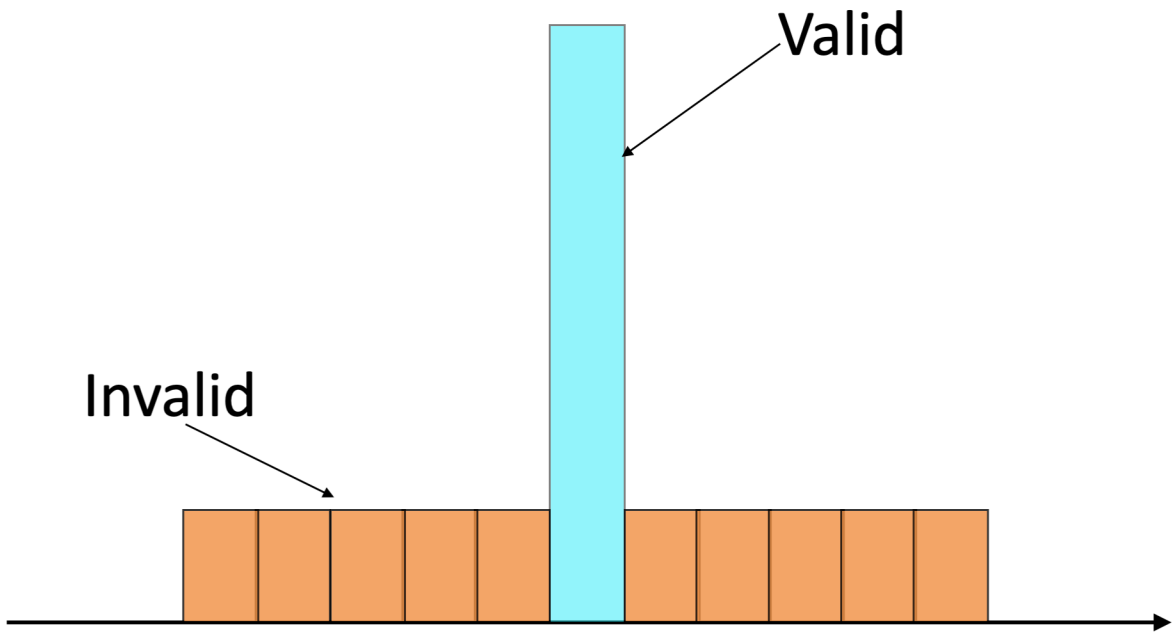
Result: On average left with 21 candidate IVs

☺ Include the 7 relevant IVs (5 valid, 2 invalid) every single time!

Step 2: Mode-finding

Solution:

All happy instruments are alike;
each unhappy instrument is unhappy in its own way



Plurality rule: Guo et al. , 2018

Step 3: Causal effect estimation

☹ The causal effect estimates are biased!

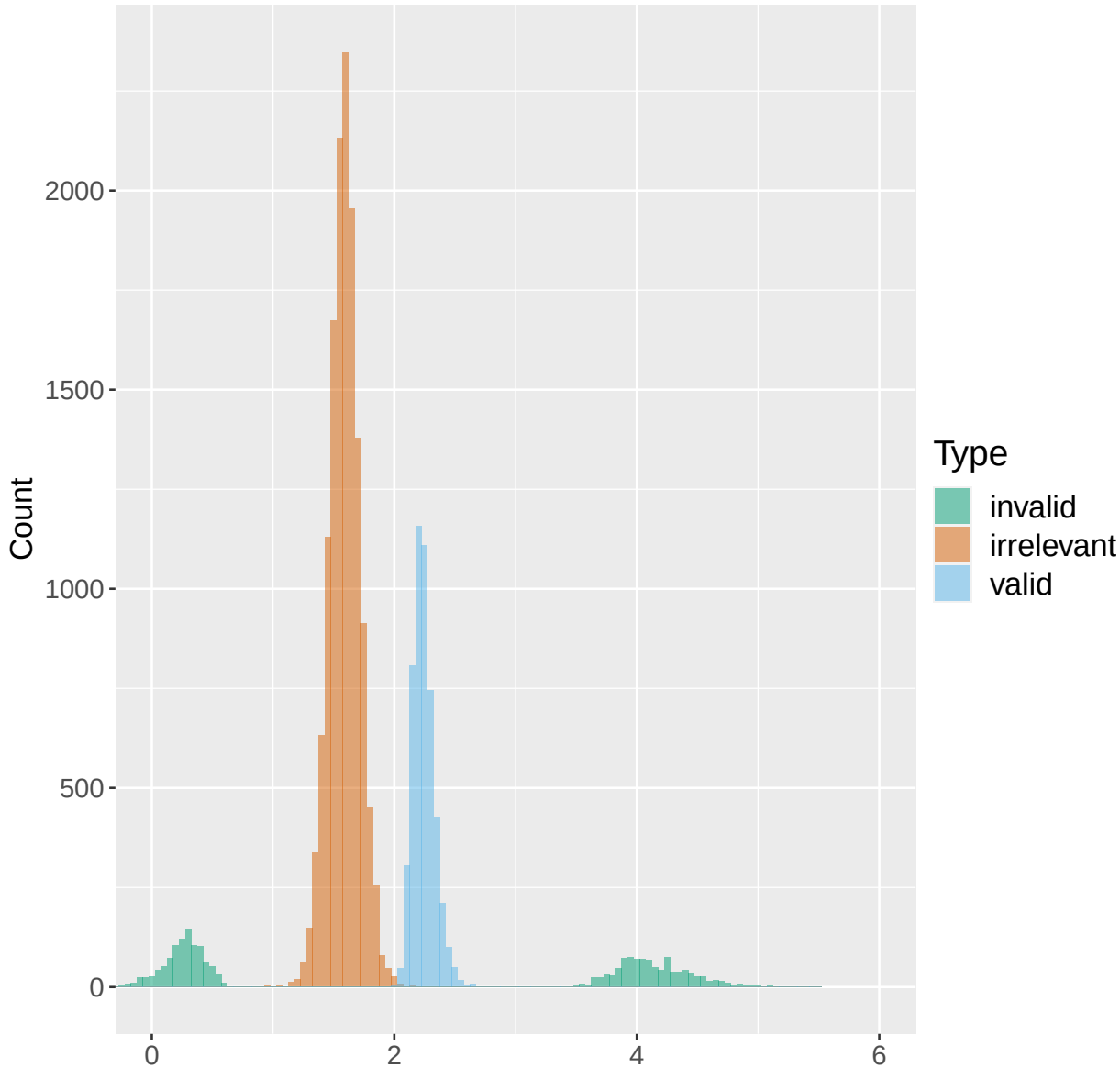
- Bias = -0.22 (SE = 0.02)
- Coverage probability = 0.07 (nominal = 0.95)

What went wrong?

Recall in Step 1: Select 21 candidate IVs

- 5 valid
- 2 invalid
- **14 irrelevant with spurious correlation**

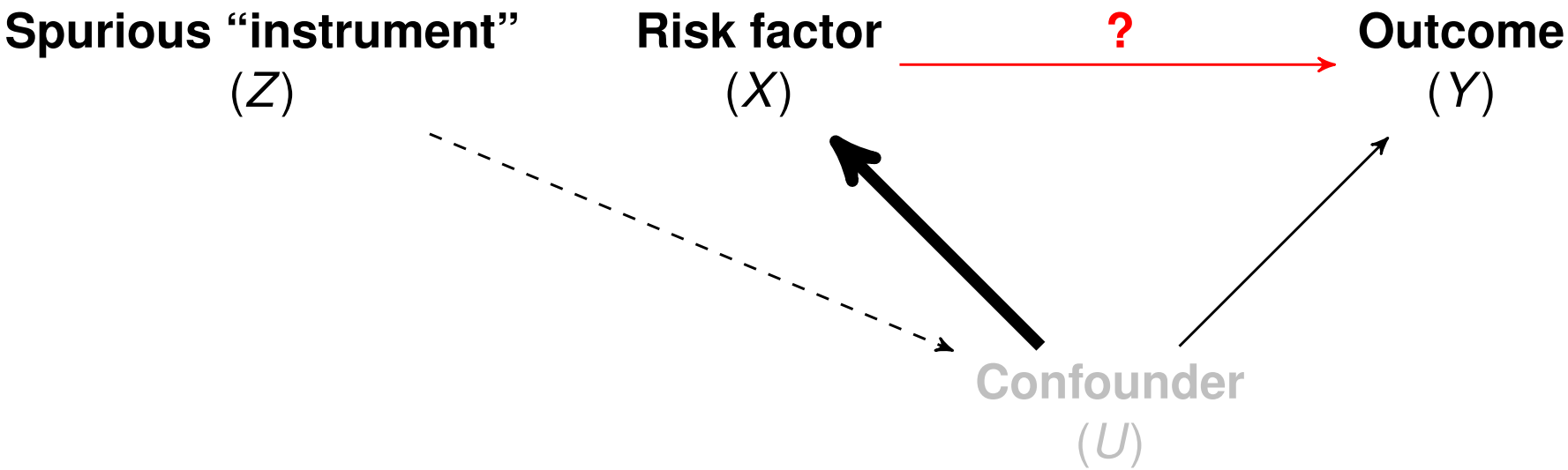
What happened in Step 2



Distribution of causal effect estimates for candidate IVs passing joint thresholding across 1000 Monte Carlo runs

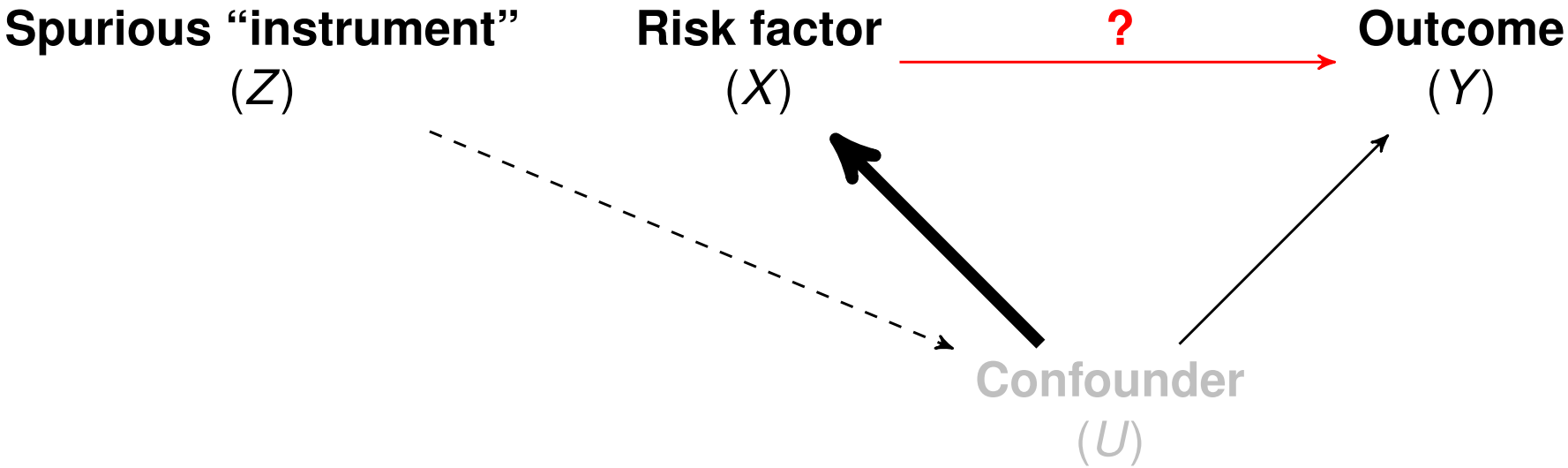
All irrelevant “instruments” are alike. Why?

The simplest case (extreme confounding)



- Single unmeasured confounder U
- X is only determined by U

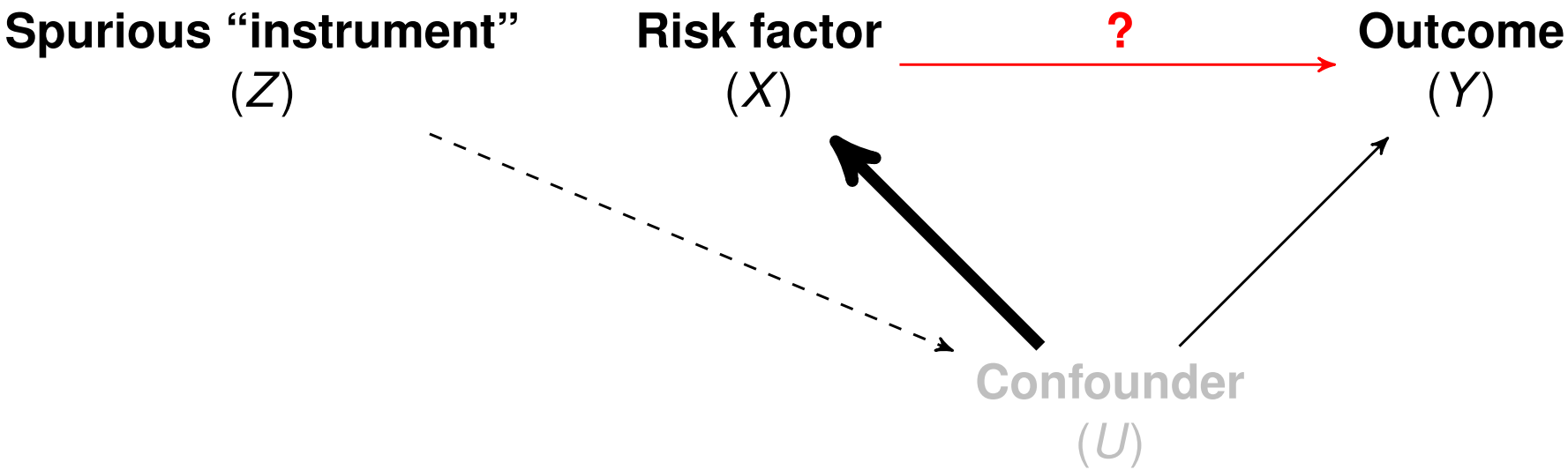
The simplest case (extreme confounding)



Recall that under the IV framework,

$$X \to Y = \frac{Z \to X \to Y}{Z \to X} = \frac{Cov(Z, Y)}{Cov(Z, X)}$$

The simplest case (extreme confounding)



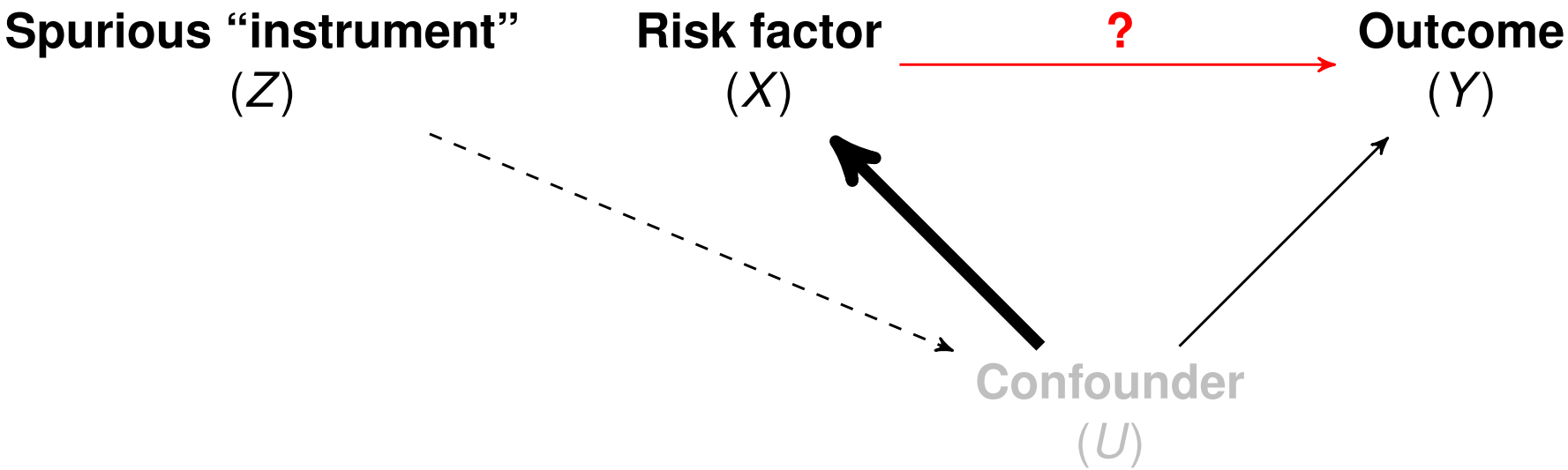
In the case of irrelevant "instrument,"

- $Cov(Z, Y) = Z \to U \to X \to Y \quad + \quad Z \to U \to Y$
- $Cov(Z, X) = Z \to U \to X$

$$\text{So} \quad \frac{Cov(Z, Y)}{Cov(Z, X)} = X \to Y \quad + \quad \frac{U \to Y}{U \to X}$$

The bias is the same for all irrelevant "instruments"!

The simplest case (extreme confounding)



In the case of irrelevant "instrument,"

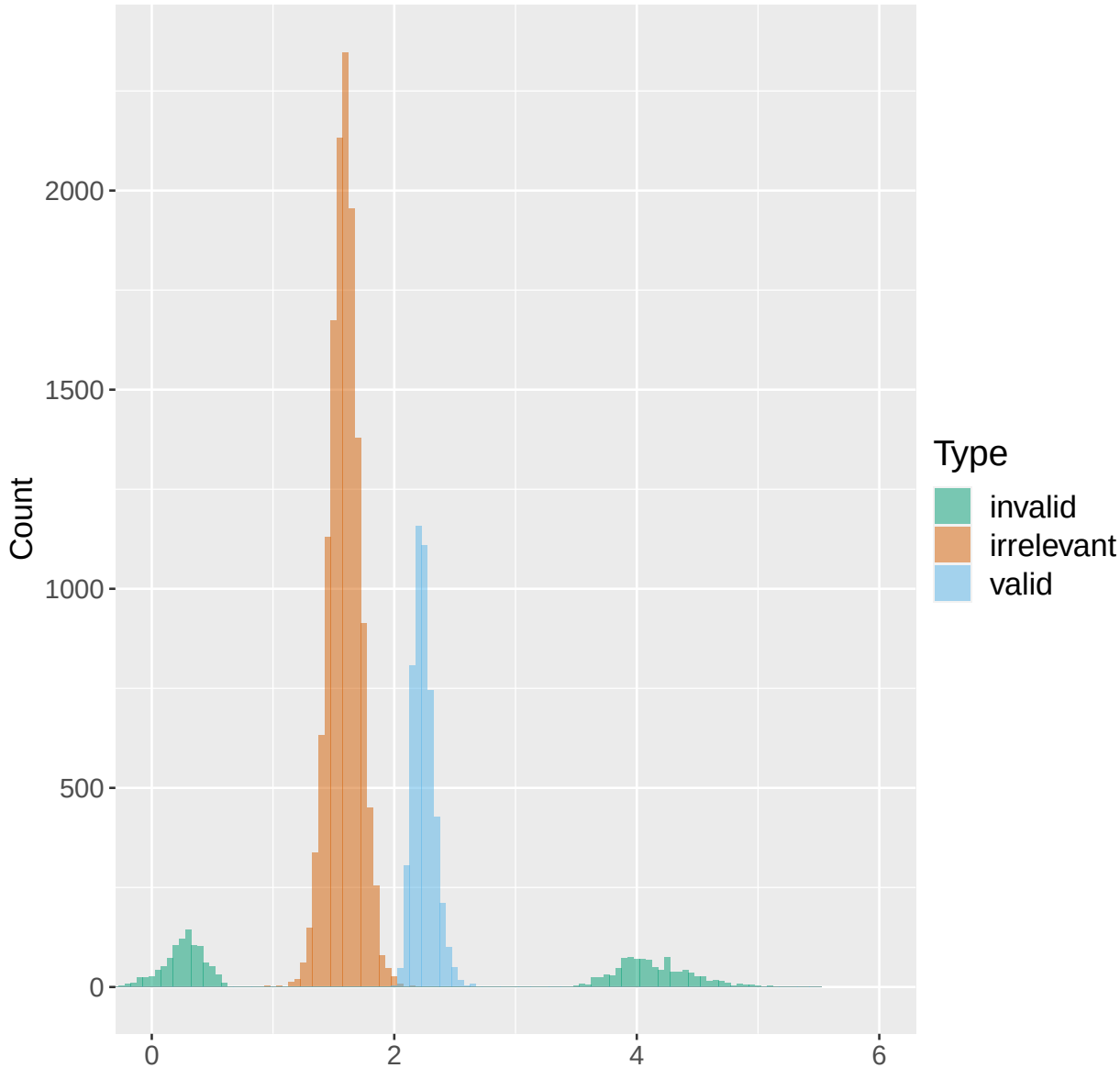
- $Cov(Z, Y) = Z \to U \to X \to Y \quad + \quad Z \to U \to Y$
- $Cov(Z, X) = Z \to U \to X$

$$\text{So} \quad \frac{Cov(Z, Y)}{Cov(Z, X)} = X \to Y \quad + \quad \frac{U \to Y}{U \to X}$$

The bias is the same for all irrelevant "instruments"!

- In fact, also same as the bias from $Y \sim X$

What happened in Step 2



Distribution of causal effect estimates for candidate IVs passing joint thresholding across 1000 Monte Carlo runs

What we know so far:

- **All valid instruments are alike**
- **All irrelevant instruments are alike**
- Invalid instruments are different from each other

What we know so far:

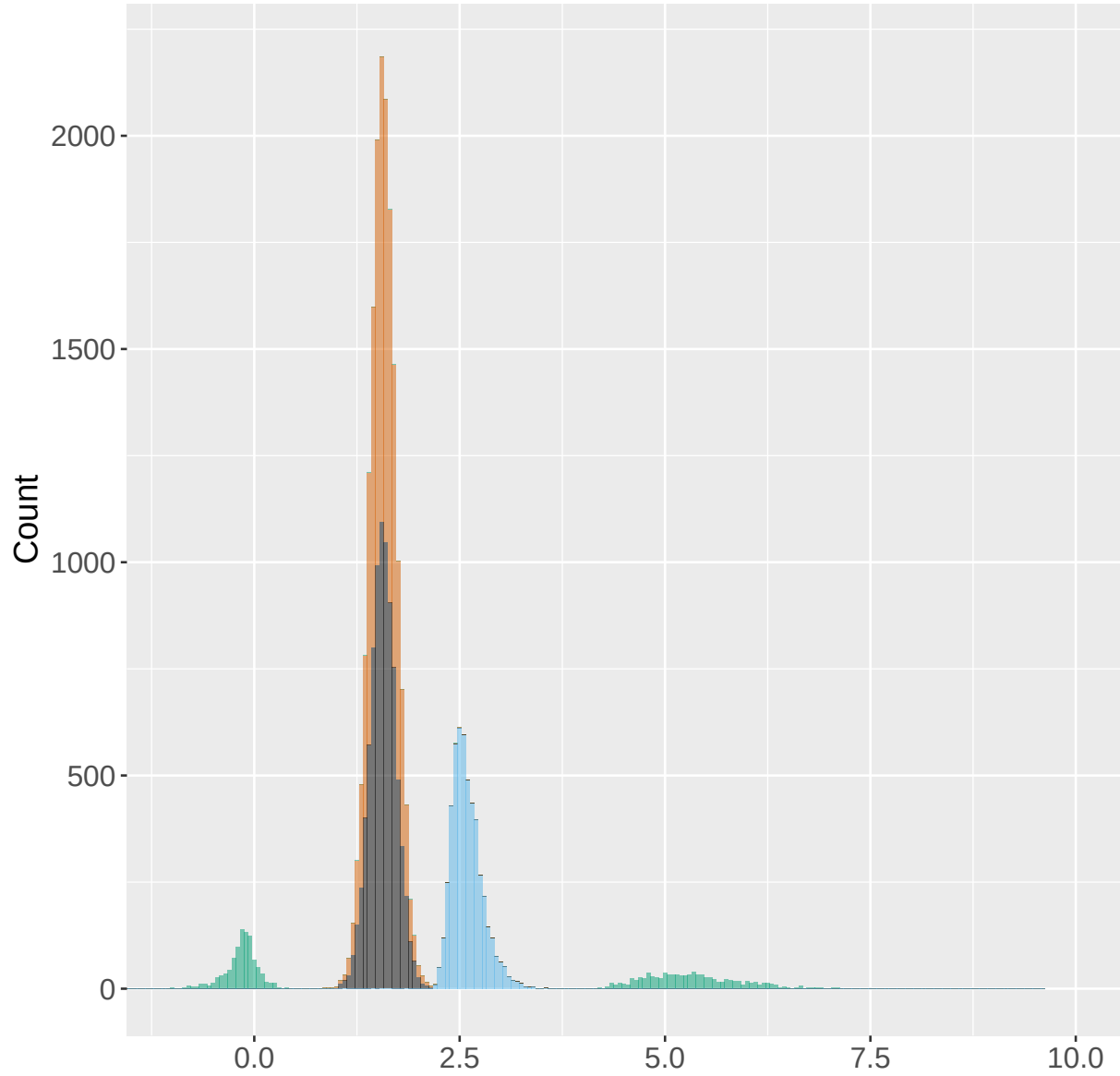
- **All valid instruments are alike**
- **All irrelevant instruments are alike**
- Invalid instruments are different from each other

How to distinguish valid instruments from irrelevant ones?

Key idea: Generate pseudo variables by permuting rows of Z, denoted by Z^*

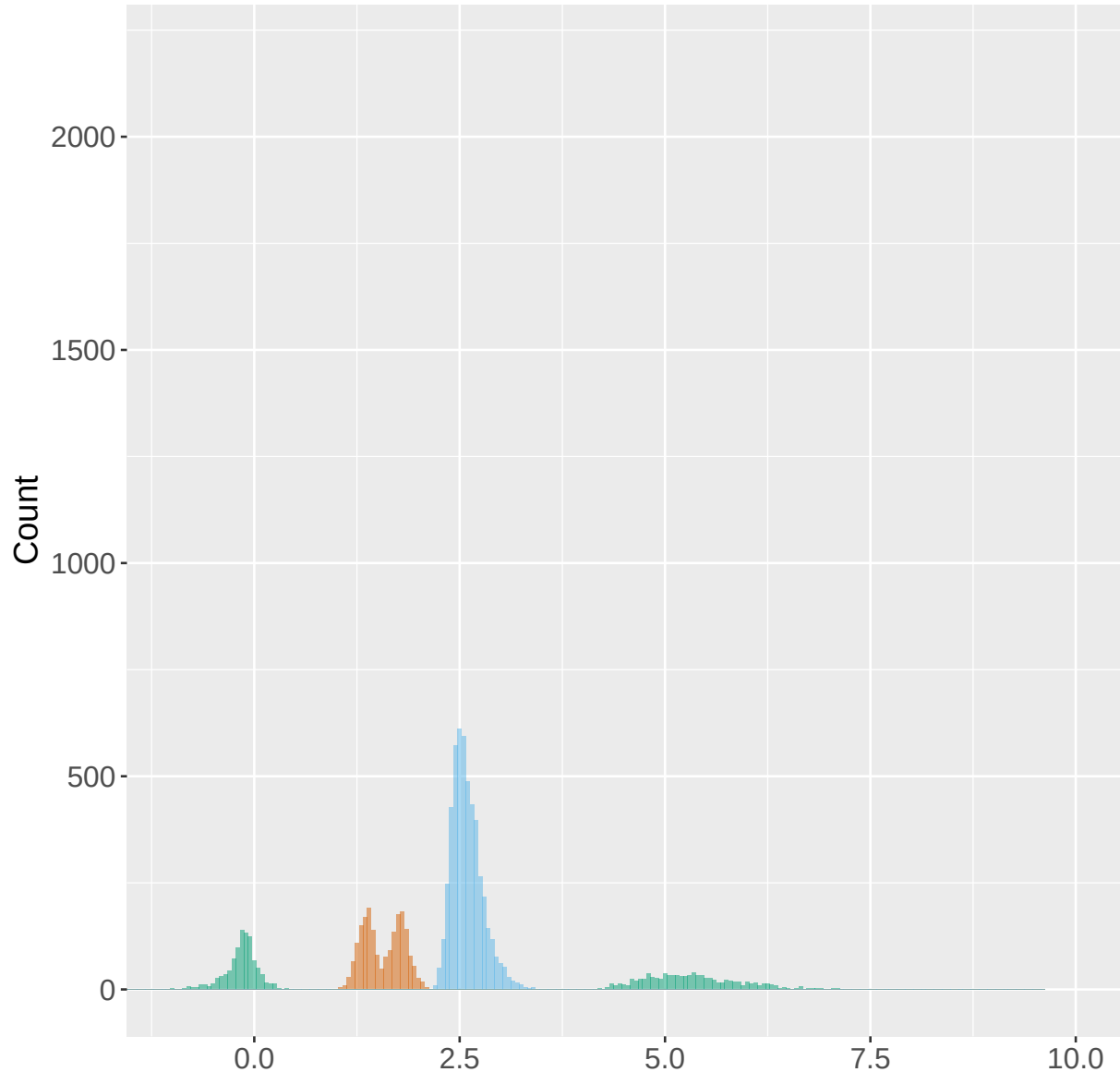
- **Group 1: valid instruments in Z**
- **Group 2: invalid instruments in Z**
- **Group 3: irrelevant instruments in Z**
- **Group 4: irrelevant instruments in Z^***

Groups 3 and 4 are alike! We can track variables in Group 4.



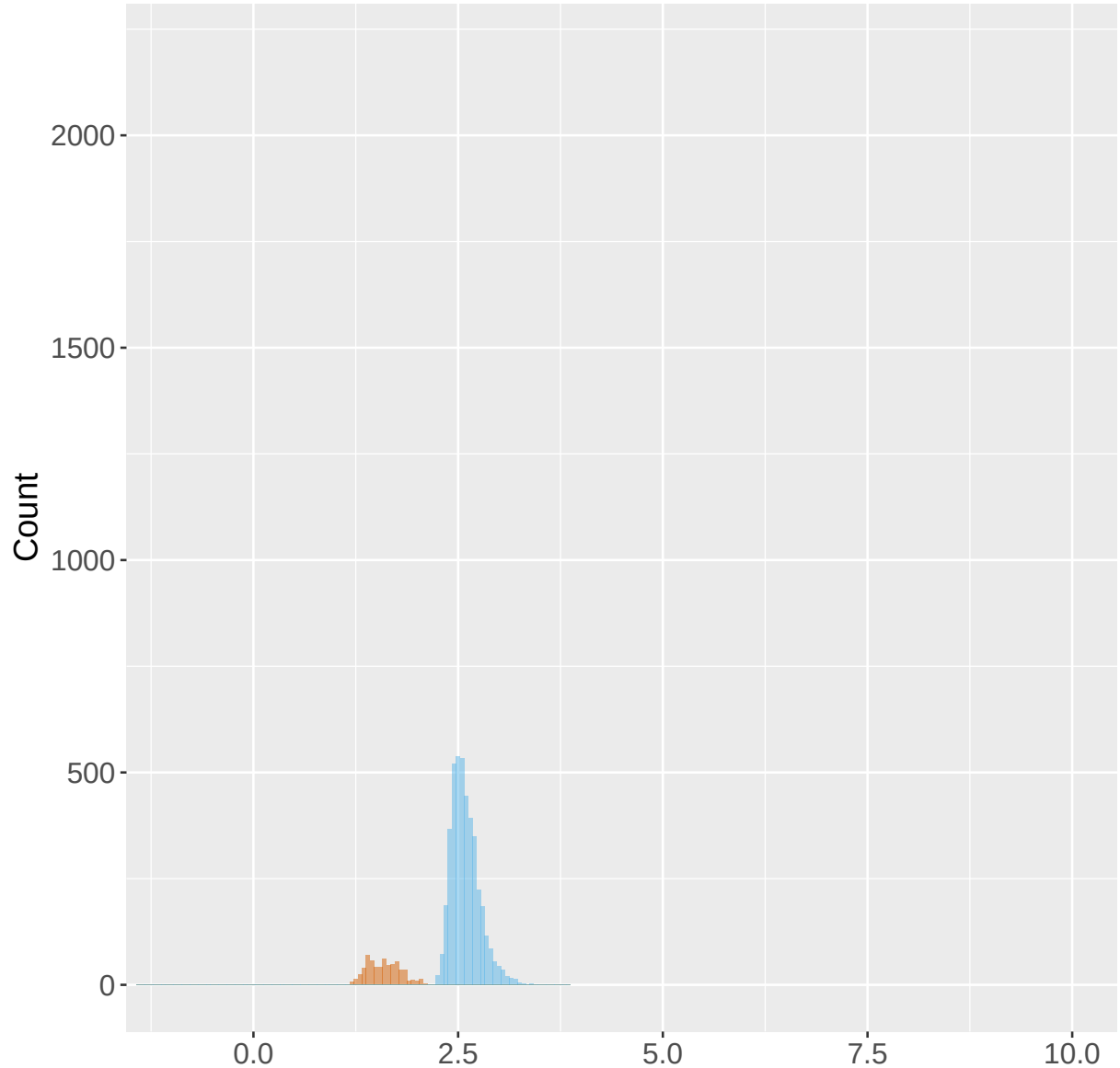
Distribution of causal effect estimates for candidate IVs passing joint thresholding across 1000 Monte Carlo runs

Step 2: Remove causal effect estimates inside the range of pseudo effect estimates



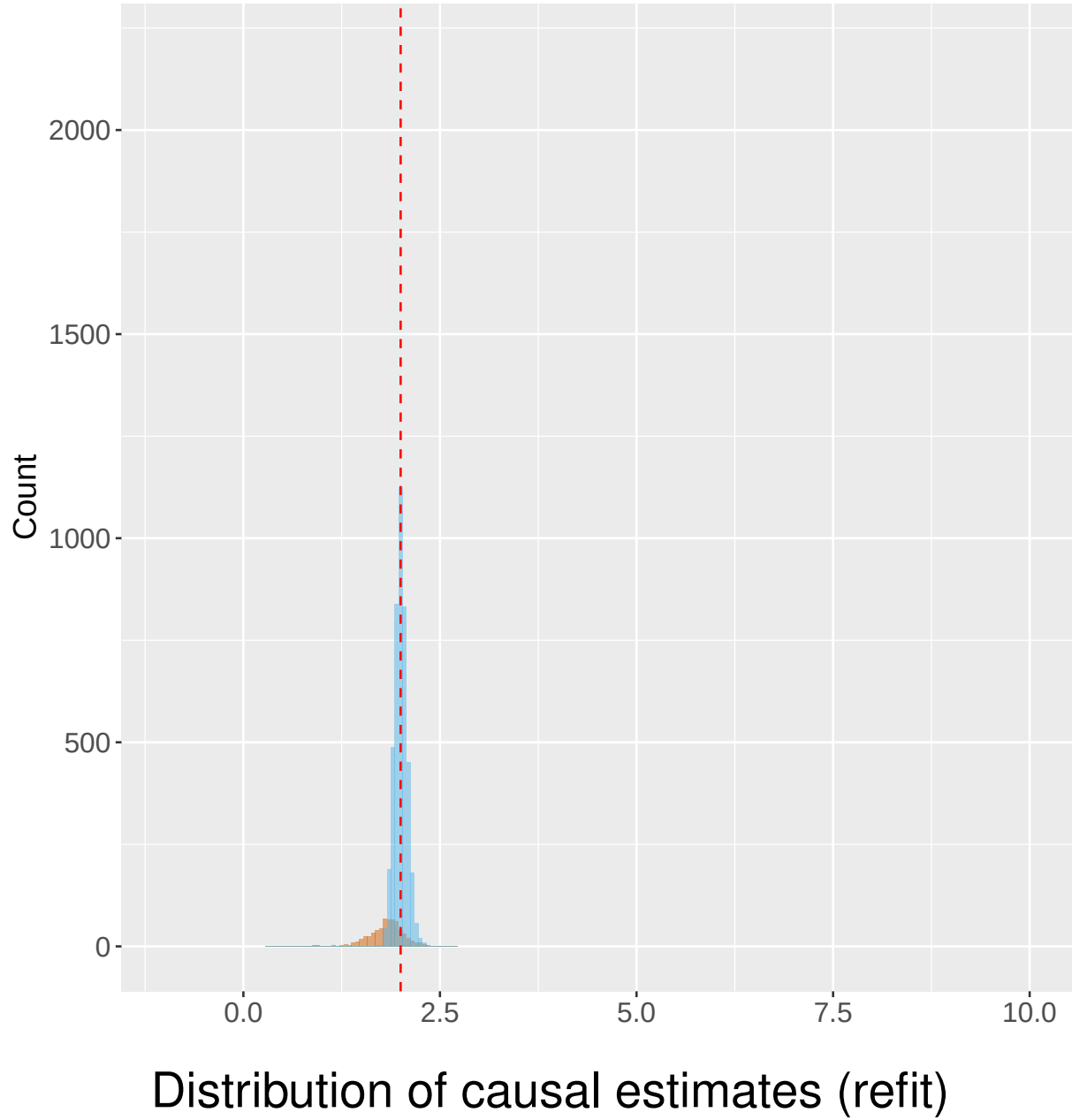
Distribution of causal estimates after removing spurious IVs

Step 3: Develop a mode-finding algorithms to identify the valid instruments



Distribution of causal estimates after mode-finding

Step 4: Use the selected valid instruments to estimate the causal effects



Theoretical guarantees

Theorem 1 (Zhang et al., 2022)

(Spurious IV) Under some regularity conditions, with probability tending to 1, causal estimates for spurious IVs are concentrated within

$$[\beta^* + C_* - d,\ \beta^* + C_* + d]$$

for $d > L\left(\omega, \sigma_X^2, \sigma_u^2, \sigma_Y^2, \alpha_X^*, \alpha_Y^*\right)$, where $C_* = (\alpha_Y^* \alpha_X^* \sigma_u^2)/(\alpha_X^{*2} \sigma_u^2 + \sigma_X^2)$.

Theorem 2 (Zhang et al., 2022)

(Valid IV) Under some regularity conditions, with probability tending to 1, causal estimates for valid IVs are concentrated around β^*.

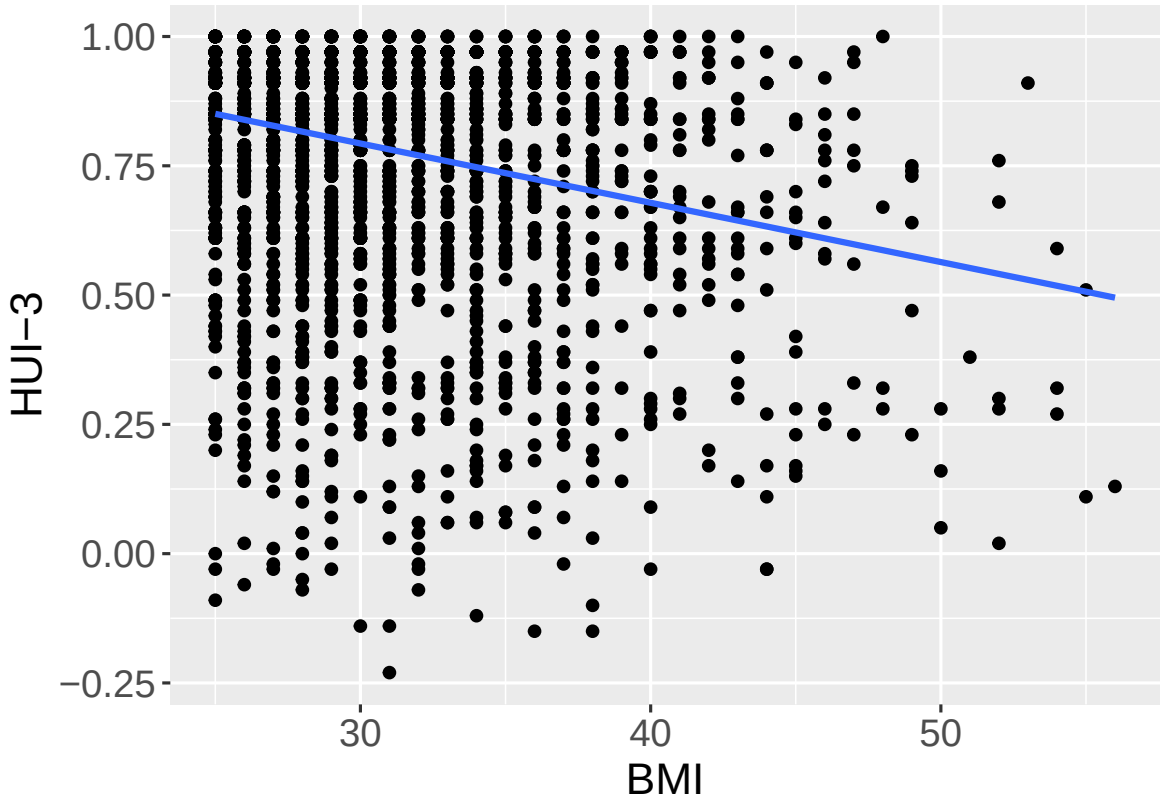
Separation: If $d < |C_*|$, then spurious IVs and valid IVs are separable.

Easier separation if stronger U OR sufficiently large ω

Wisconsin Longitudinal Study

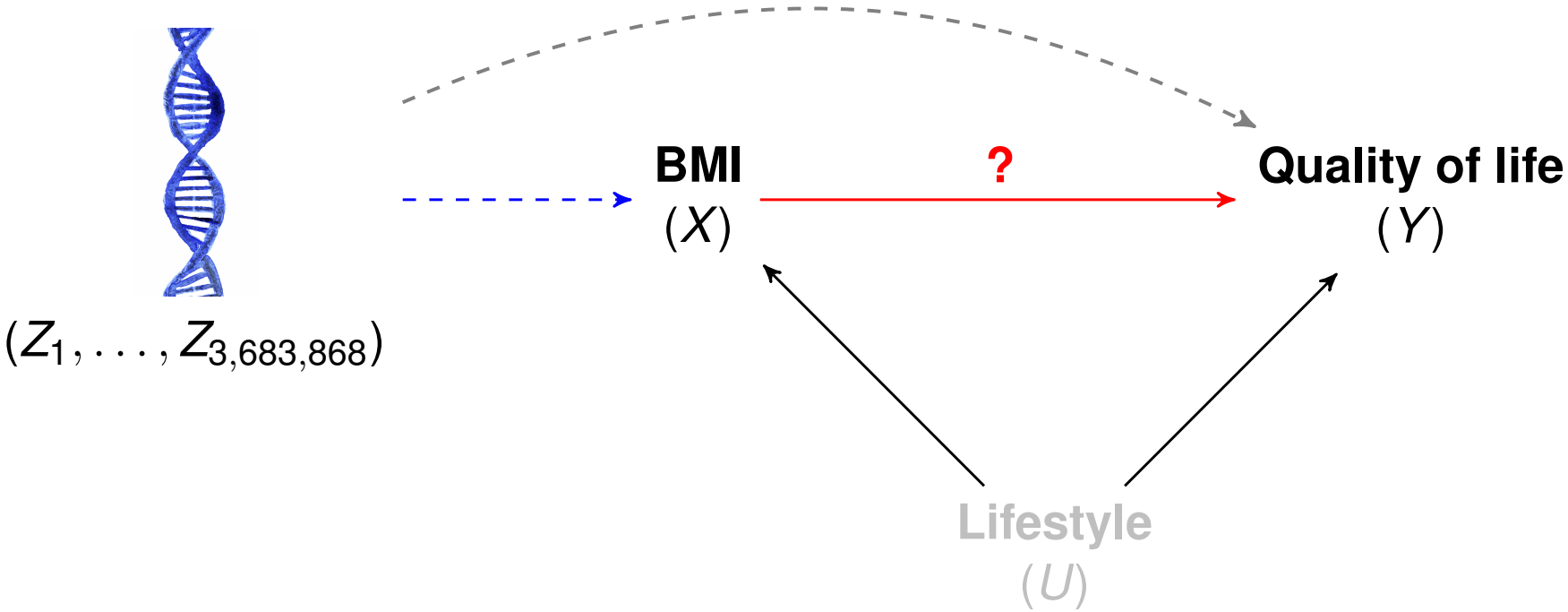
- Participants were graduates from Wisconsin high schools in 1957 and their siblings
- Use 3023 unrelated individuals re-interviewed in 2011
- Exposure: BMI (only those ≥ 25)
  - mean (SD): 30.6 (4.9)
- Outcome: Health Utility Index Mark 3 (HUI-3)
  - between -0.22 and 1
  - mean (SD): 0.79 (0.23)

Crude analysis



After adjusting for age, gender and education, one unit BMI increase associated with 0.011 unit decrease in HUI-3 (95% CI = [0.009, 0.013])

Conceptual Mendelian randomization model

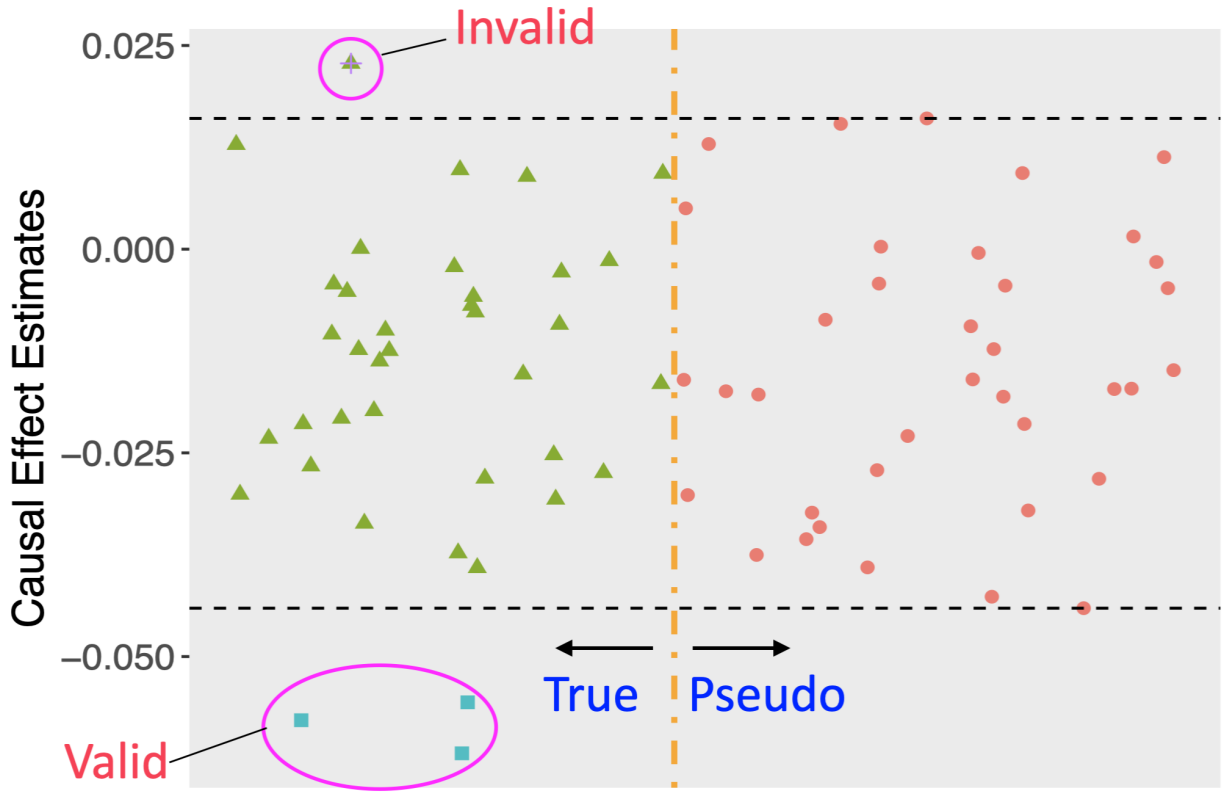


- Consider baseline covariates including age, gender, education and population stratification (top 6 principal components)

Data analysis: Prepossessing + Step 1

- Quality control ☞ 3,683,868 SNPs left
- Generate the same number of pseudo SNPs by permuting the rows of the Z matrix
- After GWAS: 44 SNPs $+$ 42 pseudo

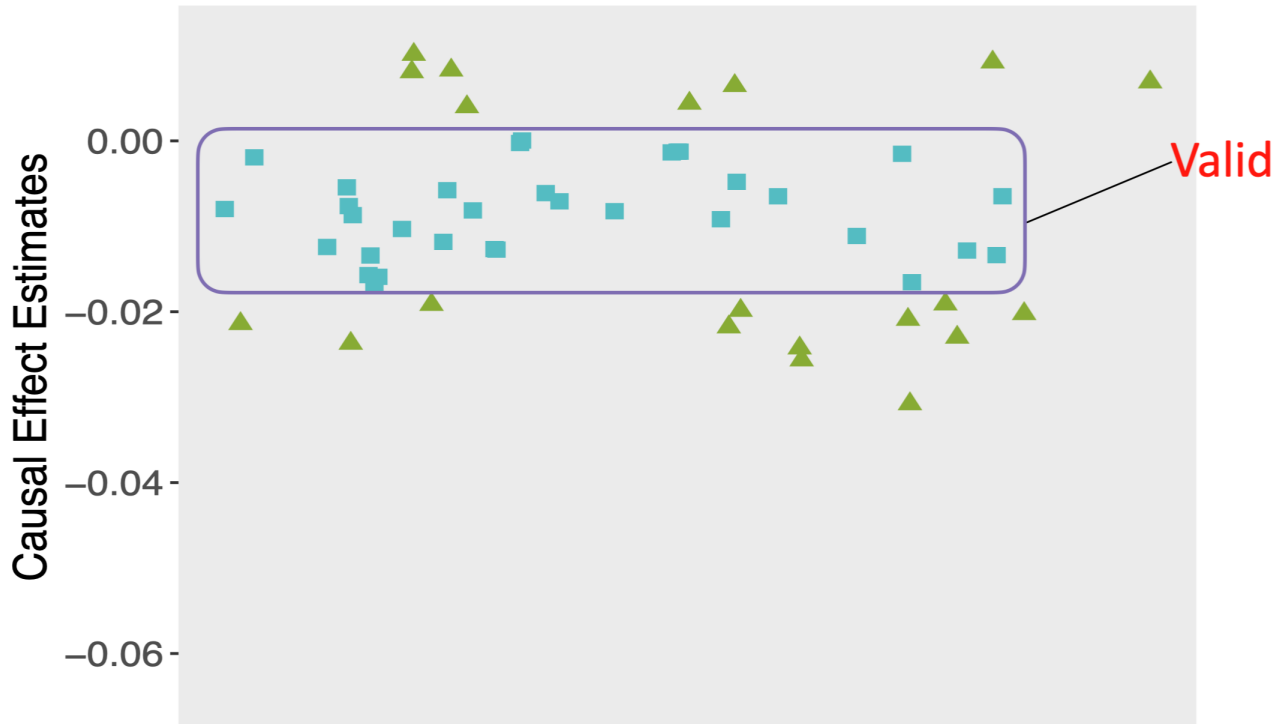
Data analysis: Steps 2-4



Causal estimate: one unit increase in BMI will result in 0.039 unit decrease of HUI-3 (95% CI $= [-0.052, -0.025]$)

Data analysis: Ignoring spurious instruments

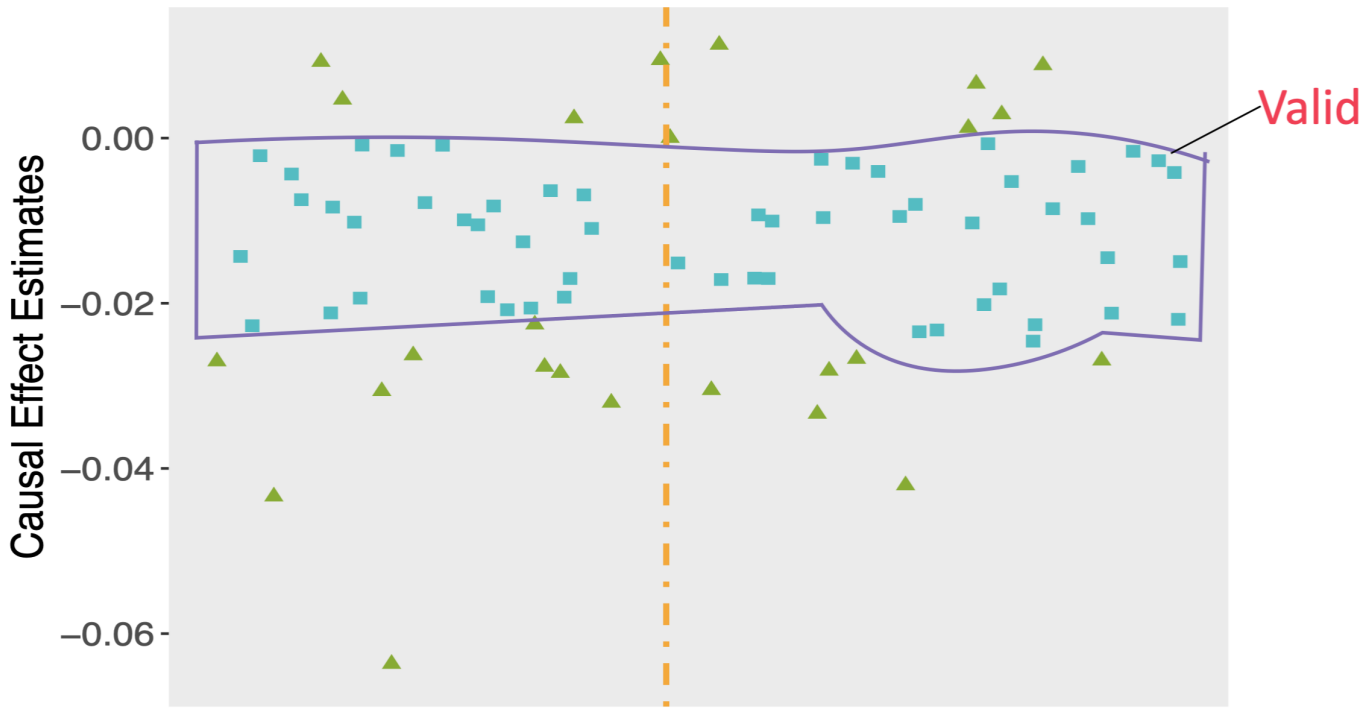
Two-stage hard thresholding with voting (Guo et al., 2018): true SNPs only



Causal estimate: one unit increase in BMI will lead to 0.008 unit decrease in HUI-3 (95% CI $= [-0.013, -0.003]$)

Data analysis: Ignoring spurious instruments

Two-stage hard thresholding with voting (Guo et al., 2018): same data set as the proposed (true + pseudo)



Causal estimate: one unit increase in BMI will lead to 0.011 unit decrease in HUI-3 (95% CI $= [-0.015, -0.008]$)

Comparison of results

Proposed:

1 SD increase in BMI leads to 1 SD (roughly) decrease in HUI-3

Guo et al. (2018)'s method (ignore spurious instruments; true only):

1 SD increase in BMI leads to 0.2 SD (roughly) decrease in HUI-3

Guo et al. (2018)'s method (ignore spurious instruments; true + pseudo):

1 SD increase in BMI leads to 0.2 SD (roughly) decrease in HUI-3

Crude analysis (OLS):

1 SD increase in BMI is associated with 0.2 SD (roughly) decrease in HUI-3

Summary

- Mendelian randomization is a powerful tool for causal effect estimation
- Challenges for MR studies
  - Find relevant IVs ☞ GWAS
  - Deal with pleiotropy ☞ find the mode of effect estimates
  - **Bias due to spurious variables in GWAS**
- Spurious IV bias is potentially a serious problem
  - Ignoring Spurious IV bias has a similar effect as ignoring unmeasured confounding bias
- **Fight noise with noise**
  - Use pseudo variables to correct for Spurious IV bias

References I

Guo, Z., Kang, H., Cai, T. T. & Small D. S. (2018). Confidence intervals for causal effects with invalid instruments by using two-stage hard thresholding with voting. J. R. Statist. Soc. B. 80, 793–815.

Kolesar, M., Chetty, R., Friedman, J., Glasser, E. & Imbens, G. W. (2015). Identification and inference with many invalid instruments. J. Bus. Econ. Statist. 33, 474–484.

Kang, H., Zhang, A., Cai, T. T. & Small , D. S. (2016). Instrumental variables estimation with some invalid instruments and its application to Mendelian randomization. J. Am. Statist. Ass. 111, 132–144.

Zhang, X., Wang, L., Volgushev, S., & Kong, D. (2022). Fighting Noise with Noise: Causal Inference with Many Candidate Instruments. arXiv preprint arXiv:2203.09330.