Causal Diagrams for Measurement Error in Epidemiologic Research

Onyebuchi A. Arah, MD, MPH, PhD

Department of Social Medicine
Academic Medical Center, University of Amsterdam
The Netherlands
&
Department of Epidemiology
UCLA School of Public Health
Los Angeles, CA, United States

Society for Epidemiologic Research
Chicago, June 2008

BACKGROUND

Observational studies using non-experimental data often plagued by multiple biases such as measurement error, response/selection bias, and unmeasured confounders.

Yet, these biases are rarely quantitatively assessed in studies or even acknowledged.

Several techniques have been developed to handle bias from measurement error and uncontrolled confounding but are rarely applied (← complexity?)

OBJECTIVE & APPROACH

Drawing on the growing literature on directed acyclic graphs (DAGs) for depicting causal assumptions, this study shows how DAGs can be augmented with information on non-differential measurement error in exposure and confounders.

The augmented graphs can be queried visually and directly to reveal the directed paths through which target effect parameters are biased by measurement error in any variable.

OUTLINE

Scenarios covered:

- exposure and confounder ME
- correlated confounders
- uncontrolled confounding
- adjustment for unnecessary surrogates

Knowledge of DAGs (naïvely) assumed

All perfectly measured variables/parameters asterisked (*)

D^*: outcome

X^*: true exposure (continuous variable)

X: mismeasured X^* or surrogate where $X = X^* + \varepsilon_X$

Z^*: true confounder

Z: surrogate for true confounder Z^* where $Z = Z^* + \varepsilon_Z$

Non-differential classical ME: $V = V^* + \varepsilon_V$, where ε_V is *iid* with mean expectation 0 and variance σ^2 and uncorrelated with other error terms

Surrogacy: $V \perp\!\!\!\perp D^* \mid V^*$, that is, the surrogate provides no fresh information once the true variable is known

Attenuation factor due to surrogate V for true effect of V^*: λ_V also given by intraclass correlation coefficient

$$\lambda_V = \sigma^2_{V^*} / (\sigma^2_{V^*} + \sigma^2_{\varepsilon_V})$$

Augmented DAGs
with Measurement Error in the Exposure

Draw observed model, that is $M_{observed}$ using dashed arrows:

(a) $X \xrightarrow{\beta_X} D^*$

Draw true model, that is M_{true} using solid arrows only:

(b) $X^* \xrightarrow{\beta_{X^*}} D^*$

Expand true model to include surrogate variable:

(c) $X \xleftarrow{\lambda_X} X^* \xrightarrow{\beta_{X^*}} D^*$

$M_{observed}$ in (a) is equivalent to the expanded M_{true} in (c)

(d)

$$\left.\begin{array}{c} X \xleftarrow{\lambda_X} X^* \xrightarrow{\beta_{X^*}} D^* \\ \Updownarrow \\ X \xrightarrow{\beta_X} D^* \end{array}\right\}$$

X **acts as an "instrument" for the true** ***X*→D**** **relationship**

Draw surrogate model, that is $M_{observed}$ using dashed arrows:

(a)

$$\begin{tikzcd} X \arrow[r, dashed, "\beta_X"'] & D^* \end{tikzcd}$$

Draw true model, that is M_{true} using solid arrows only:

(b)

$$\begin{tikzcd} X^* \arrow[r, "\beta_{X^*}"'] & D^* \end{tikzcd}$$

Expand true model to include surrogate variable:

(c)

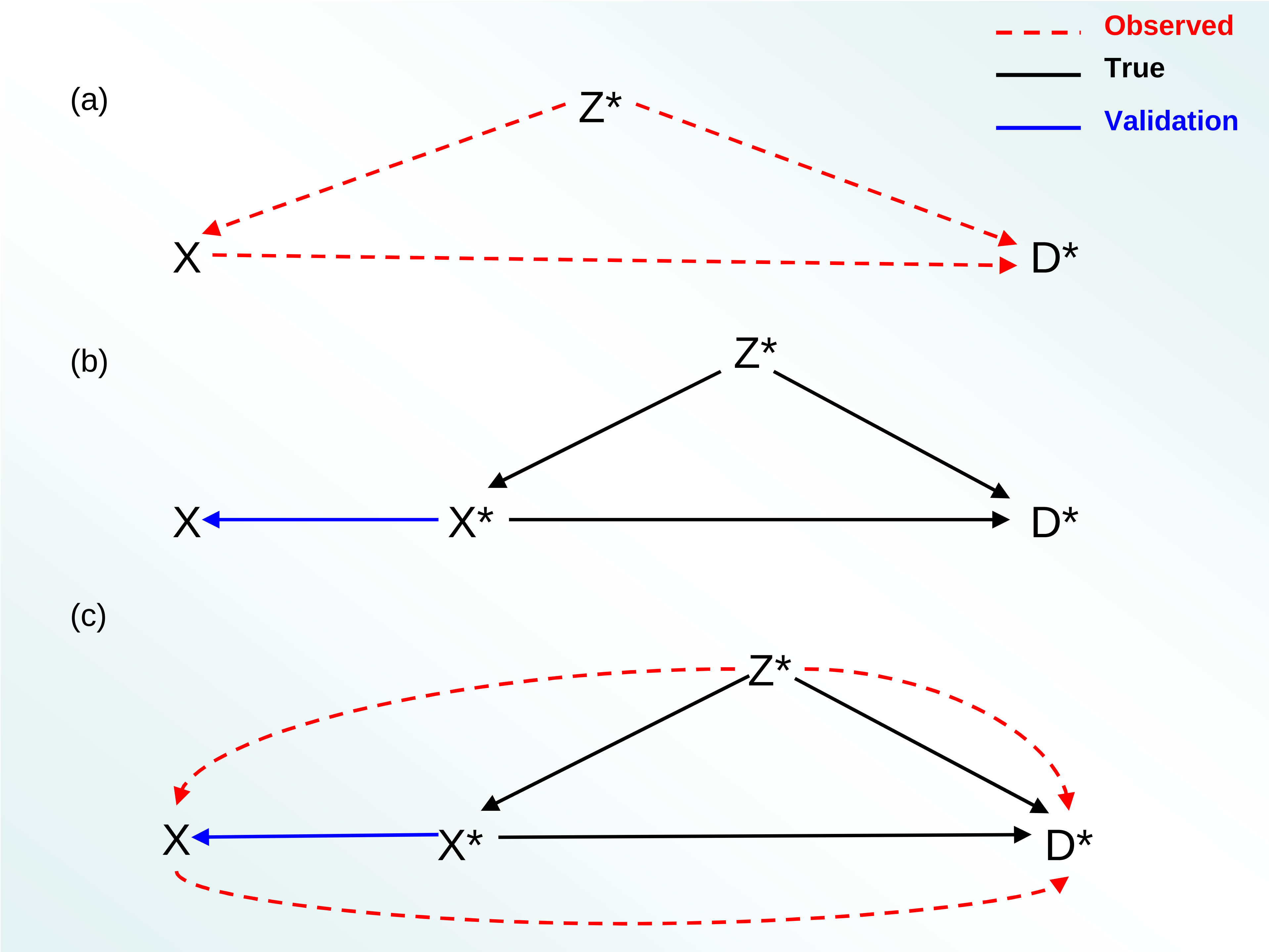
$$\begin{tikzcd} X & X^* \arrow[l, "\lambda_X"] \arrow[r, "\beta_{X^*}"'] & D^* \end{tikzcd}$$

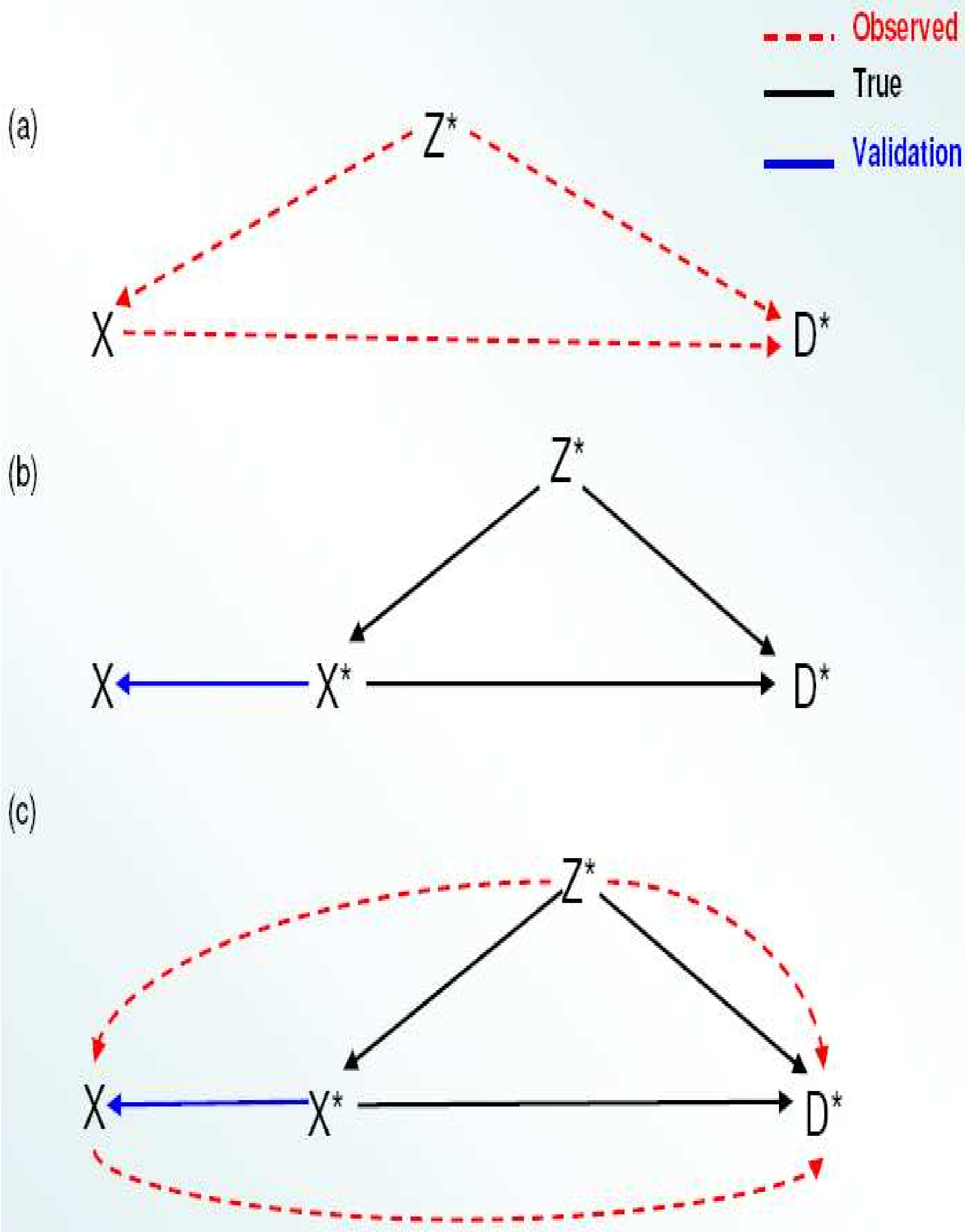
$M_{observed}$ in (a) is equivalent to the expanded M_{true} in (c)

(d)

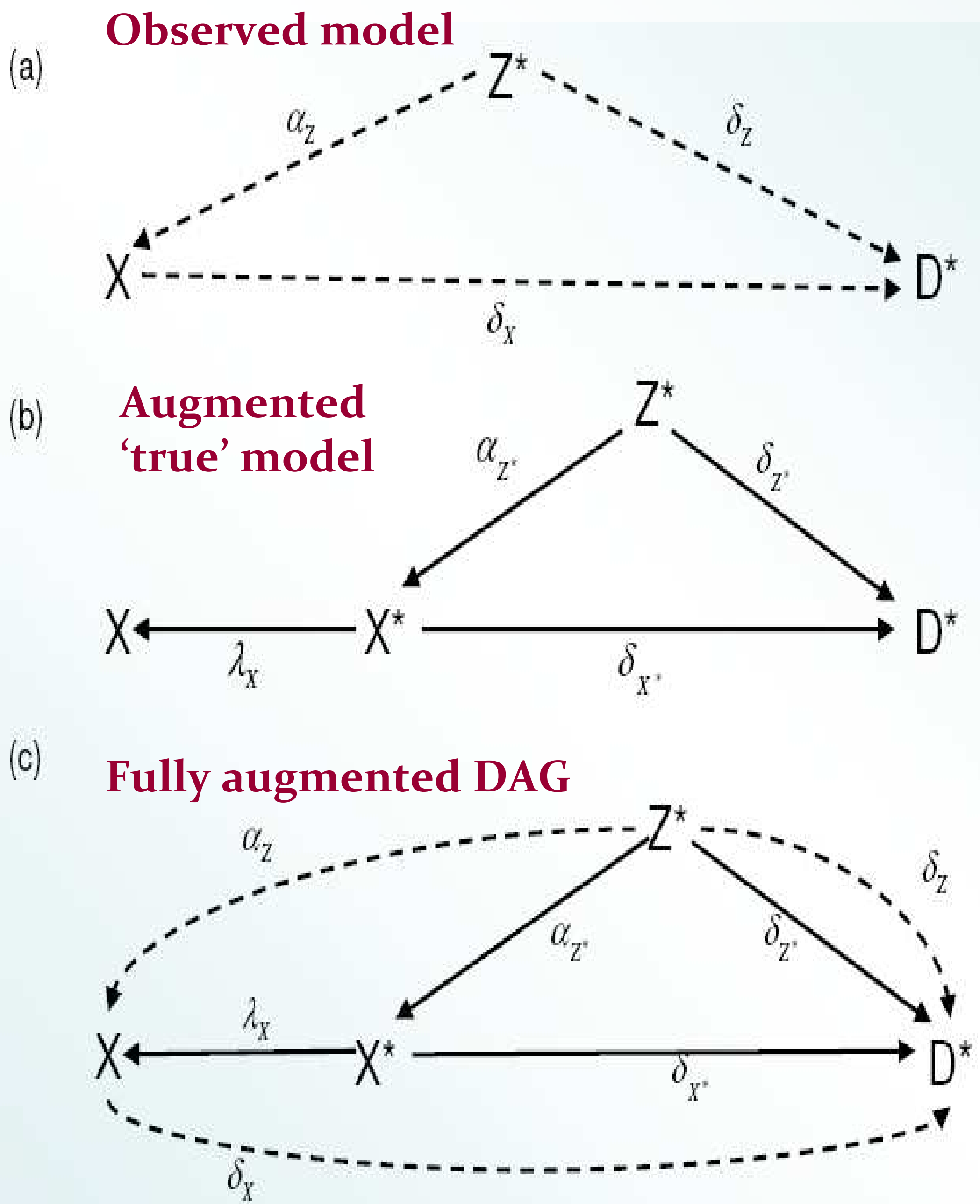
$$\begin{tikzcd} X \arrow[rr, dashed, bend right, "\beta_X"'] & X^* \arrow[l, "\lambda_X"] \arrow[r, "\beta_{X^*}"'] & D^* \end{tikzcd}$$

***X* acts as an "instrument" for the true *X*→D** relationship**


Observed
True
Validation
(a)
Z*
X
D*
(b)
Z*
X
X*
D*
(c)
Z*
X
X*
D*



- Each directed path (*broken arrow*) between any two adjacent nodes in the observed model is equivalent to all the *solid arrow paths* in the 'true' model connecting those two nodes
- For example, the observed $X \rightarrow D^*$ path is equivalent to the augmented true path $X \rightarrow X^* \rightarrow D^*$
- Observed effect of Z^* on D^* given X in (c): $Z^* \rightarrow D^* + Z^* \rightarrow X^* \rightarrow D^*$
- Support for regression calibration



- Since $X \rightarrow D^*$ implies $X \rightarrow X^* \rightarrow D^*$, then

$$\delta_{X^*} = \delta_X / \lambda_X$$

- Similarly for the observed effect of Z^* on D^* given surrogate X in (c):

$Z^* \rightarrow D^* + Z^* \rightarrow X^* \rightarrow D^*$:

$$\delta_Z = \alpha_{Z^*}\delta_{X^*} + \delta_{Z^*}$$

$$\delta_{Z^*} = \delta_Z - \alpha_{Z^*}\delta_{X^*} = \delta_Z - \alpha_{Z^*}\delta_X / \lambda_X$$

- Just as in regression calibration

Regression calibration approach

- True unobserved model:

$$g[\Pr(D^* \mid X^*, Z^*)] = \alpha^* + \beta_1^* X^* + \beta_2^* Z^*$$

- Validation sub-study model with X^*, X and Z^* (but not D^*):

$$g[\Pr(X^* \mid X, Z^*)] = \mathrm{k} + \lambda_X X + \gamma Z^*$$

- Fitted/observed model:

$$g[\Pr(D^* \mid X, Z^*)] = \mathrm{c} + \delta_1 X + \delta_2 Z^*$$

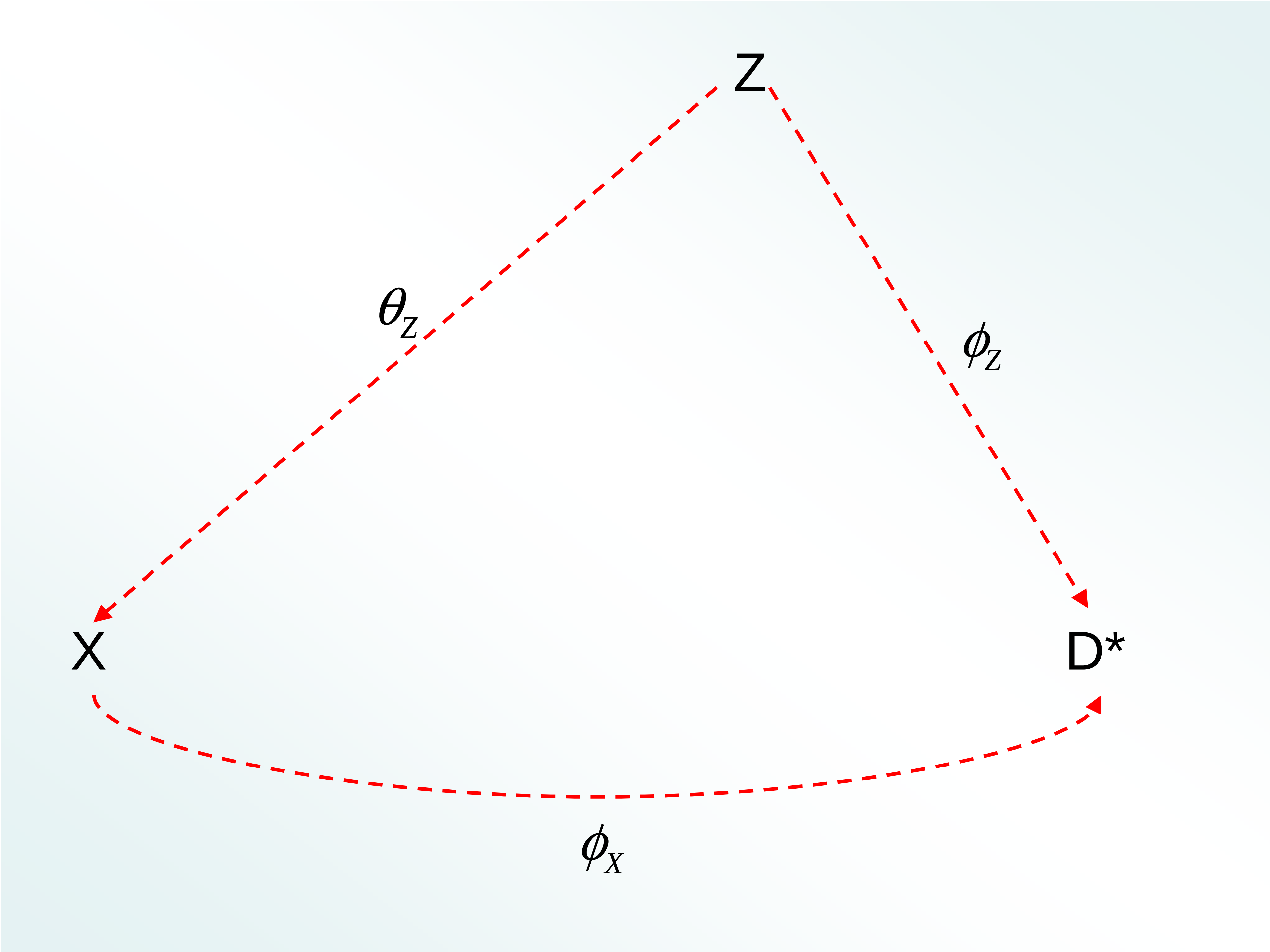
$$g[\Pr(D^* \mid X, Z^*)] = \mathrm{c} + \lambda_X \beta_1^* X + (\beta_2^* + \gamma \beta_1^*) Z^*$$

Hence, true effect of X^*: $\beta_1^* = \delta_1 / \lambda_X$

→ True exposure effect attenuated by λ_X (thus, attenuation factor)

True effect of covariate Z*: $\beta_2^* = \delta_2 - \gamma\beta_1^* = \delta_2 - \gamma\delta_1/\lambda_X$

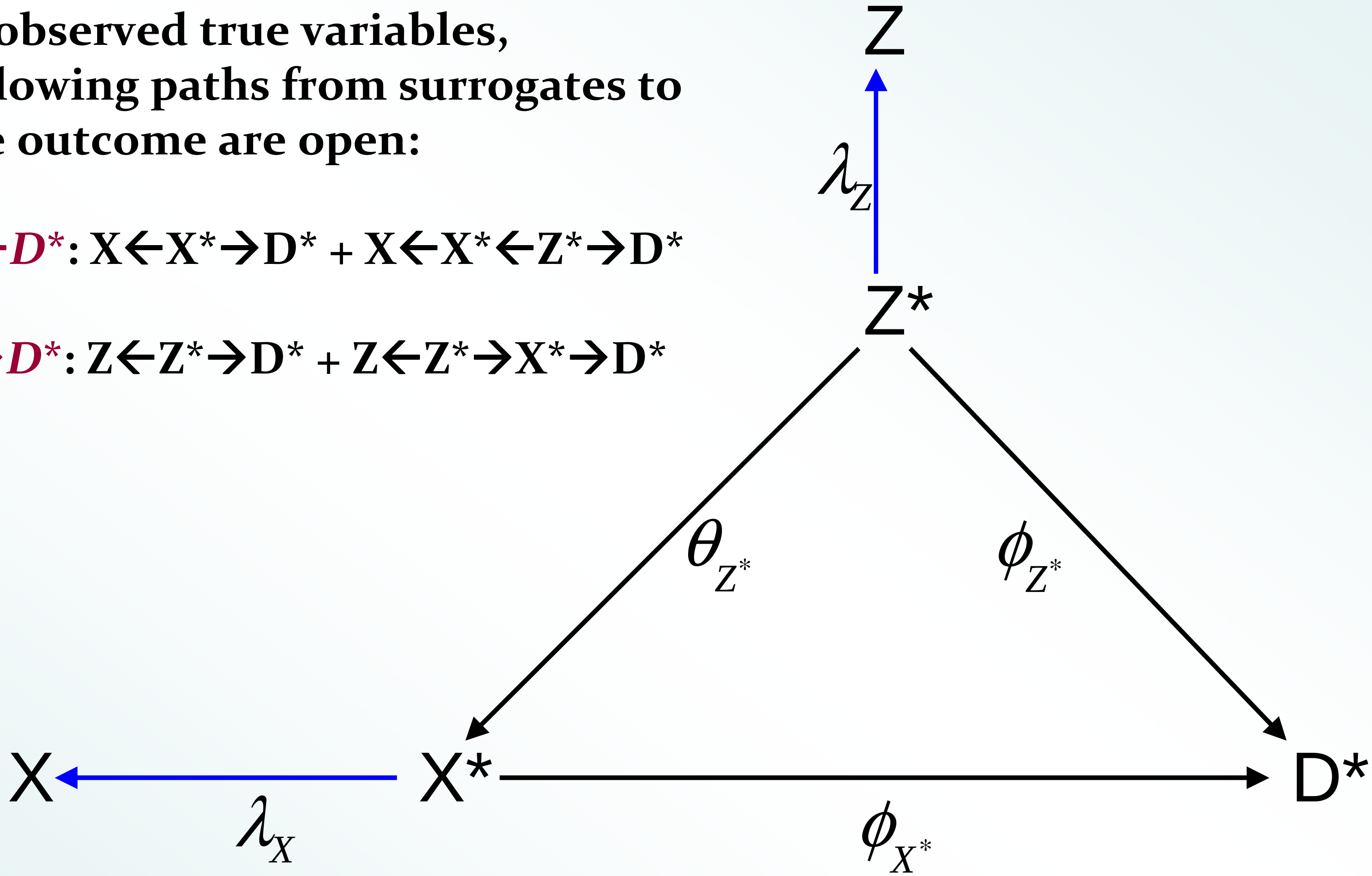
Confounder Measurement Error


Z
θ_Z
ϕ_Z
X
D*
ϕ_X

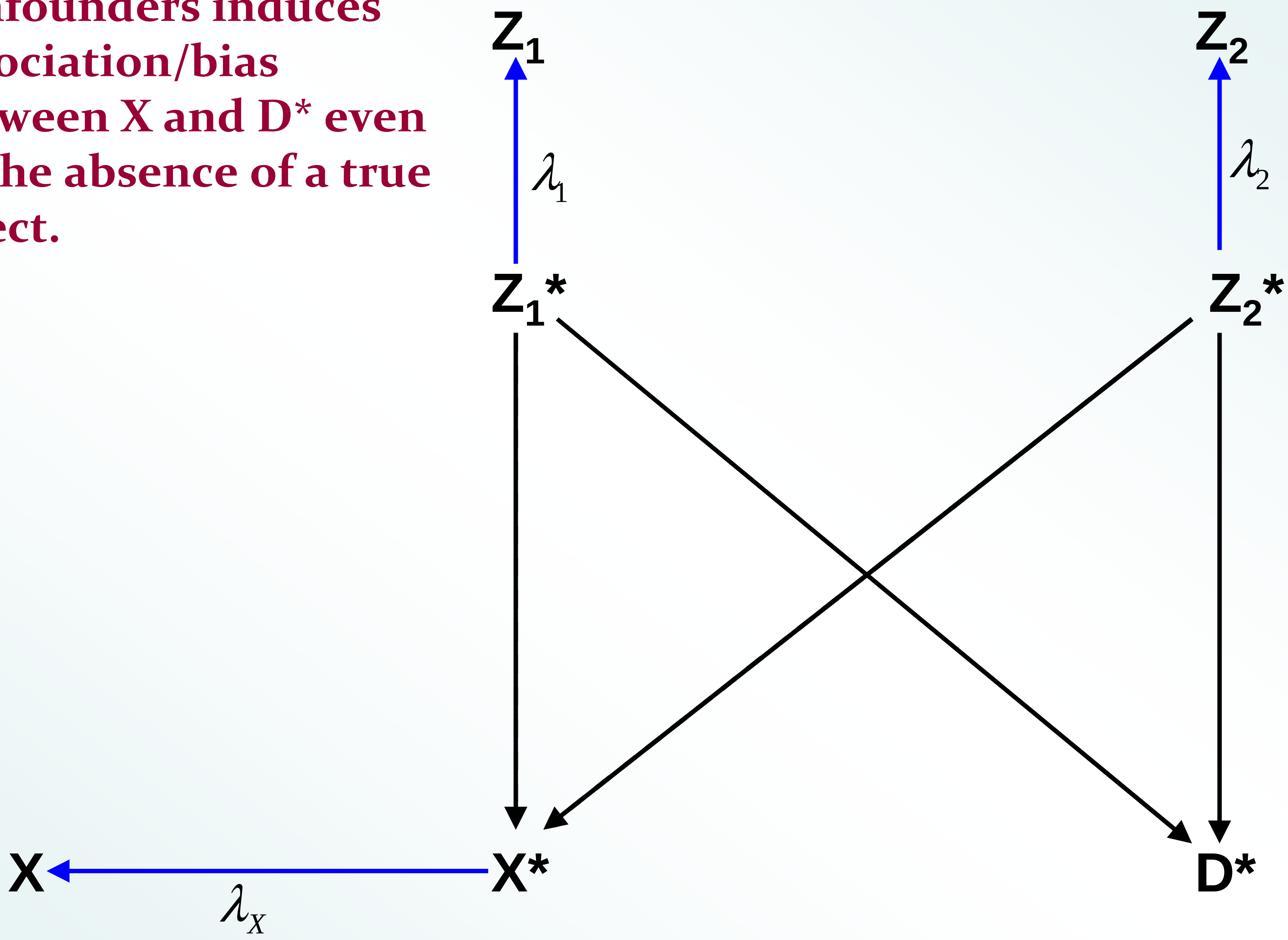
Since both *X and *Z** are unobserved true variables, following paths from surrogates to the outcome are open:**

X←D*: **X←X*→D* + X←X*←Z*→D***

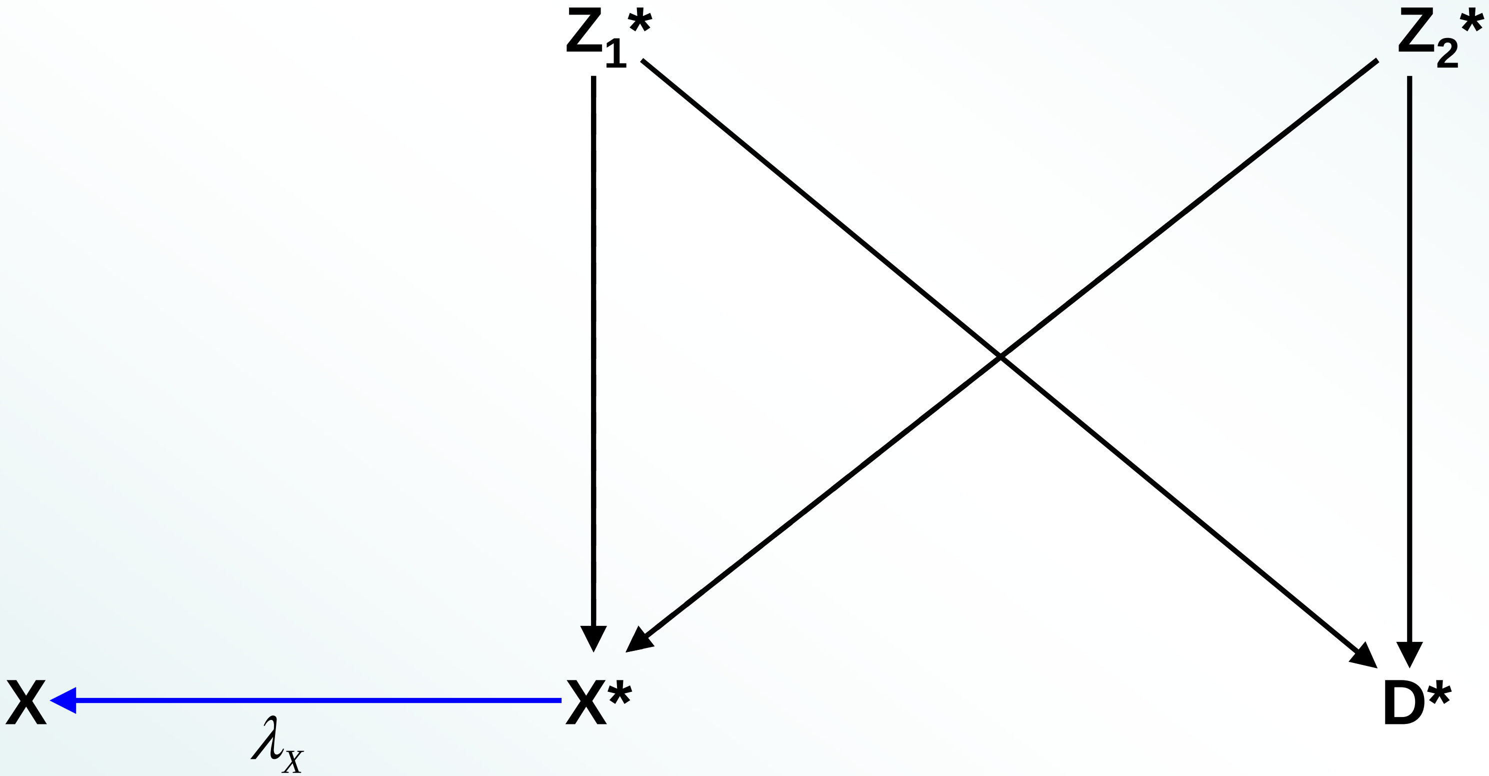
Z→D*: **Z←Z*→D* + Z←Z*→X*→D***



ME in exposure and confounders induces association/bias between X and D* even in the absence of a true effect.

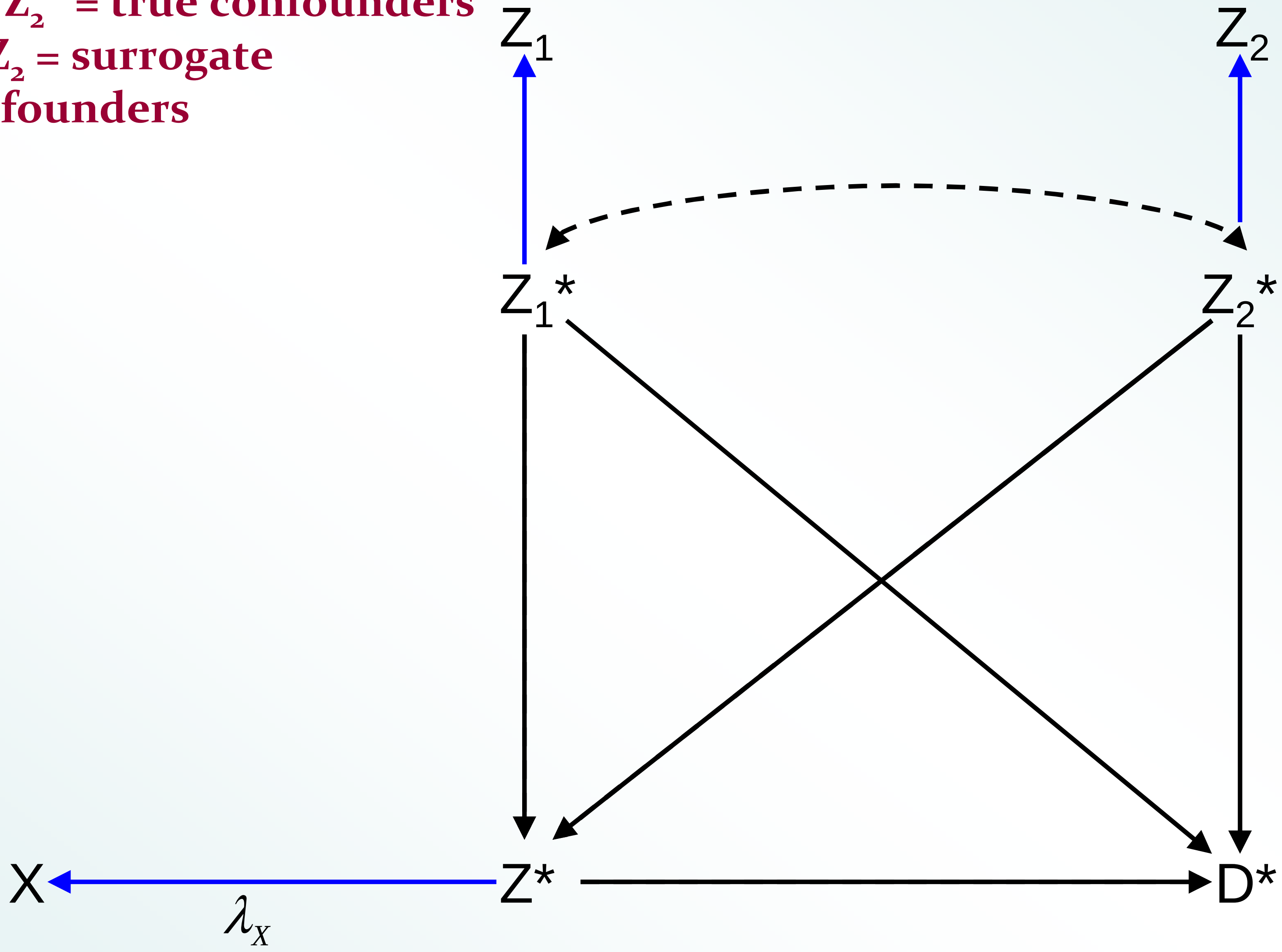


ME in exposure only DOES NOT induce association/bias between X and D* in the absence of a true effect.



Correlated Confounders with Measurement Error

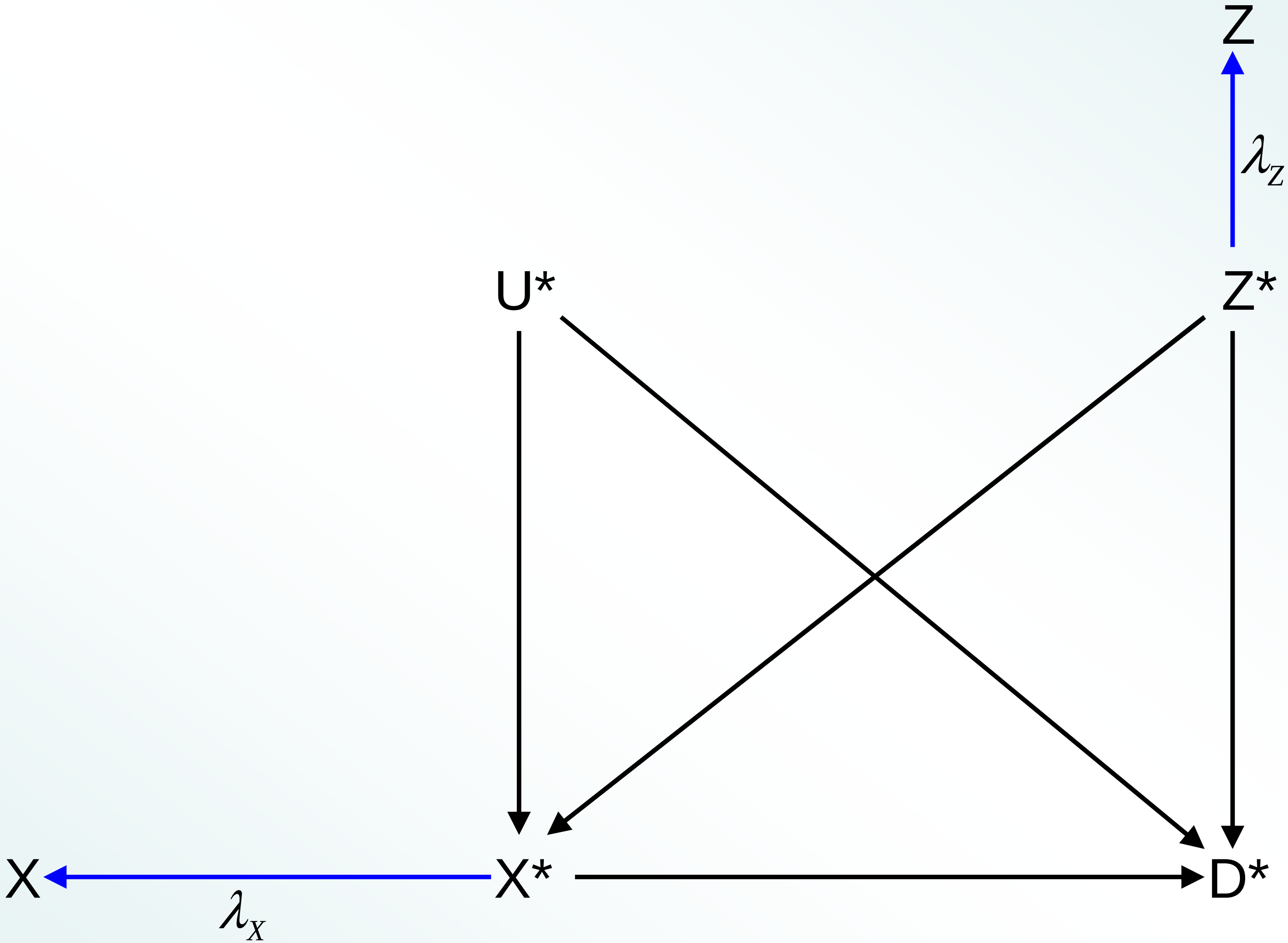
X^*/X = exposure/surrogate
Z_1^*, Z_2^* = true confounders
Z_1, Z_2 = surrogate confounders



Observed $X \rightarrow D^*$ relation:

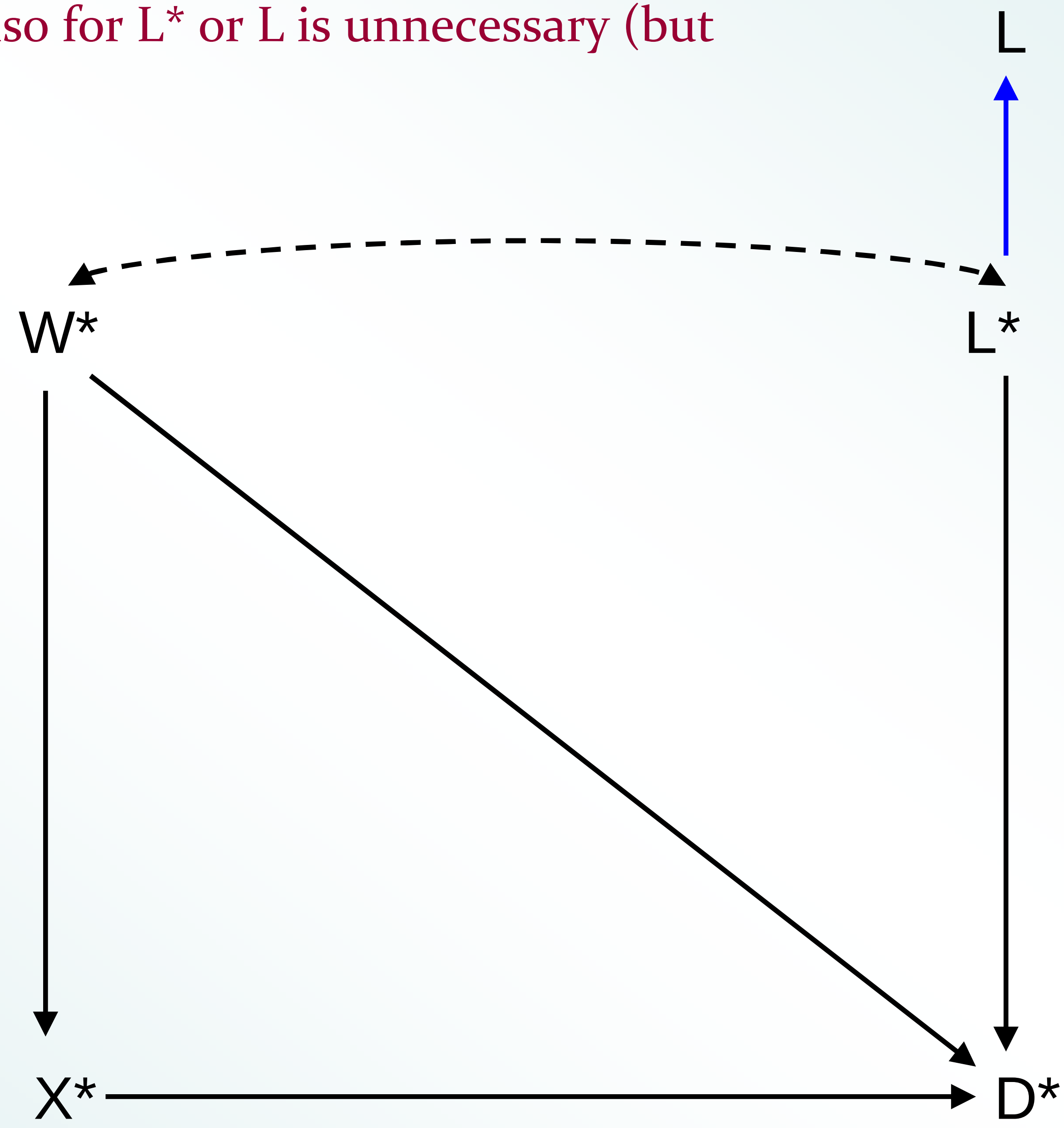
$X \leftarrow X^* \rightarrow D^* + X \leftarrow X^* \leftarrow Z_1^* \rightarrow D^* + X \leftarrow X^* \leftarrow Z_2^* \rightarrow D^* + X \leftarrow X^* \leftarrow Z_1^* \leftarrow\rightarrow Z_2^* \rightarrow D^*$

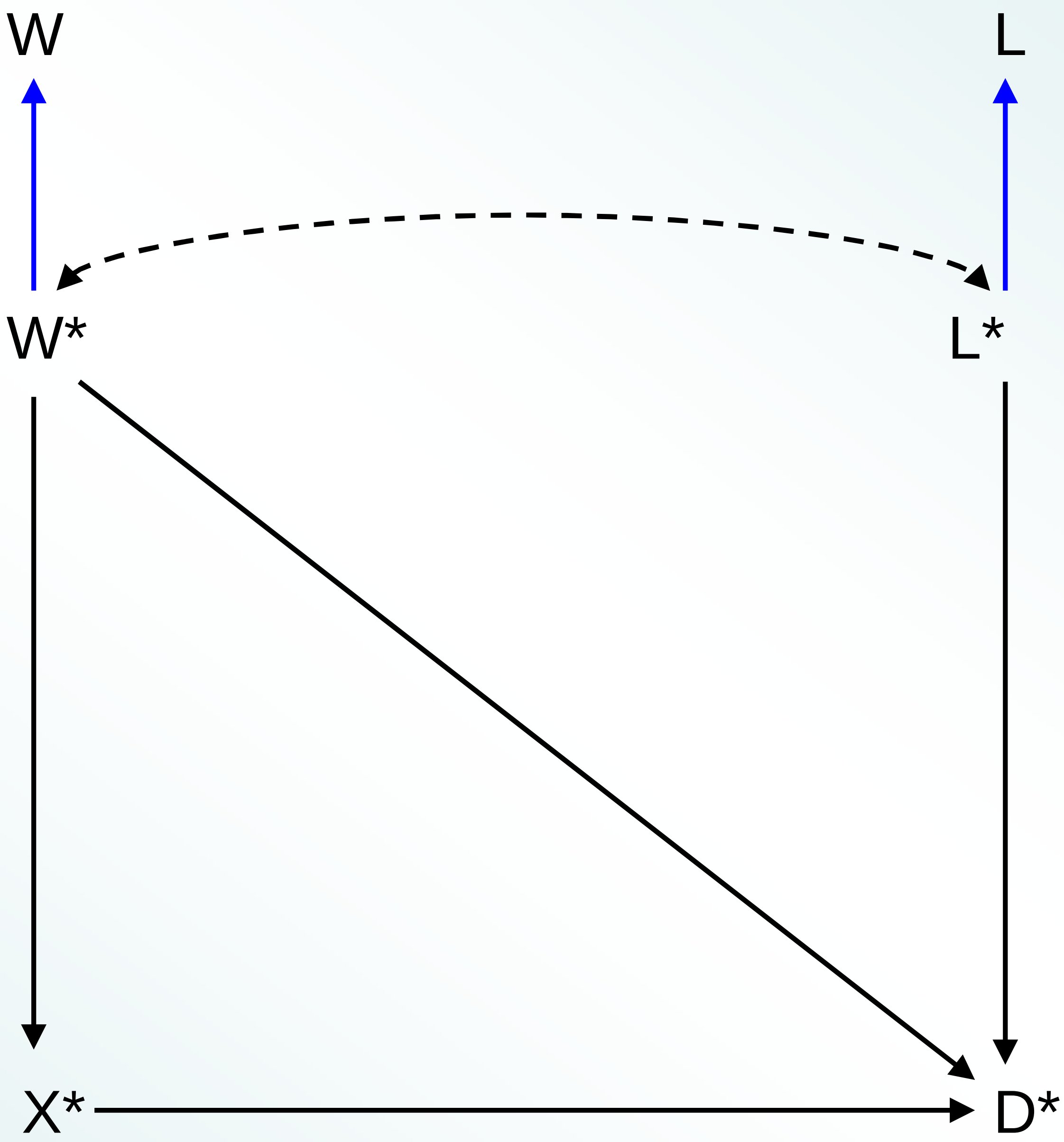
Uncontrolled Confounding with Measurement Error


Z
λ_Z
U*
Z*
X
X*
D*
λ_X

Unnecessary Surrogate

In estimating effect of X^* on D^*, we need only adjust for W^*. Adjusting also for L^* or L is unnecessary (but harmless).





If W* is mismeasured as W, then adjusting for L now further biases the X*→D* relationship. Note that adjusting for L* is helpful.

SUMMARY

Augmented graphs show that target parameters are biased through all backdoors and front-doors left open by the unobserved true variables.

Other important results include the under-appreciated finding that in the absence of measurement error in confounders, the treatment effect is only biased by the degree to which the treatment variable is mismeasured.

Furthermore, augmented graphs support the use of regression calibration (within the limits of linearity assumptions).

Funding Acknowledgement:

Rubicon Fellowship (grant #825.06.026)
Netherlands Organization for Scientific Research (NWO)