Statistical Rethinking

2023

6. Good & Bad Controls

Avoid Being Clever At All Costs

Being clever: unreliable, opaque

Given a causal model, can use logic to derive implications

Others can use same logic to verify & challenge your work

Better than clever

The Fork

$X \longleftarrow Z \longrightarrow Y$

The Pipe

$X \longrightarrow Z \longrightarrow Y$

The Collider

$X \longrightarrow Z \longleftarrow Y$

The Descendant

$$\begin{tikzcd} X \arrow[r] & Z \arrow[r] \arrow[d] & Y \\ & A & \end{tikzcd}$$

The Fork

$X \longleftarrow Z \longrightarrow Y$

X and *Y* associated unless stratify by *Z*

The Pipe

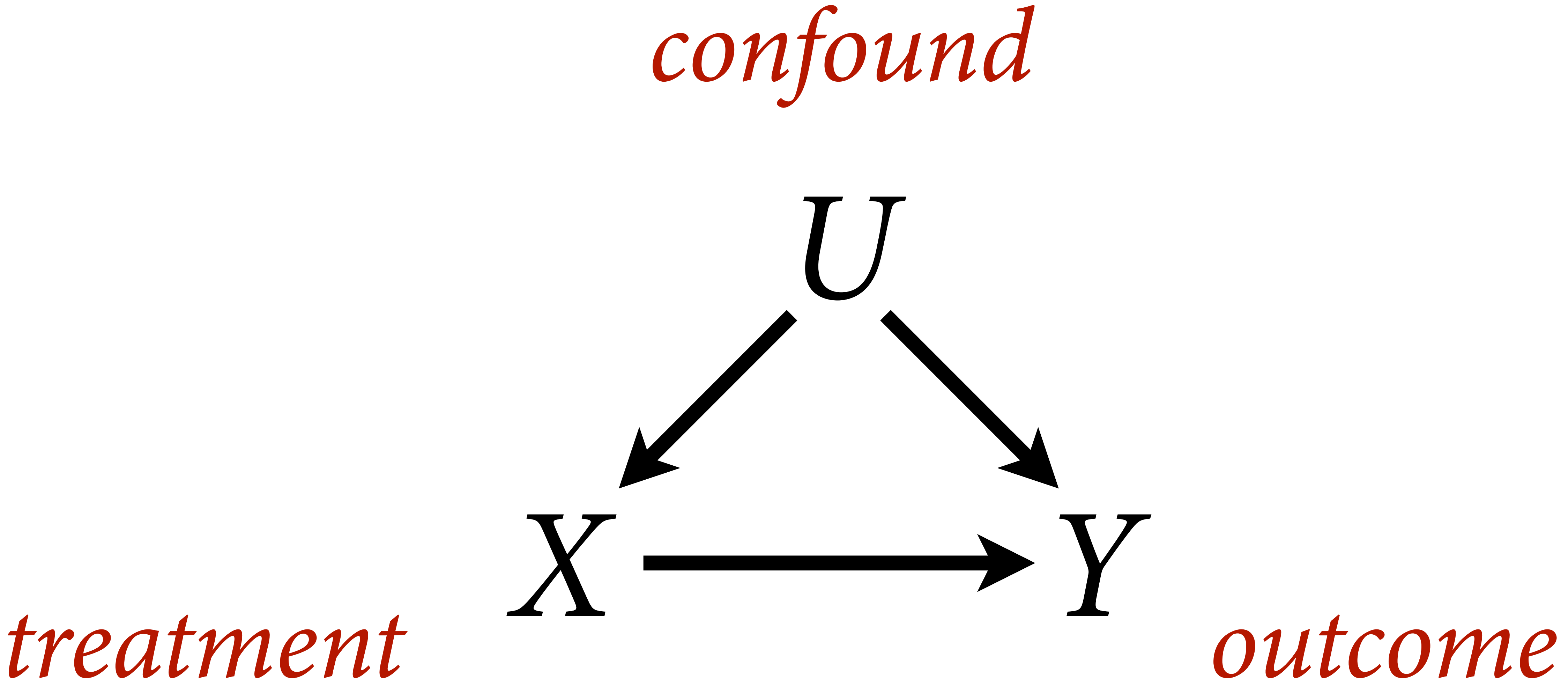
$X \longrightarrow Z \longrightarrow Y$

X and *Y* associated unless stratify by *Z*

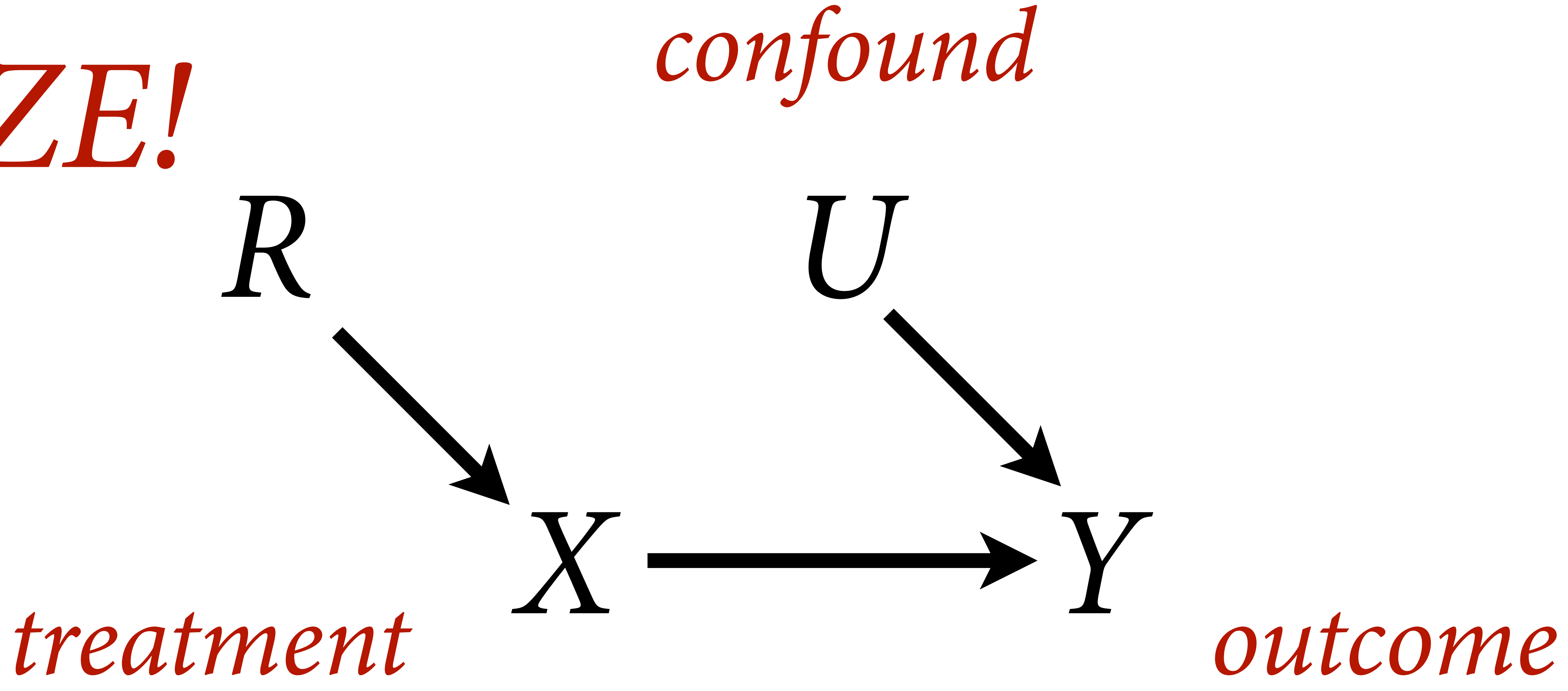
The Collider

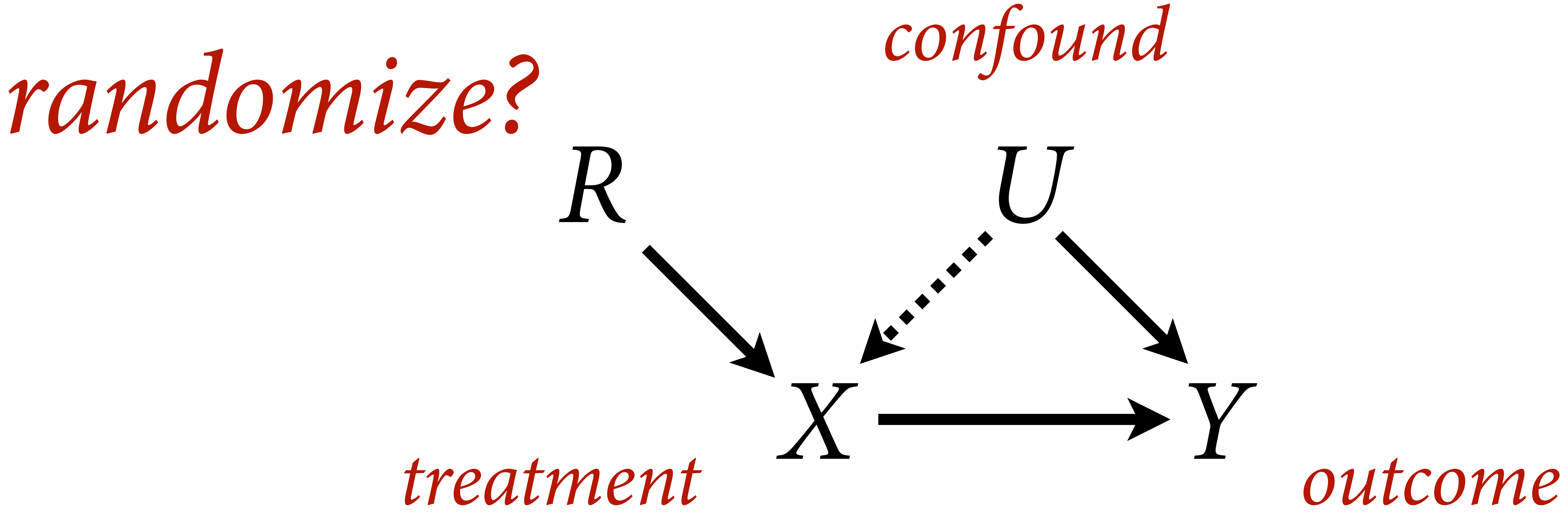
$X \longrightarrow Z \longleftarrow Y$

X and *Y* not associated unless stratify by *Z*


confound
U
X
Y
treatment
outcome

RANDOMIZE!





randomize?
confound
R
U
X
Y
treatment
outcome

Causal Thinking

In an experiment, we cut causes of the treatment

We ***randomize*** (we try at least)

So how does causal inference without randomization ever work?

Is there a statistical procedure that mimics randomization?

Without randomization

$$\begin{tikzcd} & U \arrow[dl] \arrow[dr] & \\ X \arrow[rr] & & Y \end{tikzcd}$$

With randomization

$$\begin{tikzcd} & U \arrow[dr] & \\ X \arrow[rr, "\mathrm{do}(X)"'] & & Y \end{tikzcd}$$

Causal Thinking

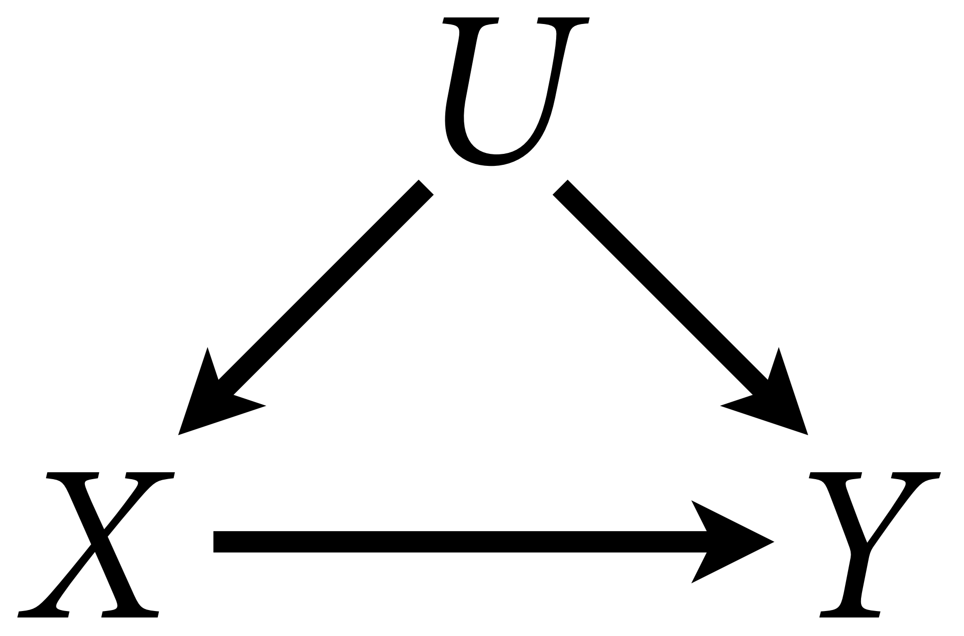
Is there a statistical procedure that mimics randomization?

$$P(Y \mid \mathrm{do}(X)) = P(Y \mid ?)$$

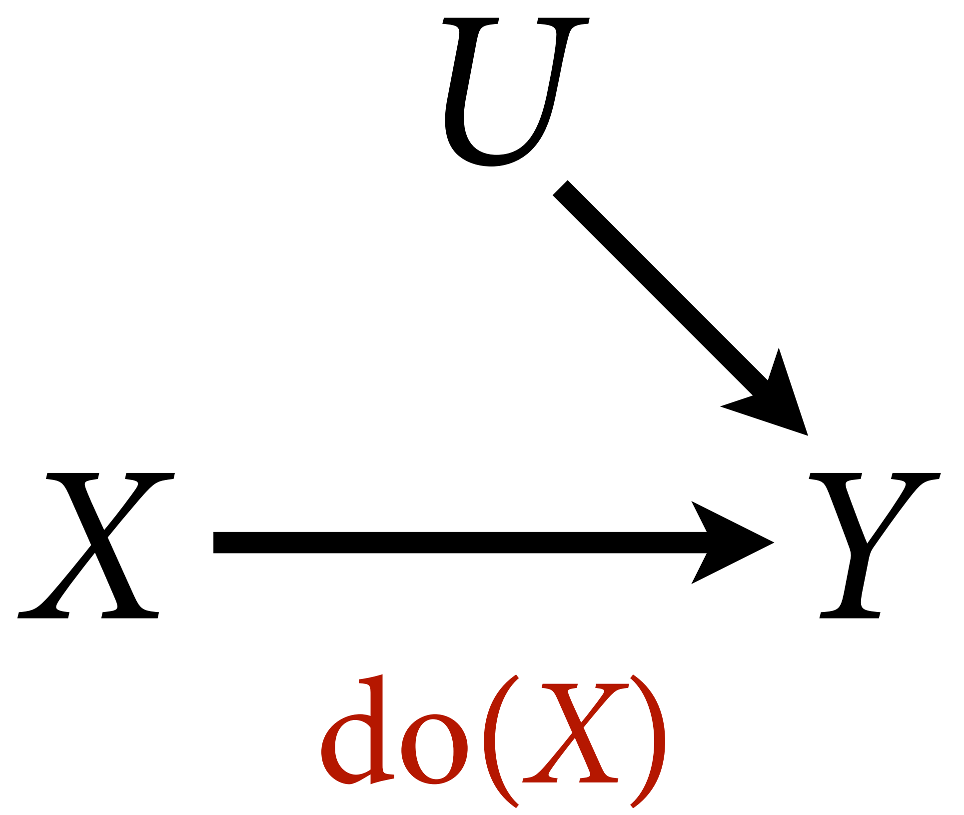
do(X) means intervene on X

Can analyze causal model to find answer (if it exists)

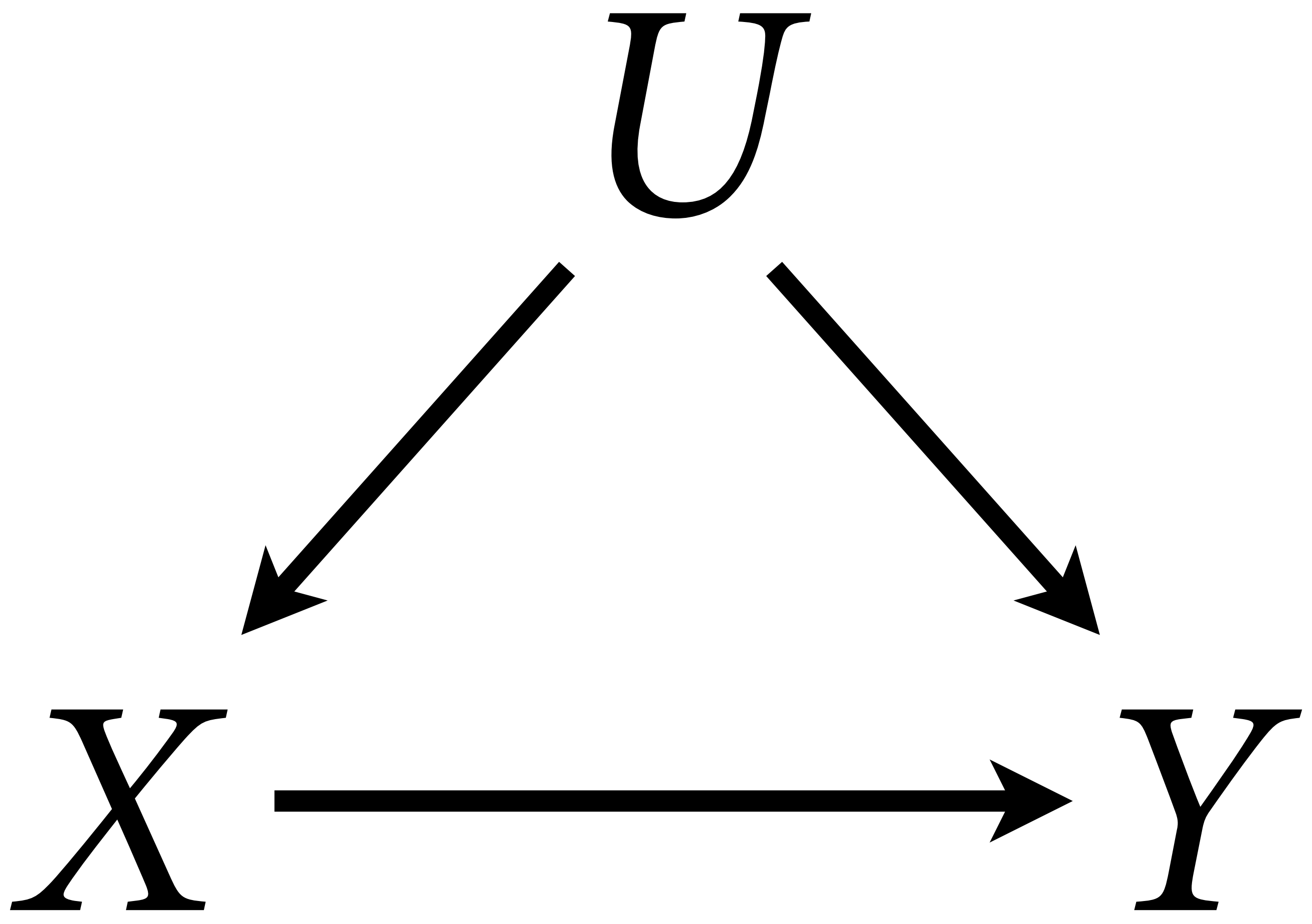
Without randomization



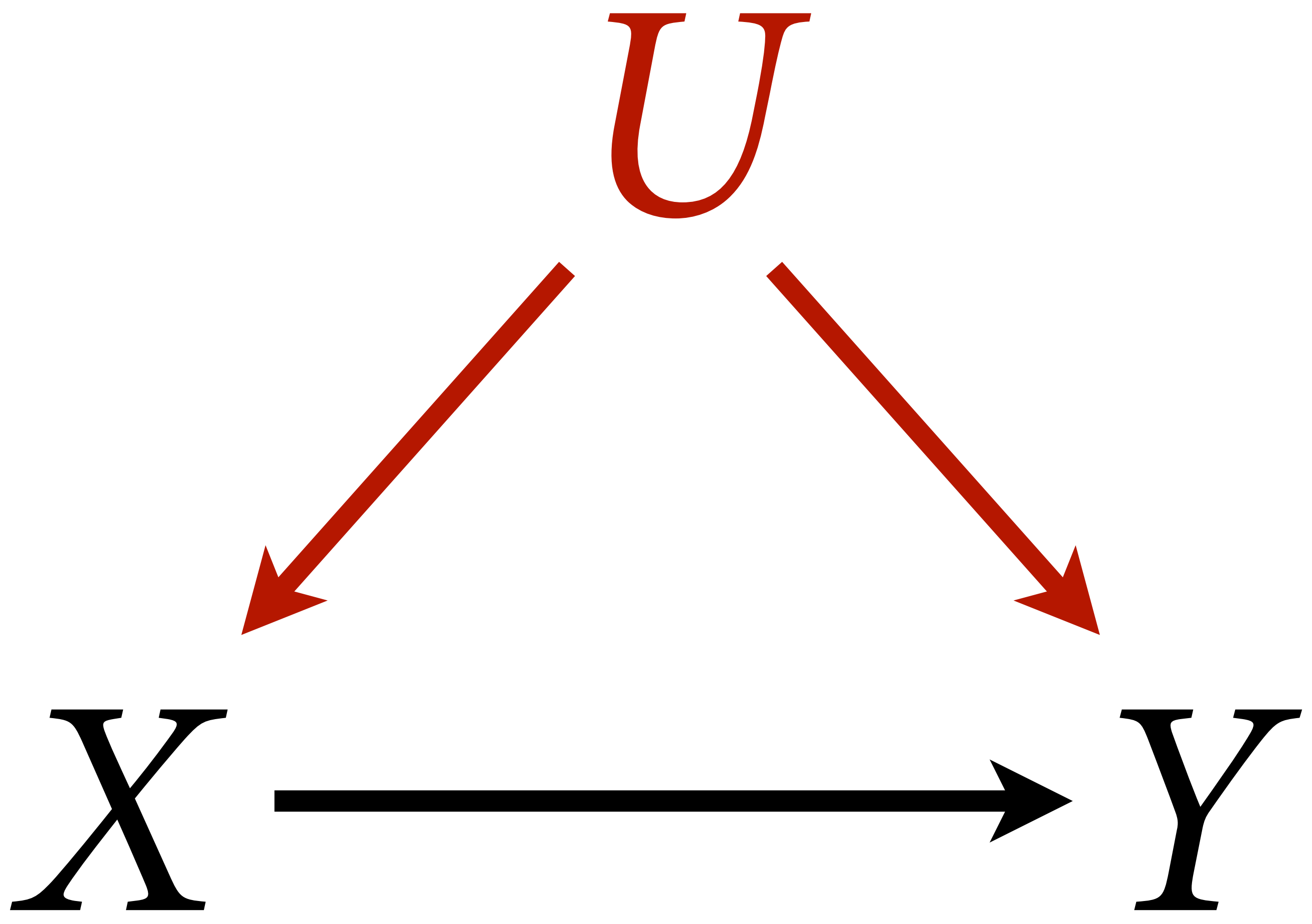
With randomization



Example: Simple Confound



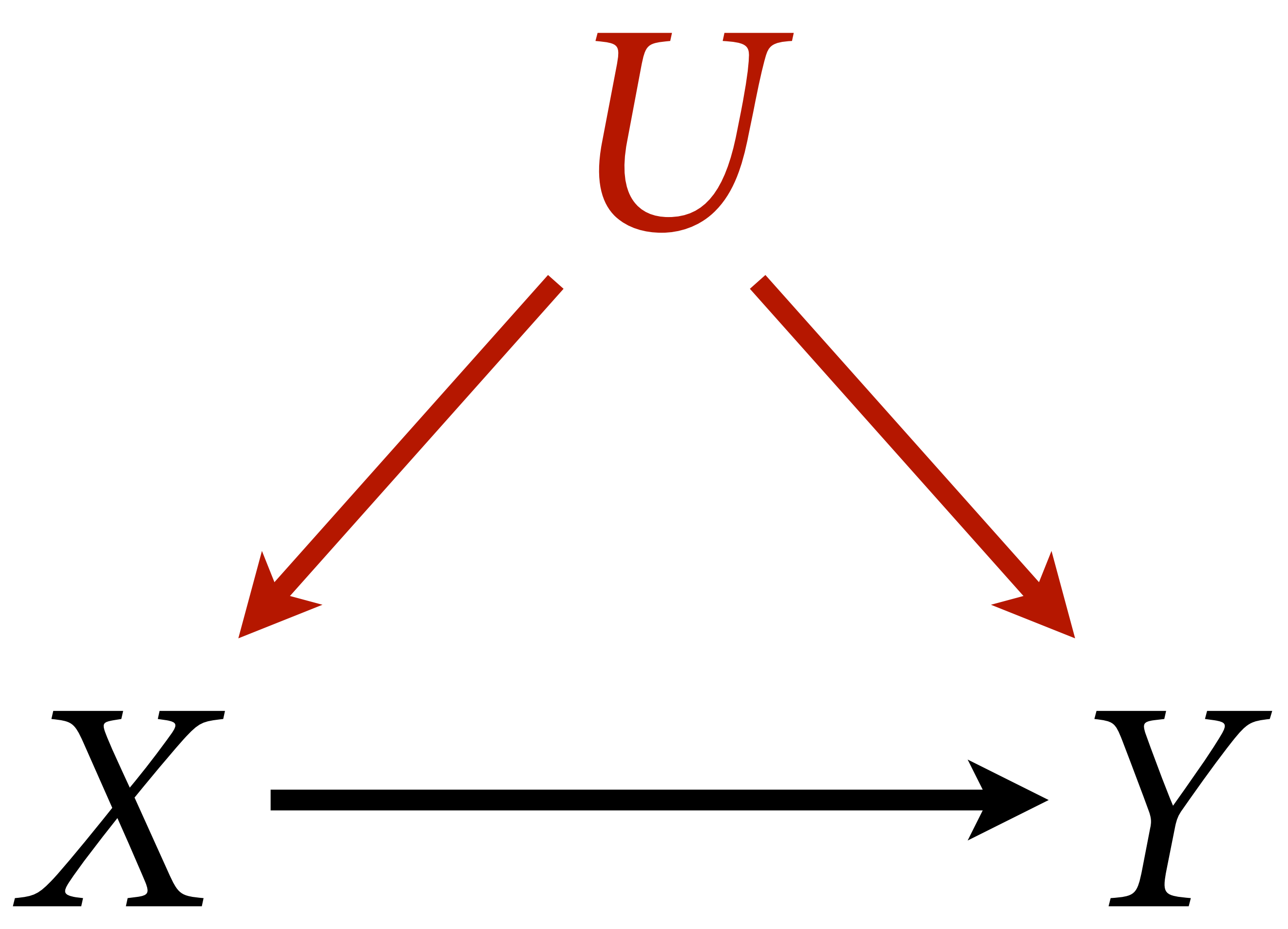
Example: Simple Confound



Non-causal path
X <– U –> Y

Close the fork!
Condition on U

Example: Simple Confound



Non-causal path
X <– U –> Y

Close the fork!
Condition on U

$$P(Y|\mathrm{do}(X)) = \sum_U P(Y|X,U)P(U) = \mathrm{E}_U P(Y|X,U)$$

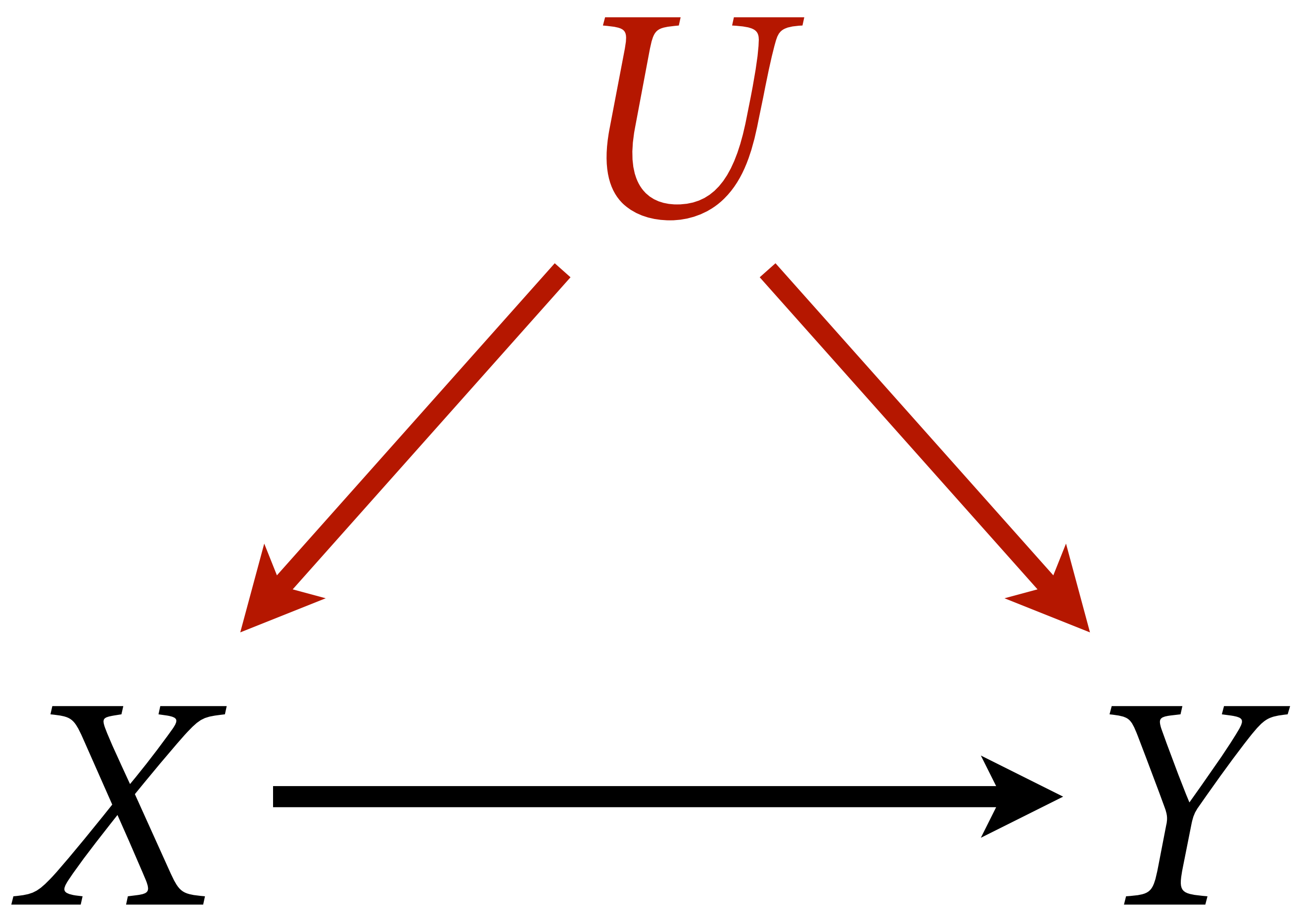
"The distribution of Y, stratified by X and U, averaged over the distribution of U."

$$P(Y|\mathrm{do}(X)) = \sum_{U} P(Y|X,U)P(U) = \mathrm{E}_{U}P(Y|X,U)$$

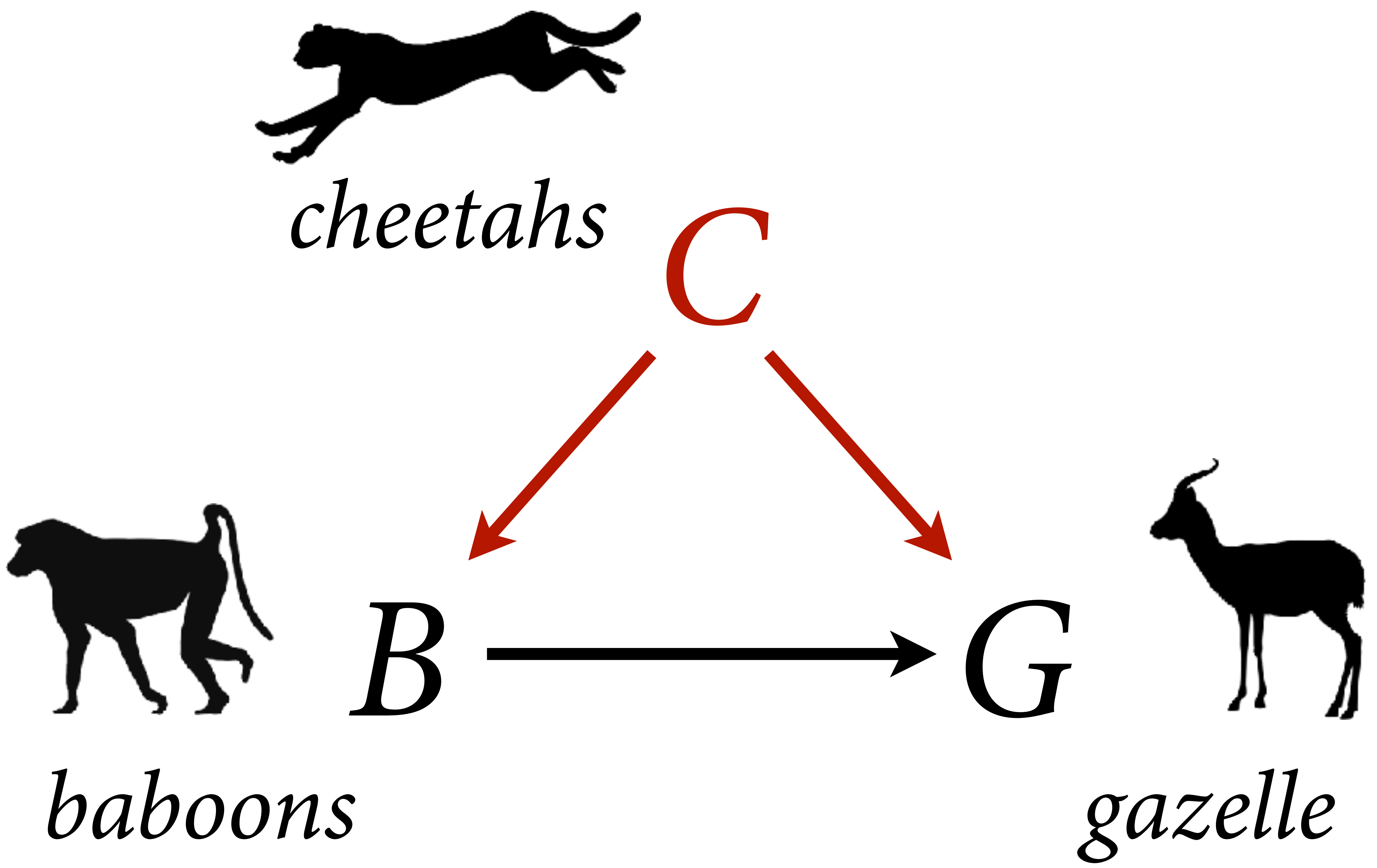
"The distribution of Y, stratified by X and U, averaged over the distribution of U."

The causal effect of X on Y is **not** (in general) the **coefficient** relating X to Y

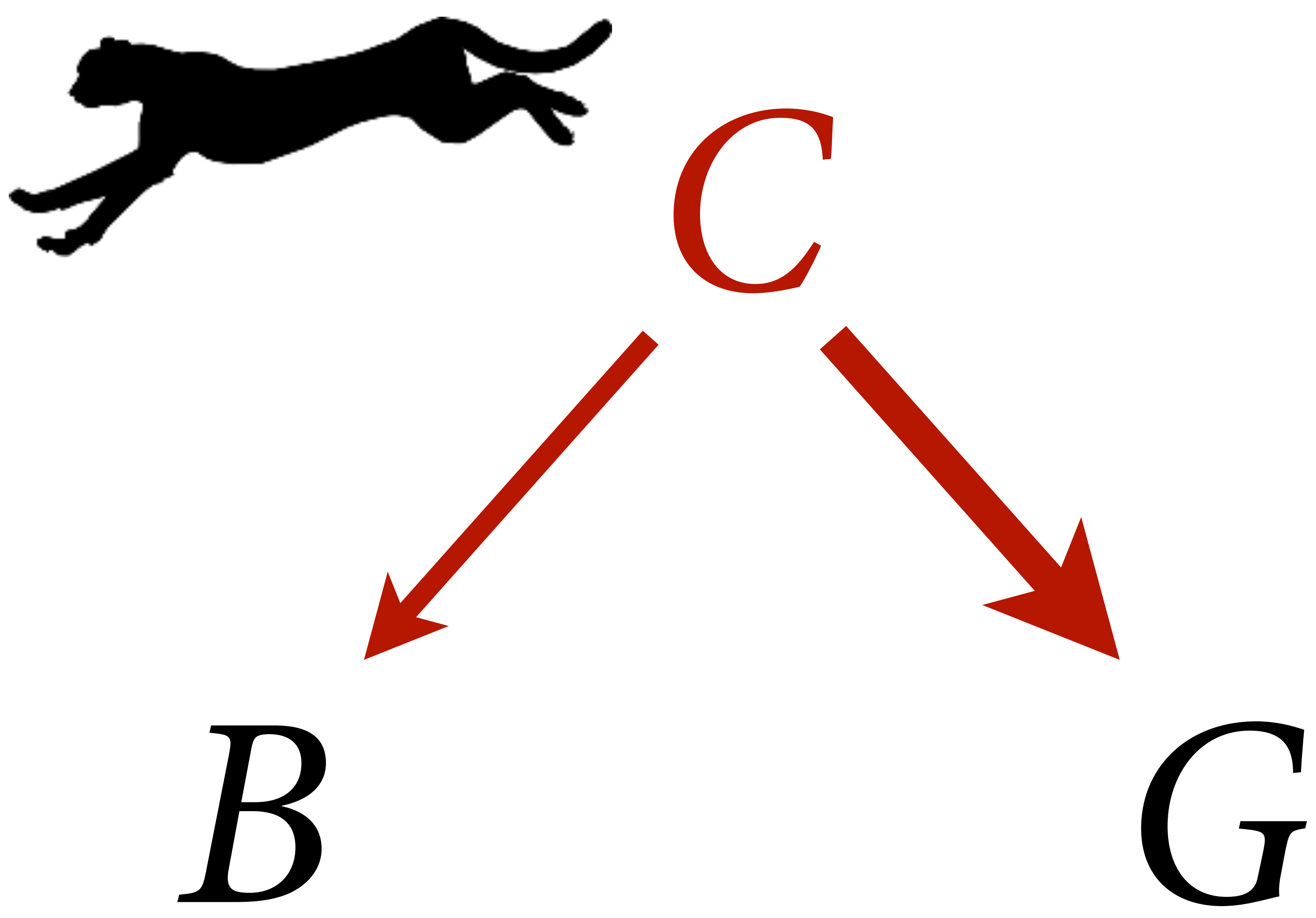
It is the distribution of Y when we change X, **averaged** over the distributions of the control variables (here U)



Marginal Effects Example



cheetahs present



cheetahs present *cheetahs absent*

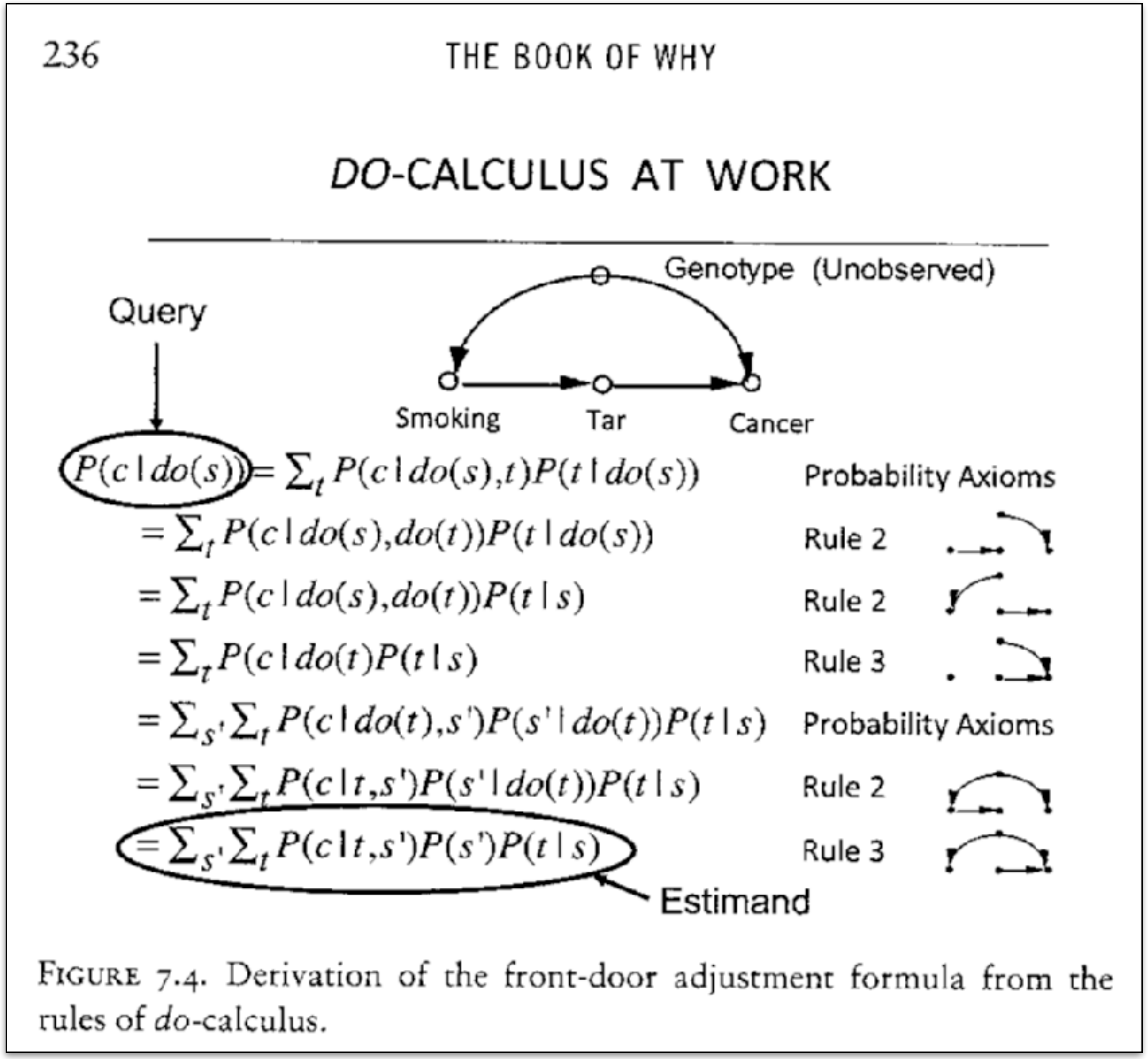
Causal effect of baboons depends upon distribution of cheetahs

do-calculus

For DAGs, rules for finding $P(Y|\text{do}(X))$ known as **do-calculus**

do-calculus says what is possible to say **before** picking functions

Justifies **graphical** analysis


236 THE BOOK OF WHY

DO-CALCULUS AT WORK

$$P(c \mid do(s)) = \sum_t P(c \mid do(s), t) P(t \mid do(s))$$ Probability Axioms

$$= \sum_t P(c \mid do(s), do(t)) P(t \mid do(s))$$ Rule 2

$$= \sum_t P(c \mid do(s), do(t)) P(t \mid s)$$ Rule 2

$$= \sum_t P(c \mid do(t) P(t \mid s)$$ Rule 3

$$= \sum_{s'} \sum_t P(c \mid do(t), s') P(s' \mid do(t)) P(t \mid s)$$ Probability Axioms

$$= \sum_{s'} \sum_t P(c \mid t, s') P(s' \mid do(t)) P(t \mid s)$$ Rule 2

$$= \sum_{s'} \sum_t P(c \mid t, s') P(s') P(t \mid s)$$ Rule 3

FIGURE 7.4. Derivation of the front-door adjustment formula from the rules of *do*-calculus.

Do calculus, not too much, mostly graphs

do-calculus

do-calculus is **worst case**:
additional assumptions often allow stronger inference

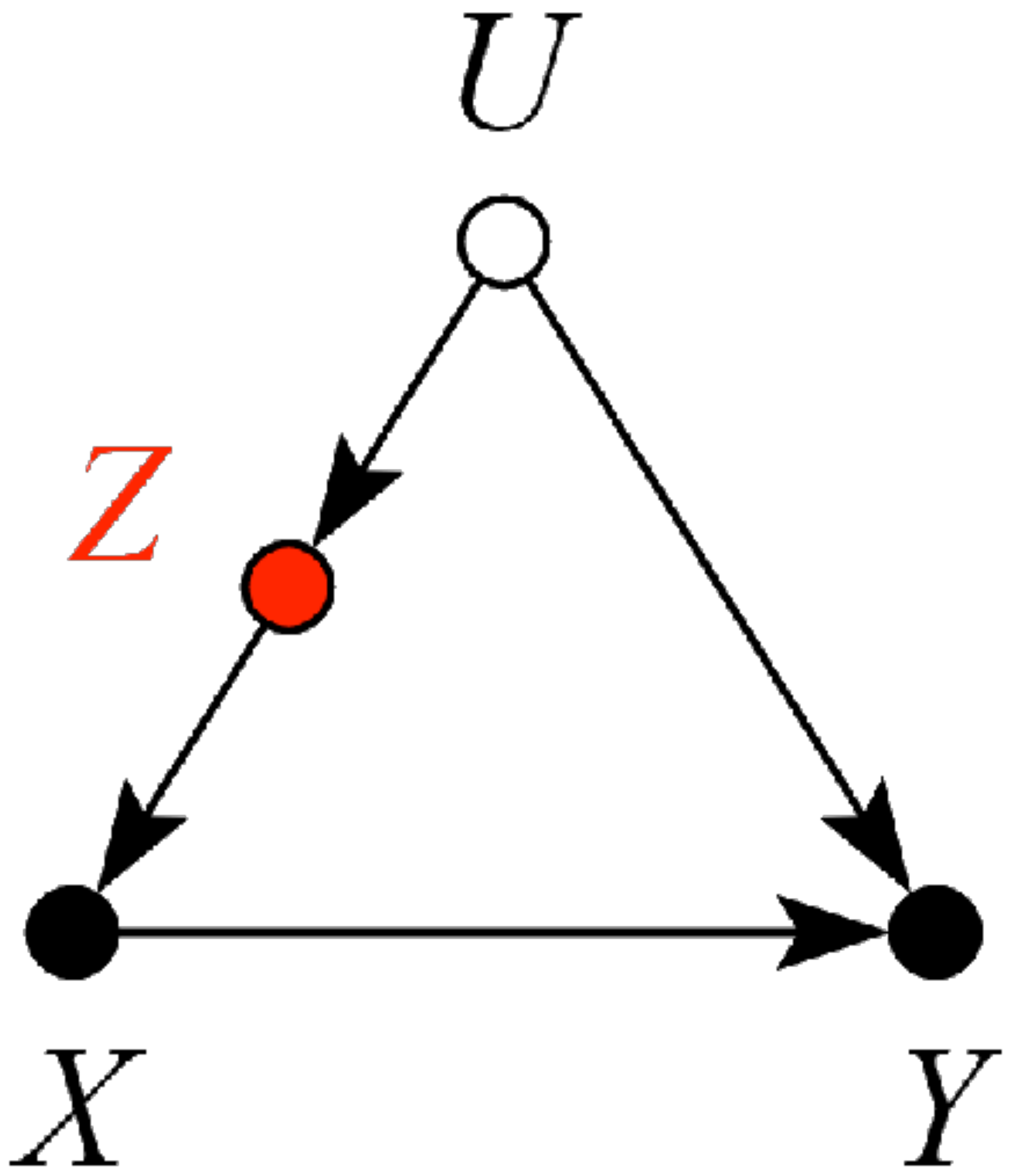
do-calculus is **best case**:
if inference possible by do-calculus, does not depend on special assumptions

Judea Pearl, father of do-calculus (1966)

Backdoor Criterion

Backdoor Criterion is a shortcut to applying (some) results of do-calculus

Can be performed with your **eyeballs**

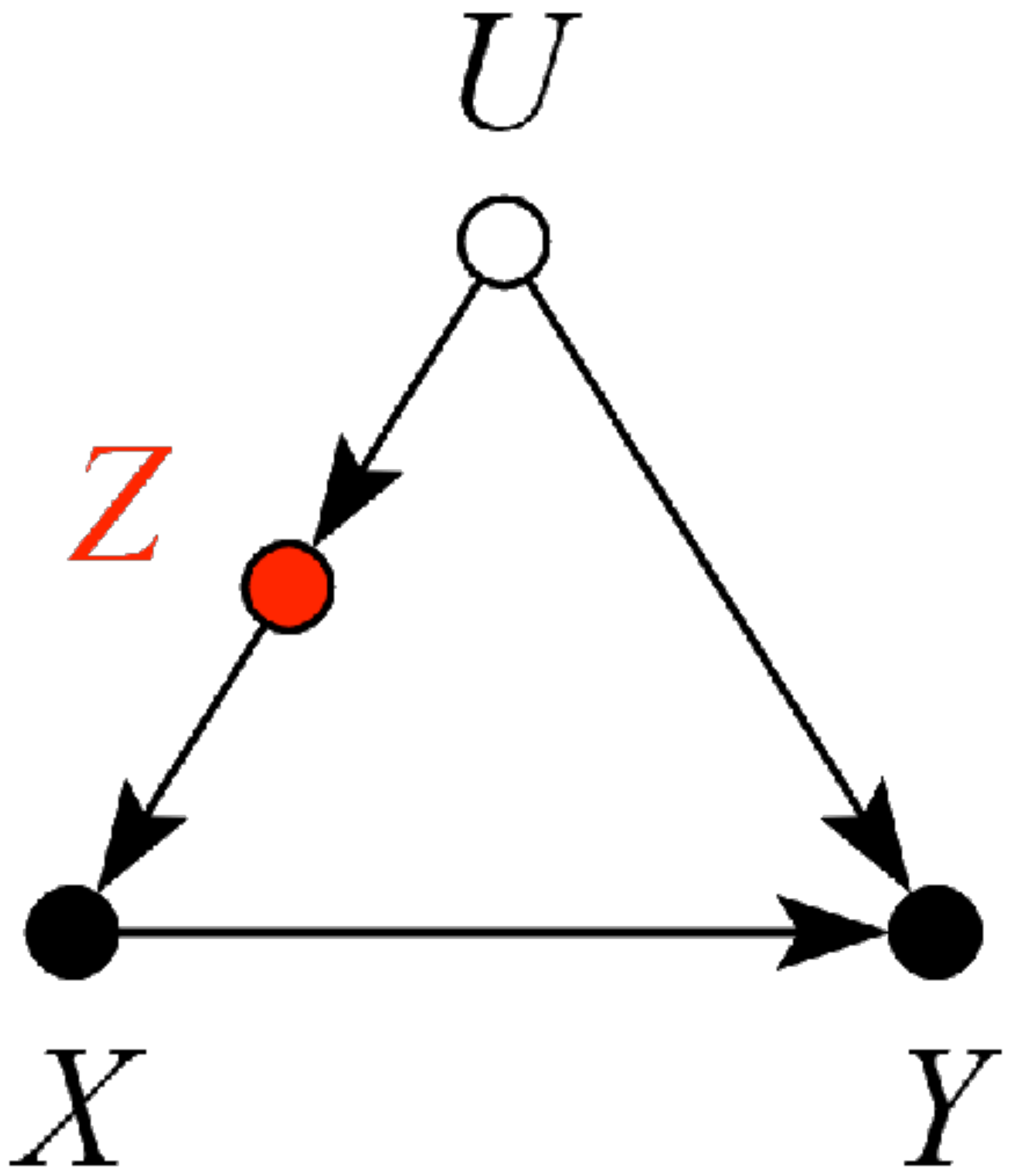


Backdoor Criterion: Rule to find a set of variables to stratify by to yield $P(Y|\text{do}(X))$

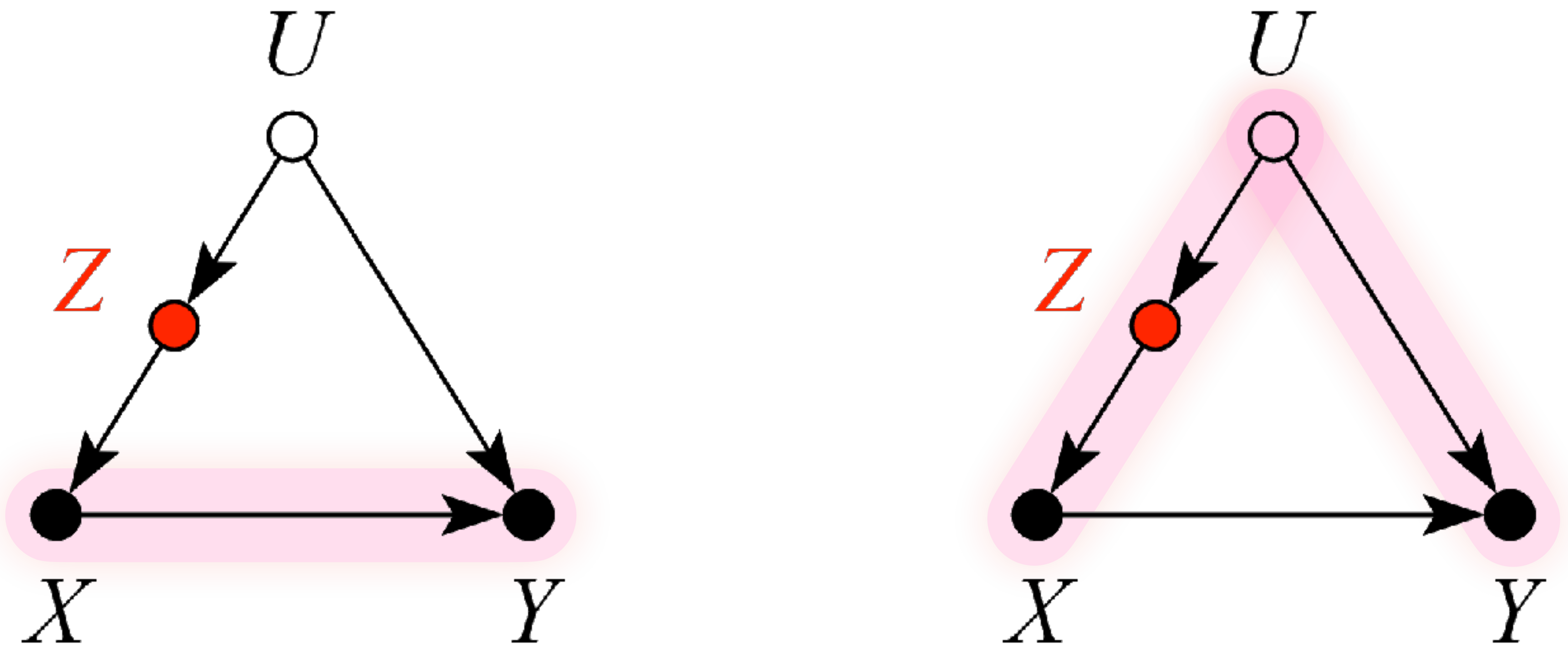
(1) Identify all **paths** connecting the treatment (X) to the outcome (Y)

(2) Paths with arrows **entering** X are backdoor paths (non-causal paths)

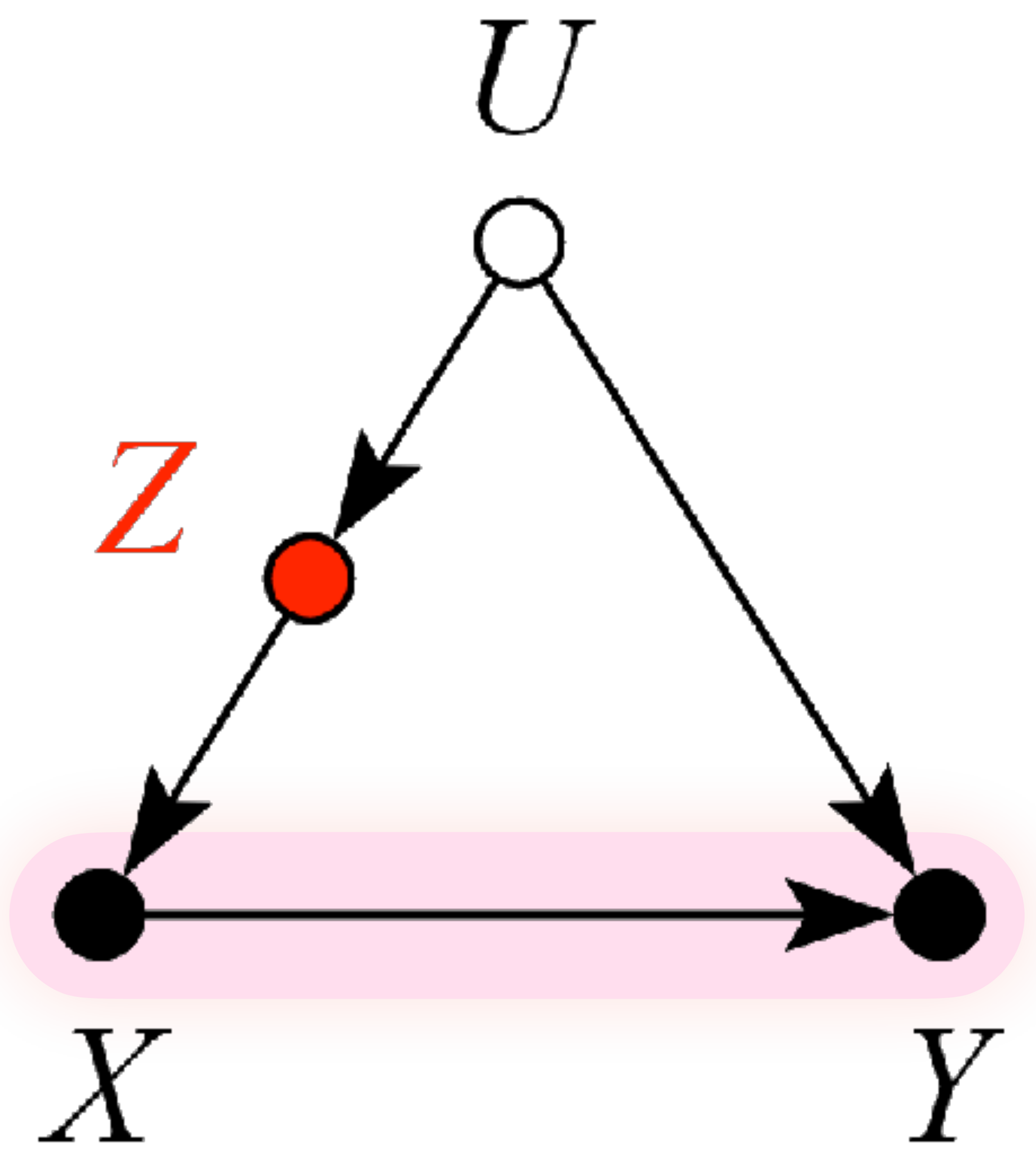
(3) Find adjustment set that closes/blocks all backdoor paths

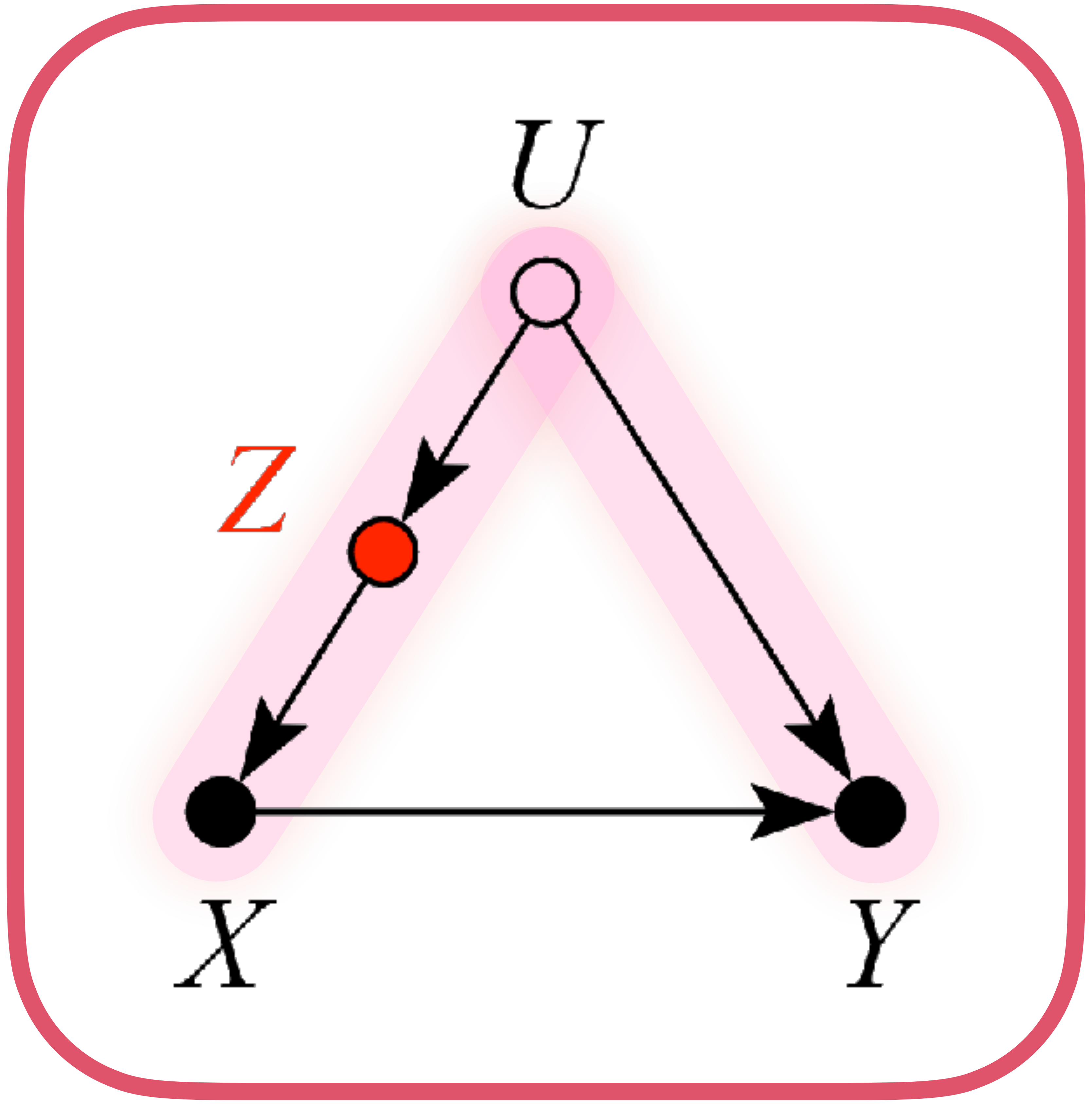


(1) Identify all **paths** connecting the treatment (X) to the outcome (Y)



(2) Paths with arrows **entering** X are backdoor paths (confounding paths)

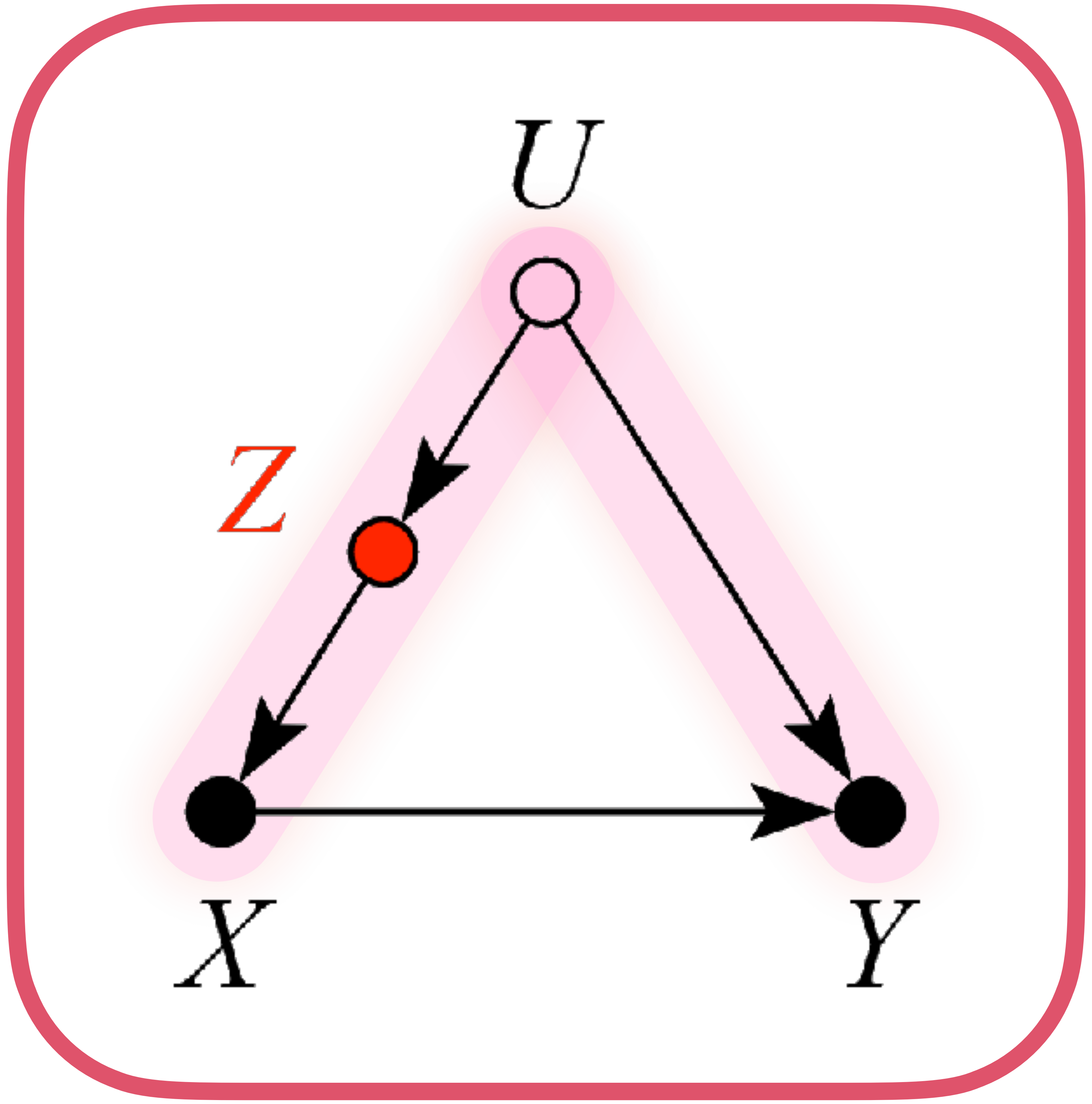




(3) Find a set of control variables that close/block all backdoor paths

Block the pipe: $X \perp\!\!\!\perp U \mid Z$

Z "knows" all of the association between X, Y that is due to U

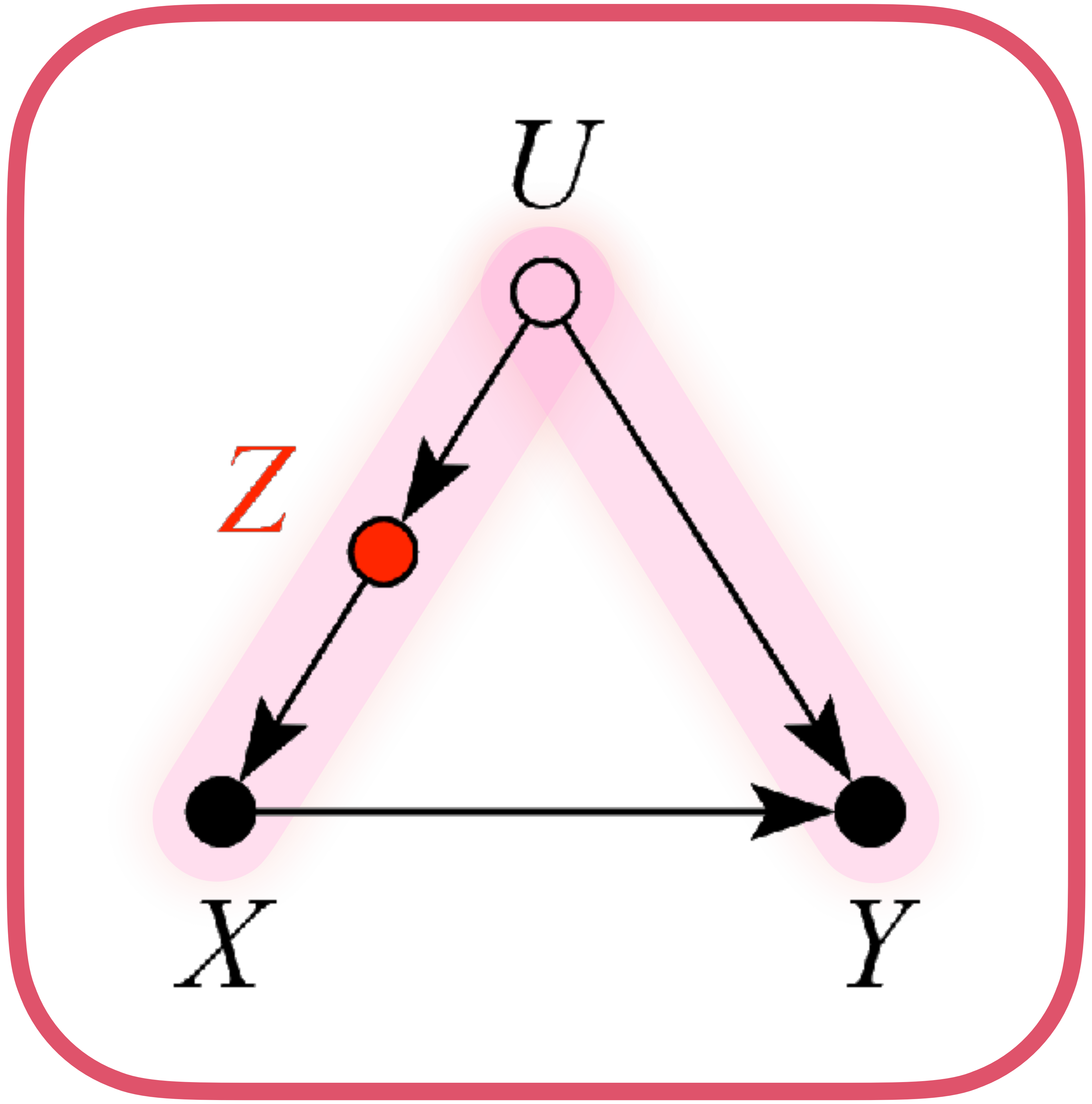


(3) Find a set of control variables that close/block all backdoor paths

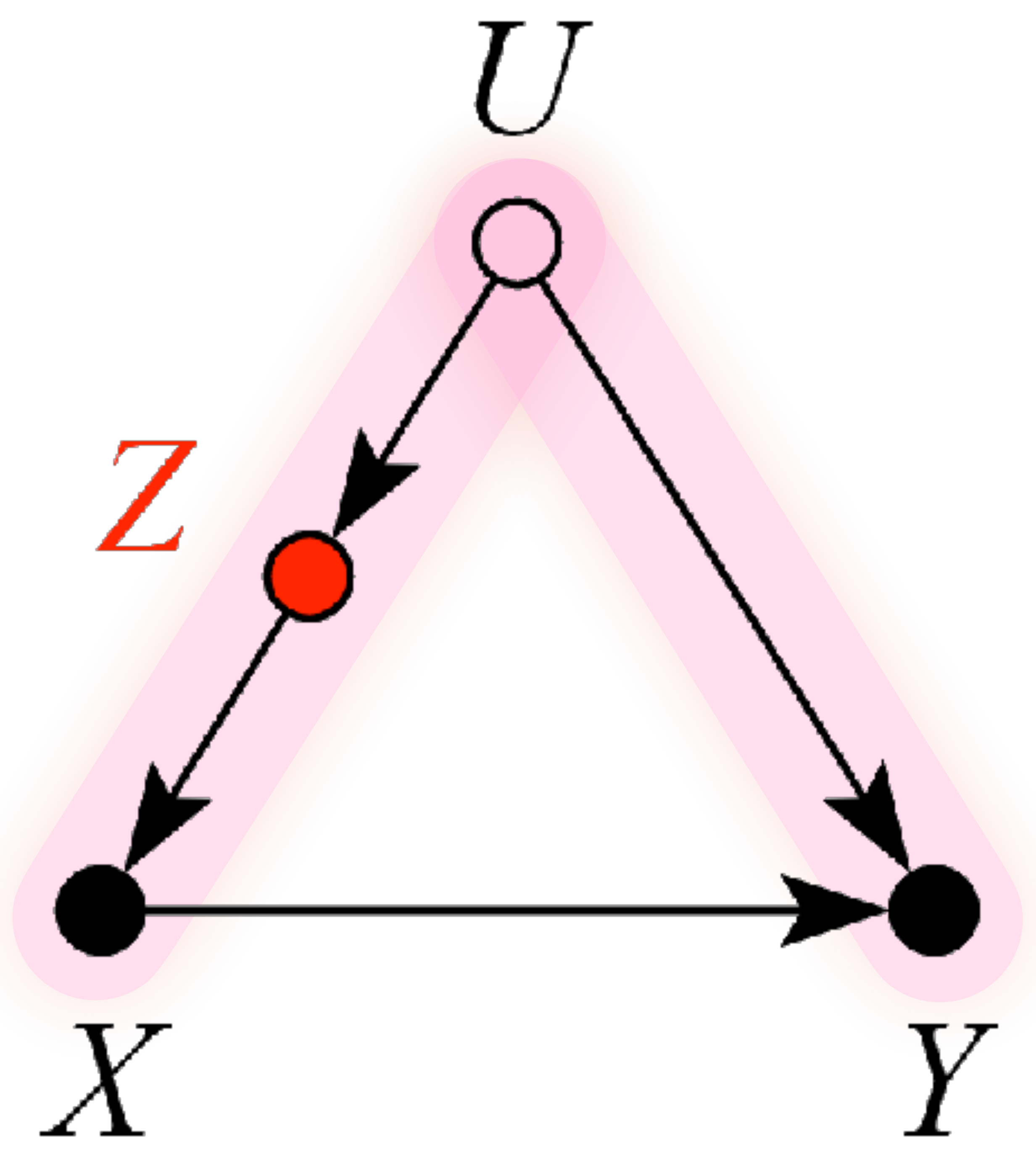
Block the pipe: $X \perp\!\!\!\perp U \mid Z$

$$P(Y \mid \mathrm{do}(X)) = \sum_z P(Y \mid X, Z) P(Z = z)$$

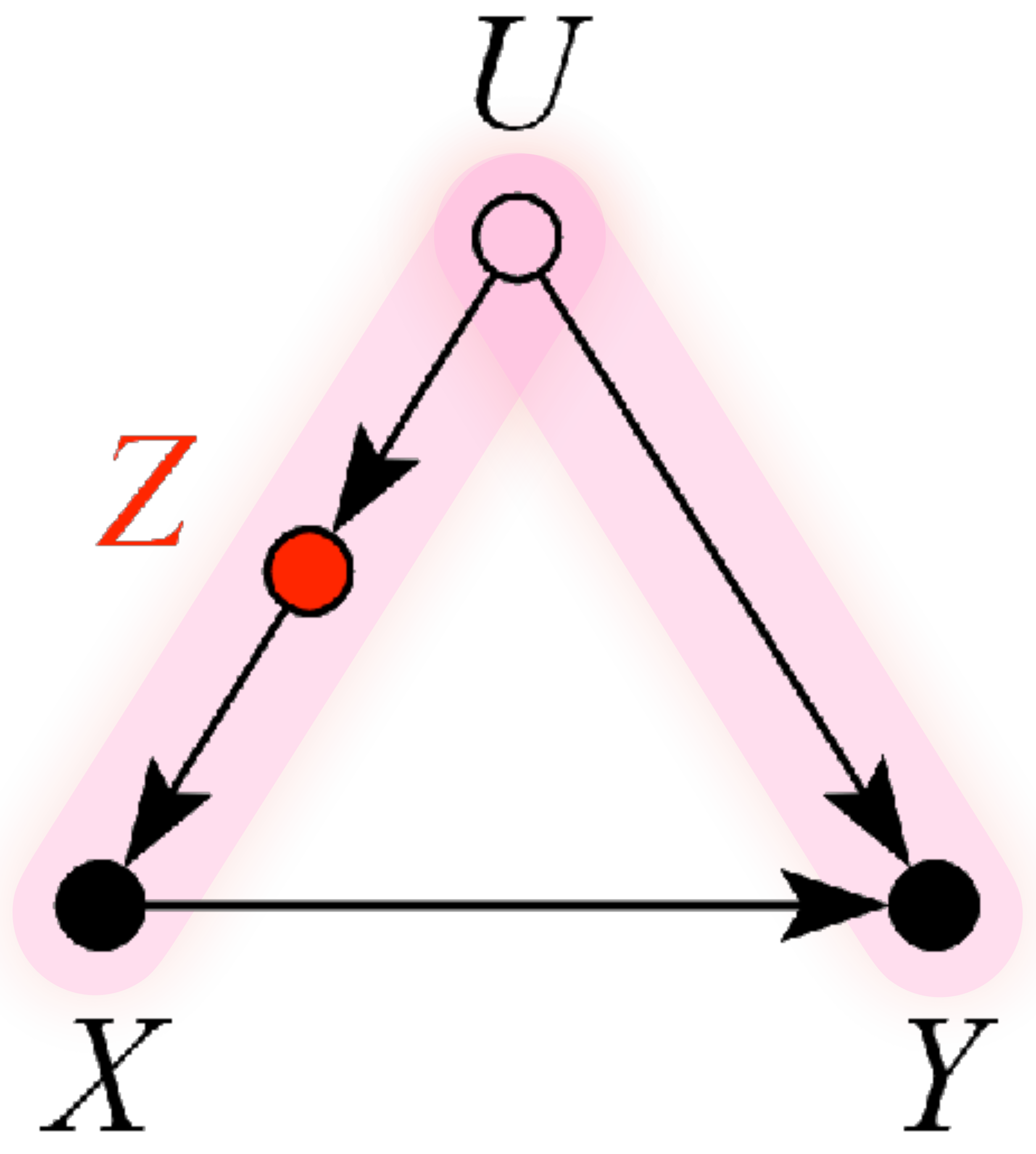
$$Y_i \sim \mathrm{Normal}(\mu_i, \sigma)$$
$$\mu_i = \alpha + \beta_X X_i + \beta_Z Z_i$$

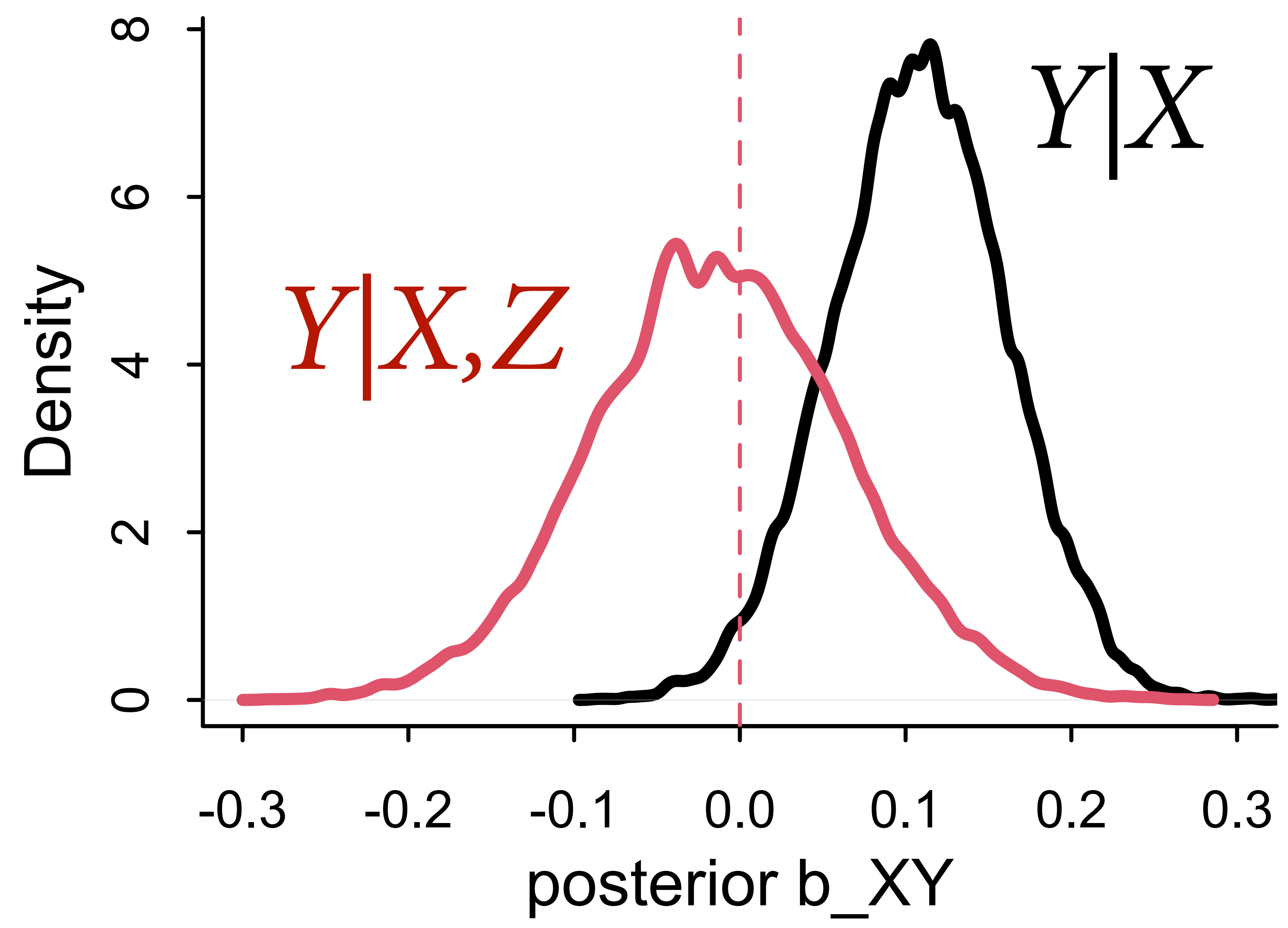


```r
# simulate confounded Y
N <- 200
b_XY <- 0
b_UY <- -1
b_UZ <- -1
b_ZX <- 1

set.seed(10)
U <- rbern(N)
Z <- rnorm(N,b_UZ*U)
X <- rnorm(N,b_ZX*Z)
Y <- rnorm(N,b_XY*X+b_UY*U)
d <- list(Y=Y,X=X,Z=Z)
```

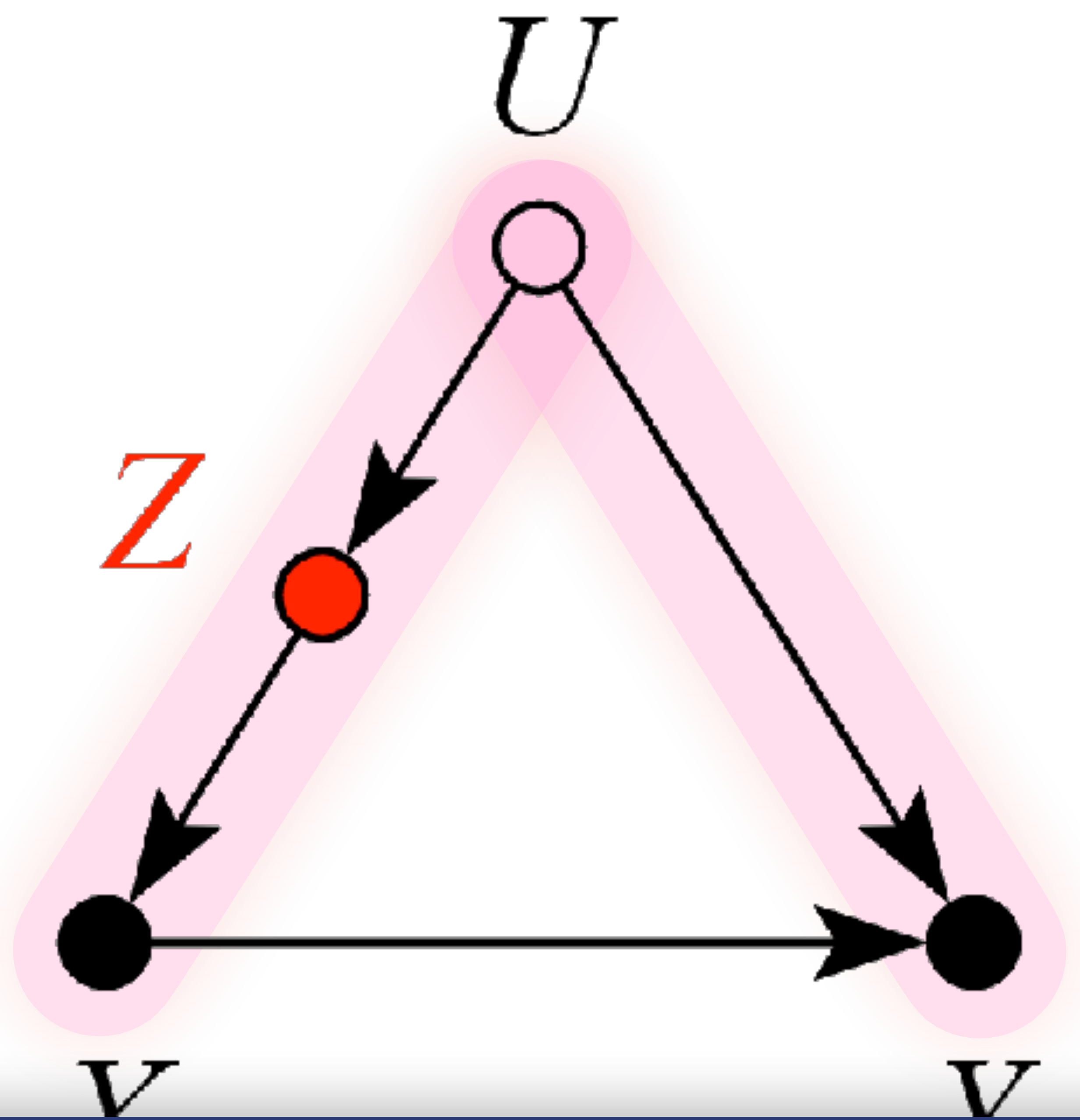


```
# ignore U,Z
m_YX <- quap(
    alist(
        Y ~ dnorm( mu , sigma ),
        mu <- a + b_XY*X,
        a ~ dnorm( 0 , 1 ),
        b_XY ~ dnorm( 0 , 1 ),
        sigma ~ dexp( 1 )
    ), data=d )

# stratify by Z
m_YXZ <- quap(
    alist(
        Y ~ dnorm( mu , sigma ),
        mu <- a + b_XY*X + b_Z*Z,
        a ~ dnorm( 0 , 1 ),
        c(b_XY,b_Z) ~ dnorm( 0 , 1 ),
        sigma ~ dexp( 1 )
    ), data=d )

post <- extract.samples(m_YX)
post2 <- extract.samples(m_YXZ)
dens(post$b_XY,lwd=3,col=1,xlab="posterior
b_XY",xlim=c(-0.3,0.3))
dens(post2$b_XY,lwd=3,col=2,add=TRUE)
```





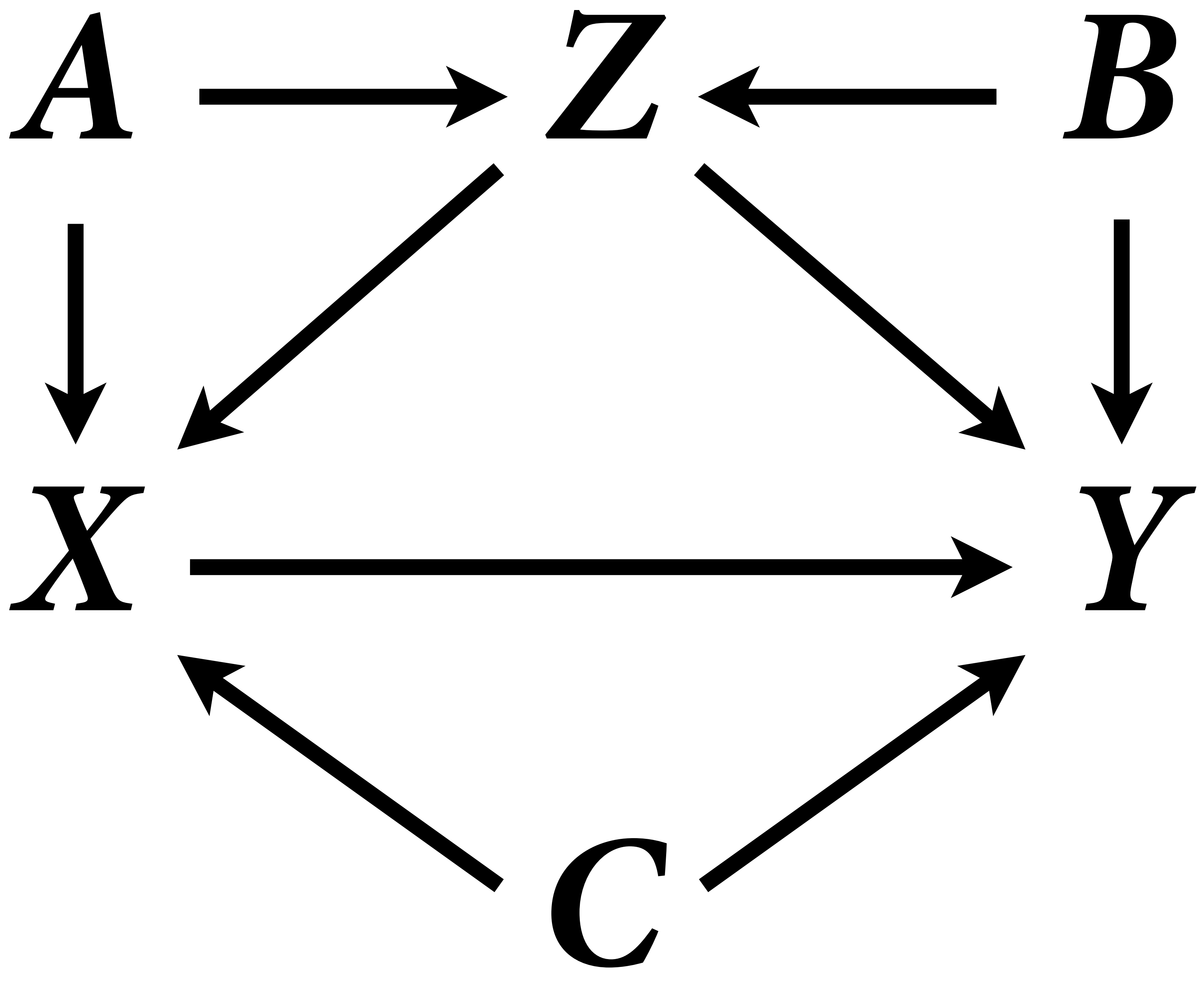
```
# ignore U,Z
m_YX <- quap(
    alist(
        Y ~ dnorm( mu , sigma ),
        mu <- a + b_XY*X,
        a ~ dnorm( 0 , 1 ),
        b_XY ~ dnorm( 0 , 1 ),
        sigma ~ dexp( 1 )
    ), data=d )

# stratify by Z
m_YXZ <- quap(
    alist(
        Y ~ dnorm( mu , sigma ),
        mu <- a + b_XY*X + b_Z*Z,
        a ~ dnorm( 0 , 1 ),
        c(b_XY,b_Z) ~ dnorm( 0 , 1 ),
        sigma ~ dexp( 1 )
    ), data=d )

post <- extract.samples(m_YX)
post2 <- extract.samples(m_YXZ)
dens(post$b_XY,lwd=3,col=1,xlab="posterior
b_XY",xlim=c(-0.3,0.3))
dens(post2$b_XY,lwd=3,col=2,add=TRUE)
```

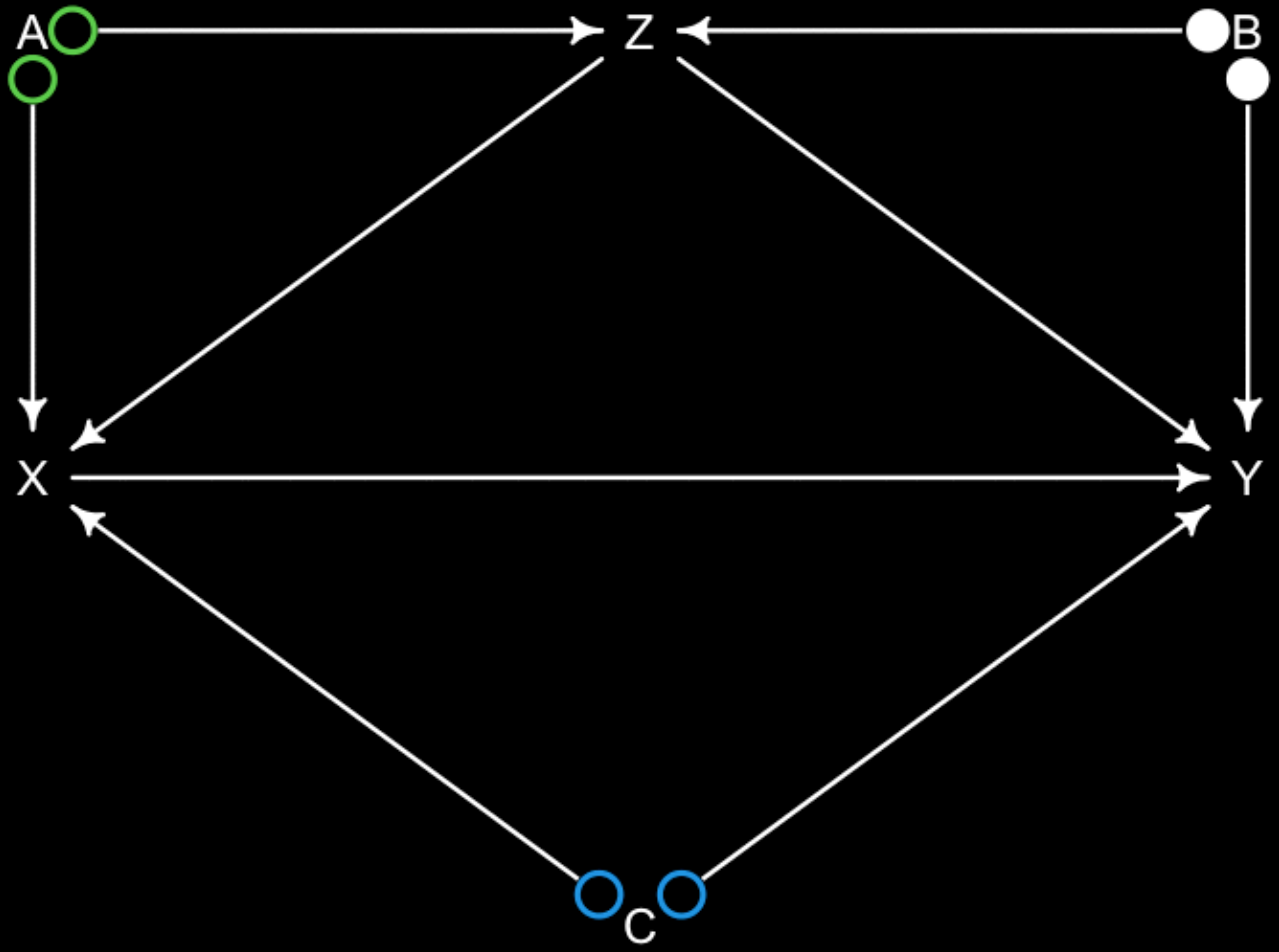


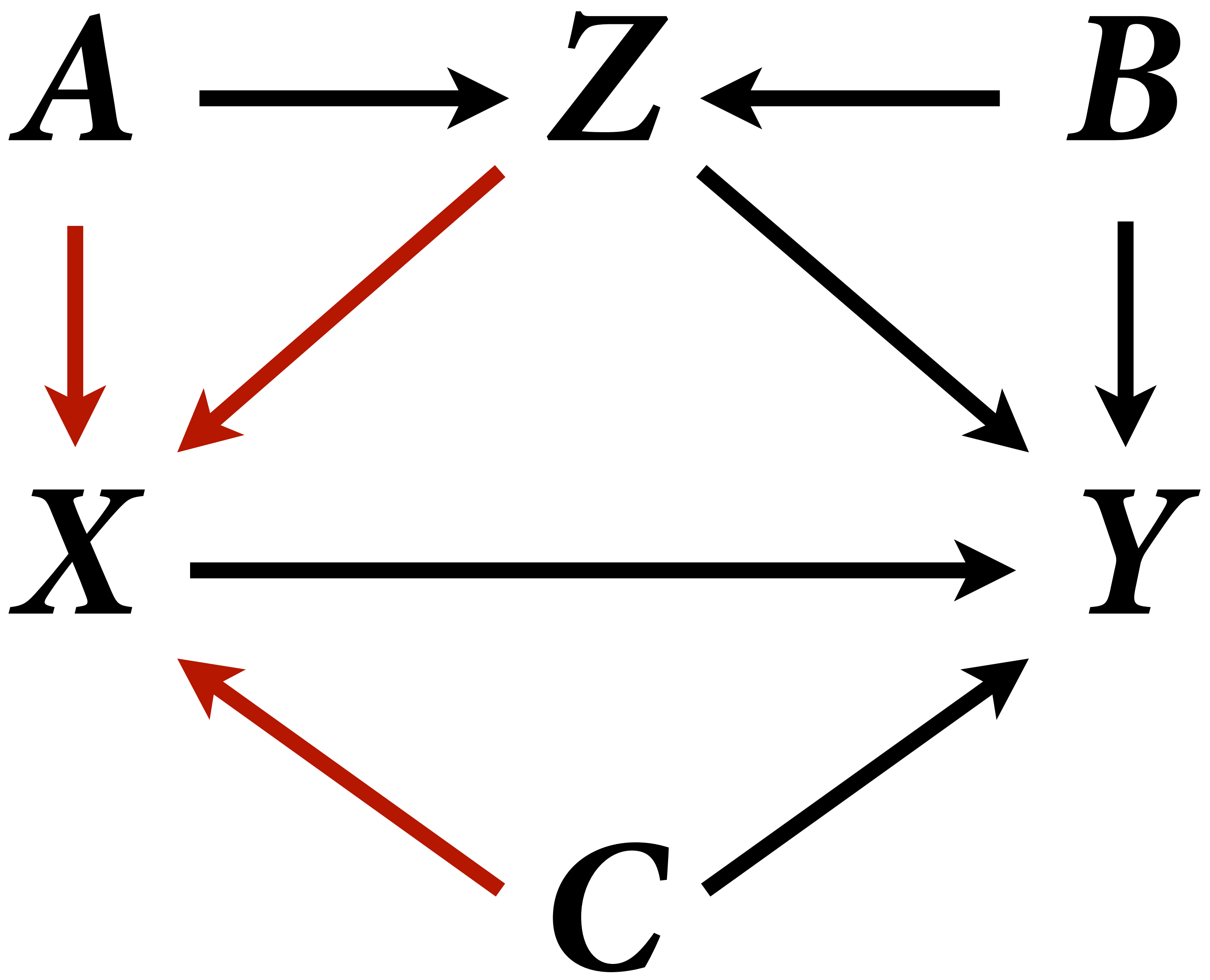

```
> precis(m_YXZ)
       mean   sd  5.5% 94.5%
a     -0.32 0.09 -0.47 -0.18
b_XY  -0.01 0.08 -0.13  0.11
b_Z    0.24 0.11  0.06  0.42
sigma  1.18 0.06  1.08  1.27
```

Coefficient on *Z* means nothing. “Table 2 Fallacy”

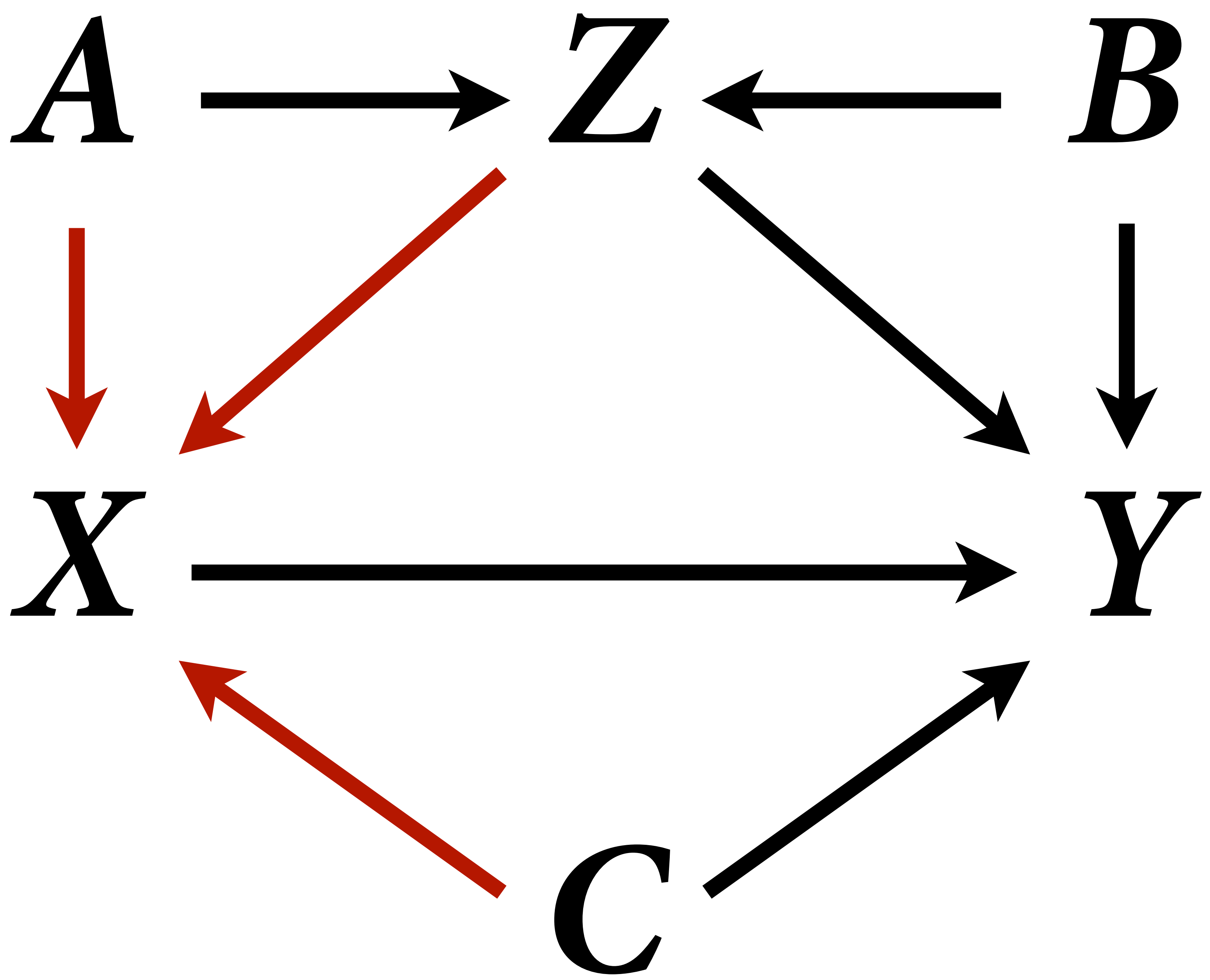
List all the paths connecting **X** and **Y**. Which need to be closed to estimate effect of **X** on **Y**?

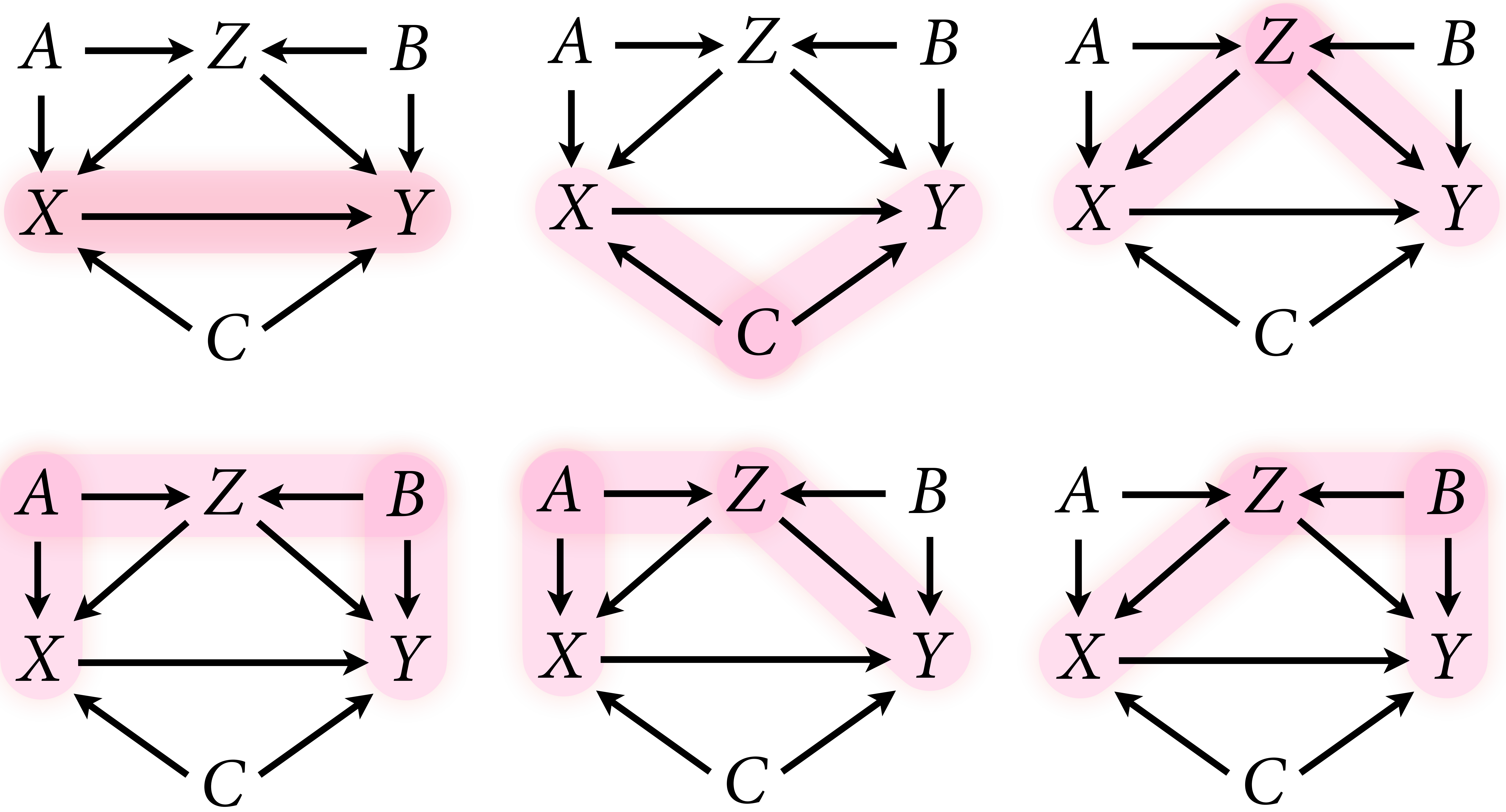


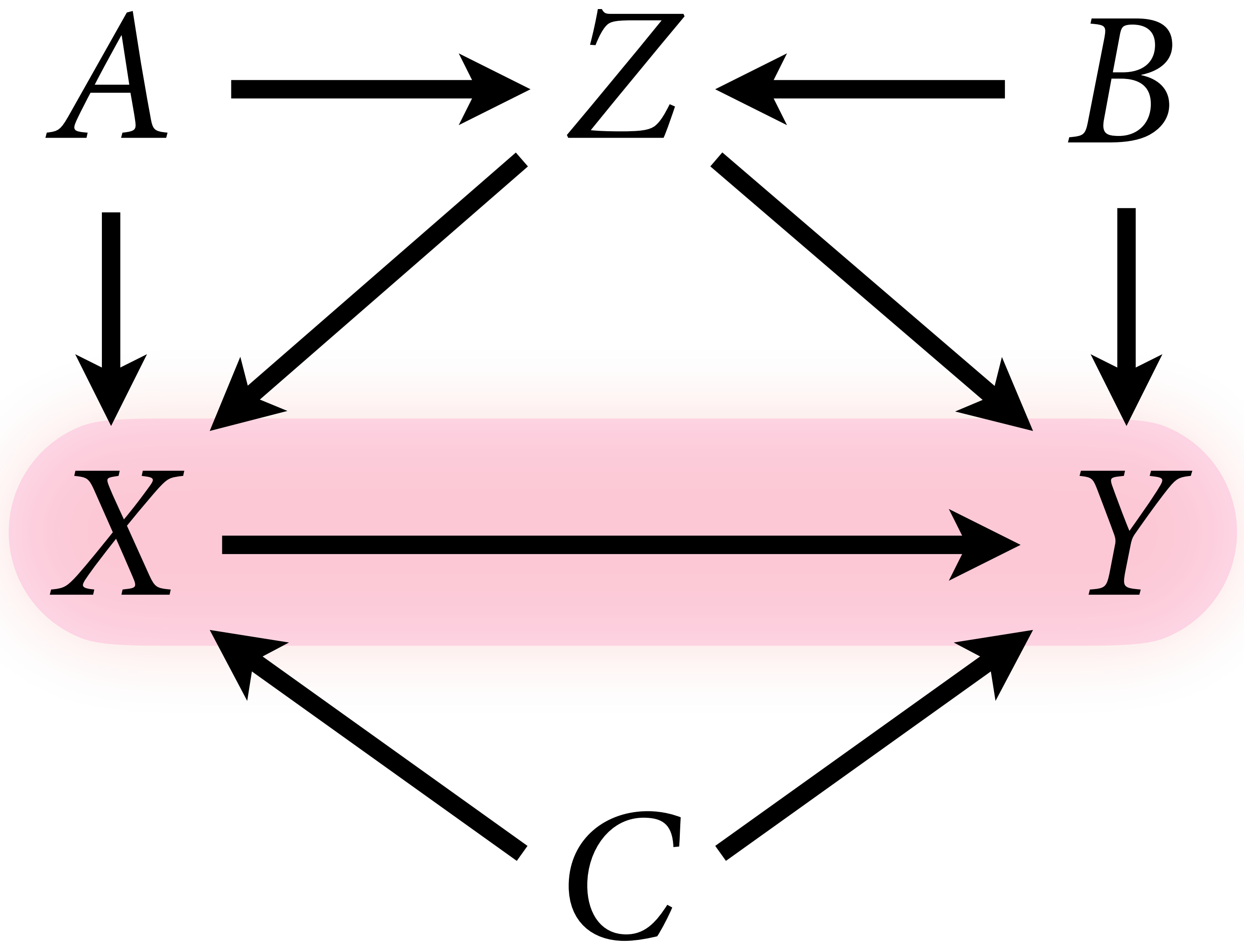

A
Z
B
X
Y
C



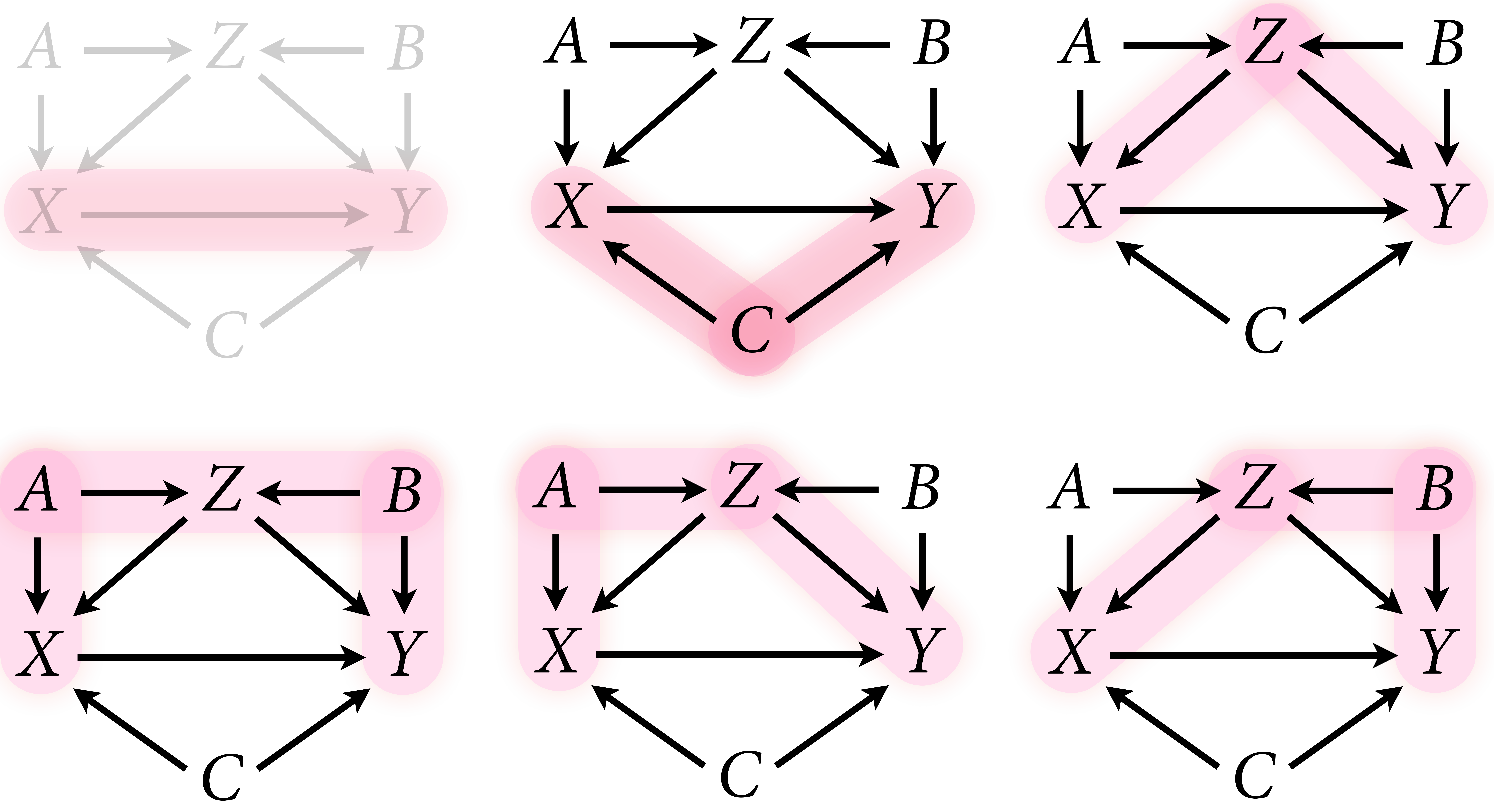
$$P(Y|\mathrm{do}(X))$$

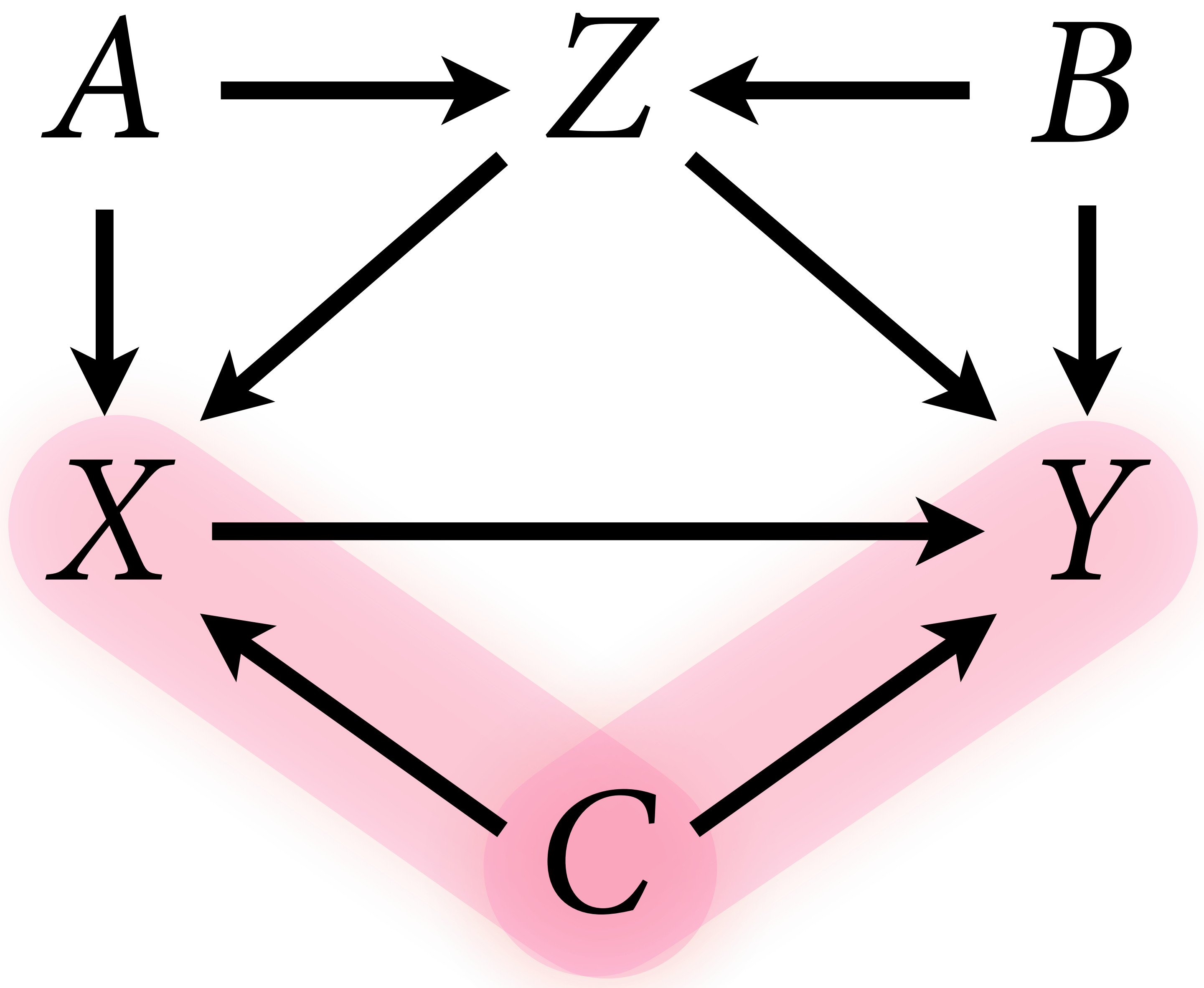



A
Z
B
X
Y
C
A
Z
B
X
Y
C
A
Z
B
X
Y
C
A
Z
B
X
Y
C
A
Z
B
X
Y
C
A
Z
B
X
Y
C



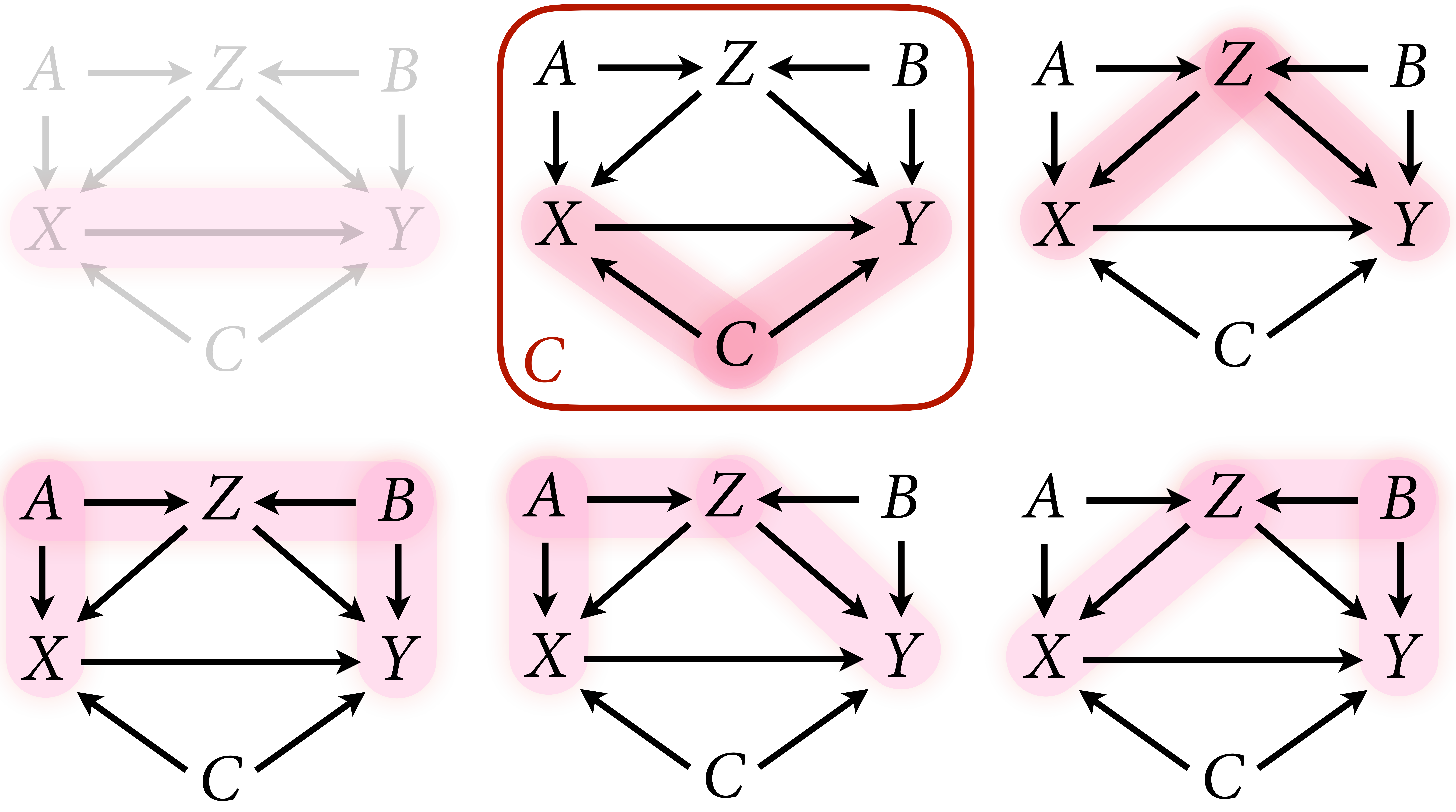
Causal path, open

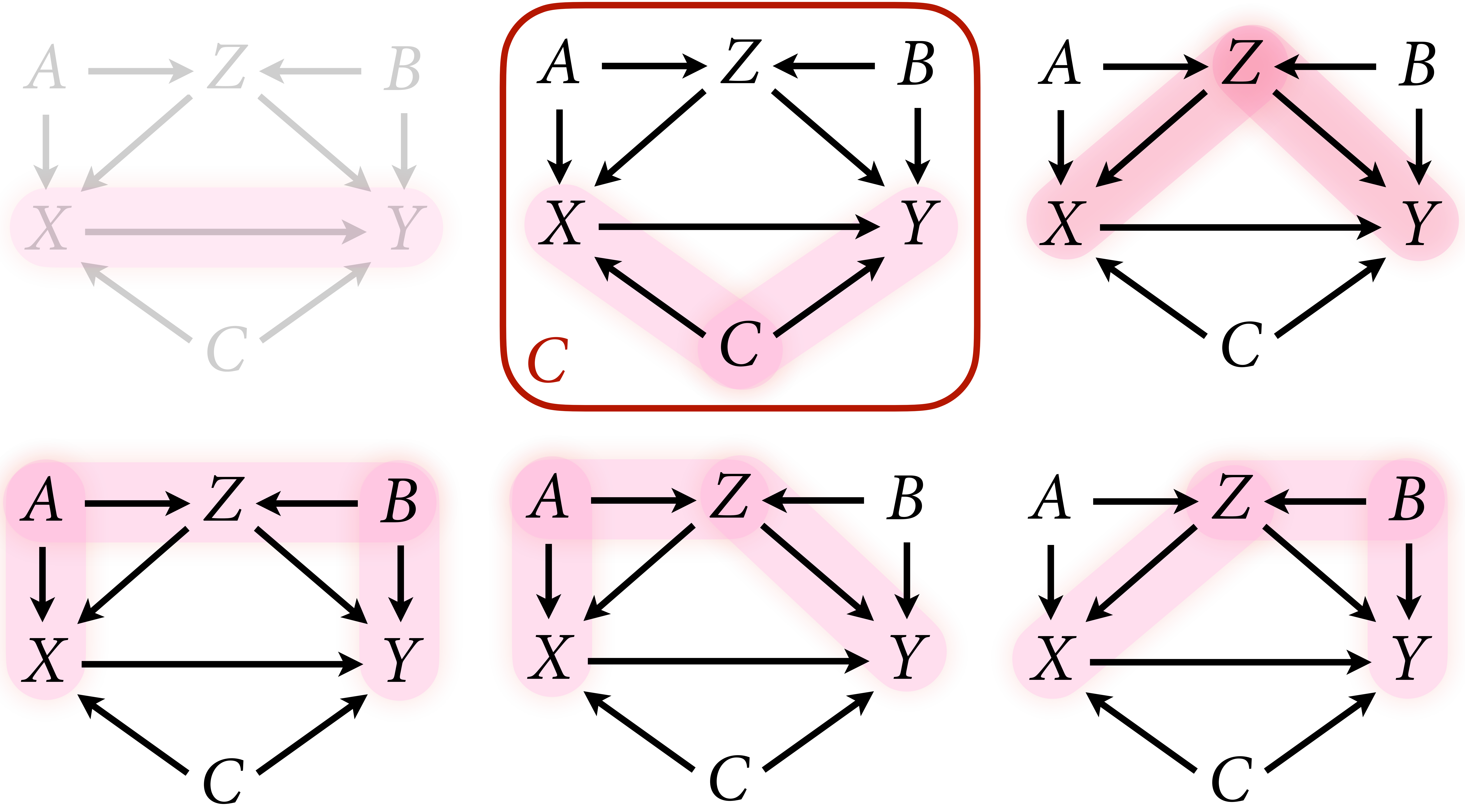

A Z B X Y C
A Z B X Y C
A Z B X Y C
A Z B X Y C
A Z B X Y C
A Z B X Y C

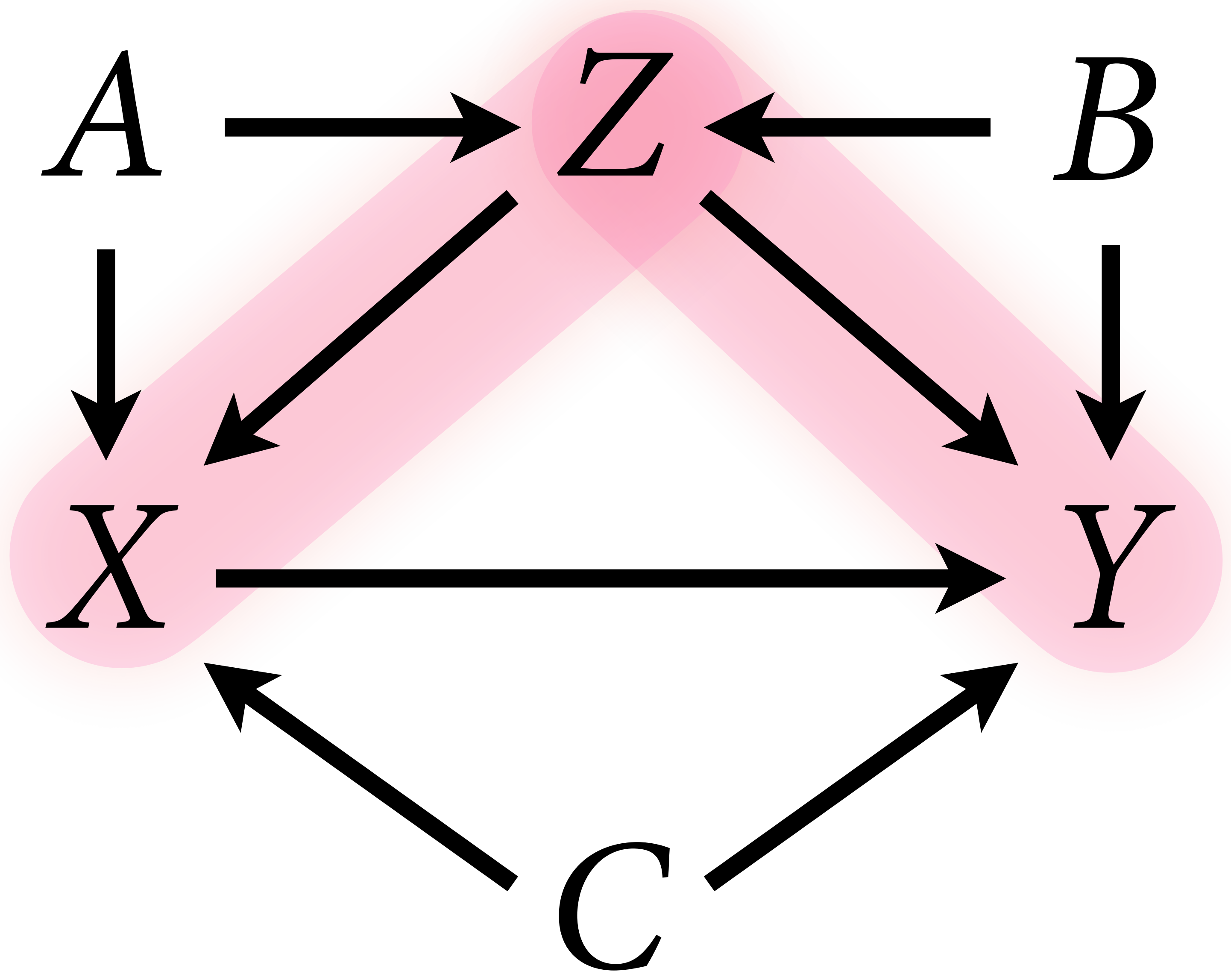


Backdoor path, open

Close with C

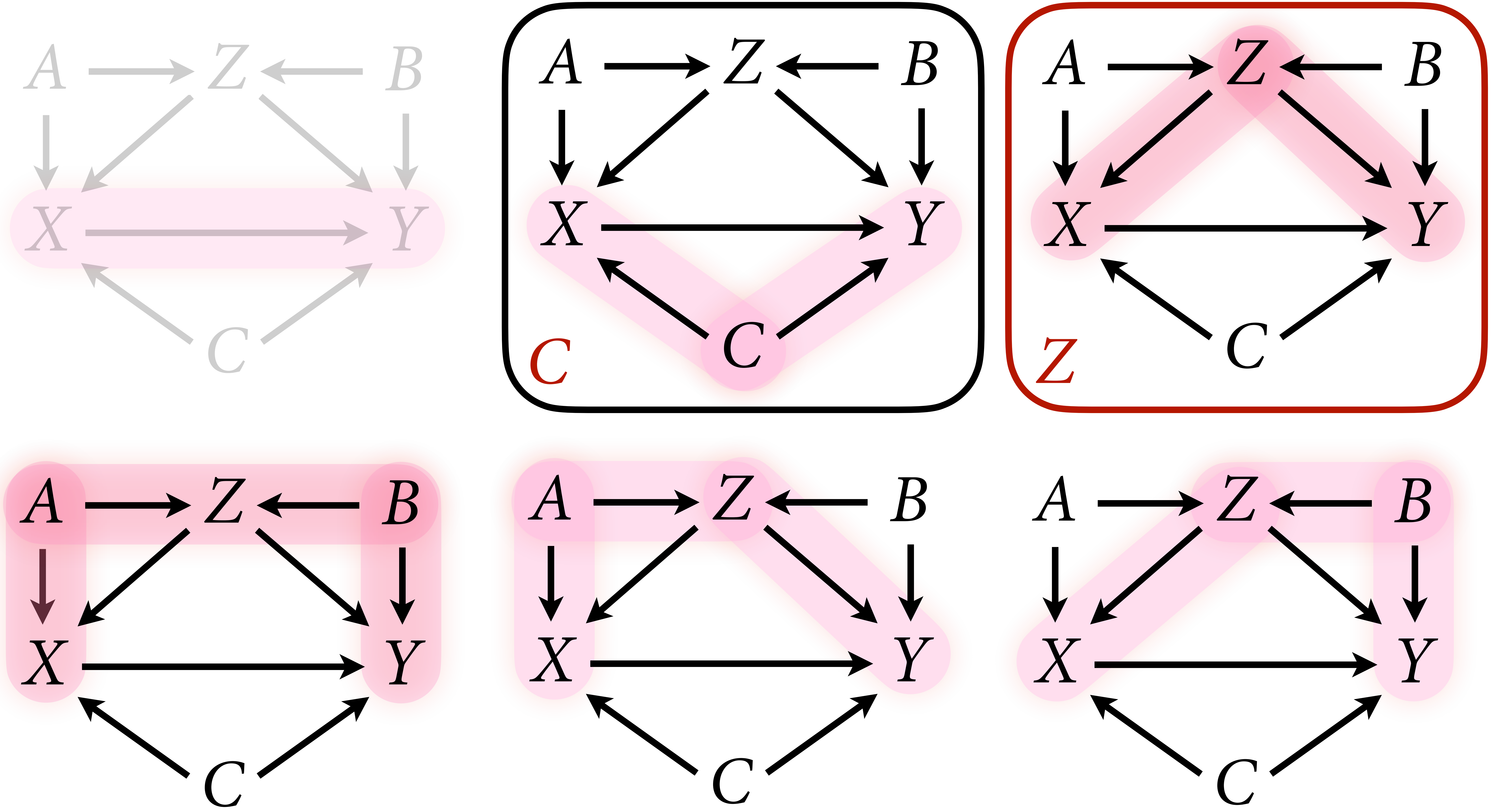

A
Z
B
X
Y
C
A
Z
B
X
Y
C
C
A
Z
B
X
Y
C
A
Z
B
X
Y
C
A
Z
B
X
Y
C
A
Z
B
X
Y
C

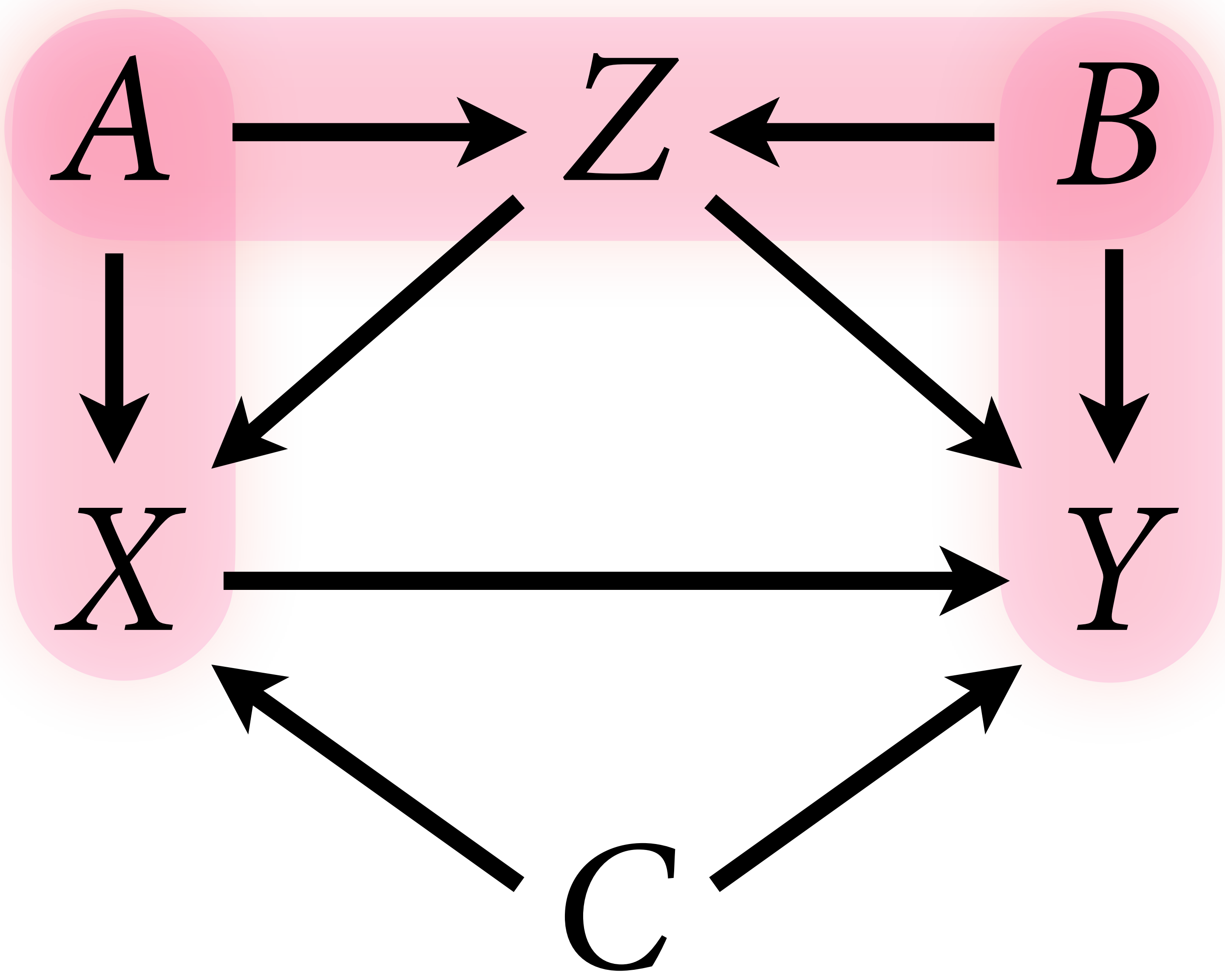

A Z B X Y C
A Z B X Y C
C
A Z B X Y C
A Z B X Y C
A Z B X Y C
A Z B X Y C



Backdoor path, open

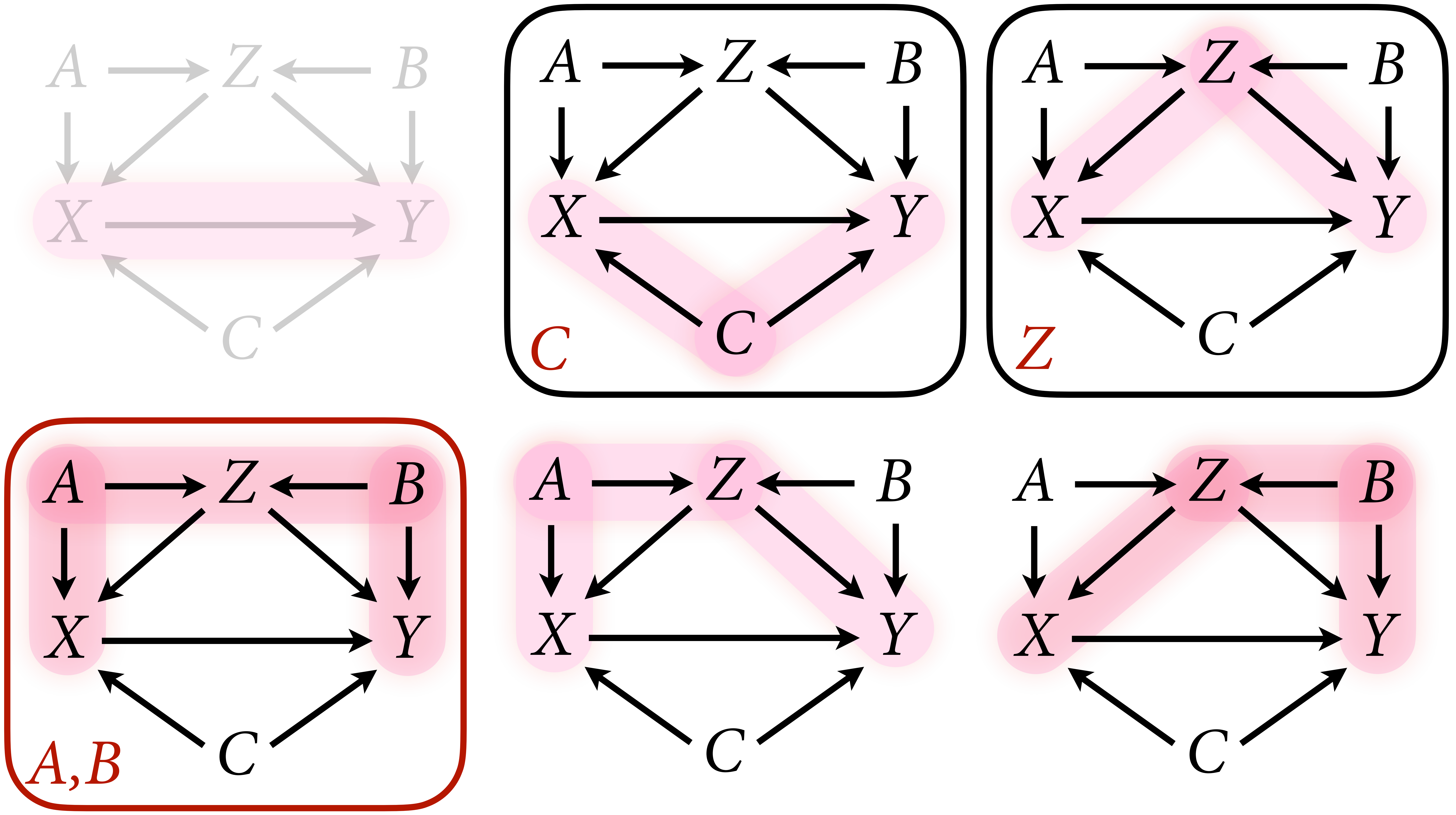
Close with *Z*


A
Z
B
X
Y
C
A
Z
B
X
Y
C
C
A
Z
B
X
Y
C
Z
A
Z
B
X
Y
C
A
Z
B
X
Y
C
A
Z
B
X
Y
C

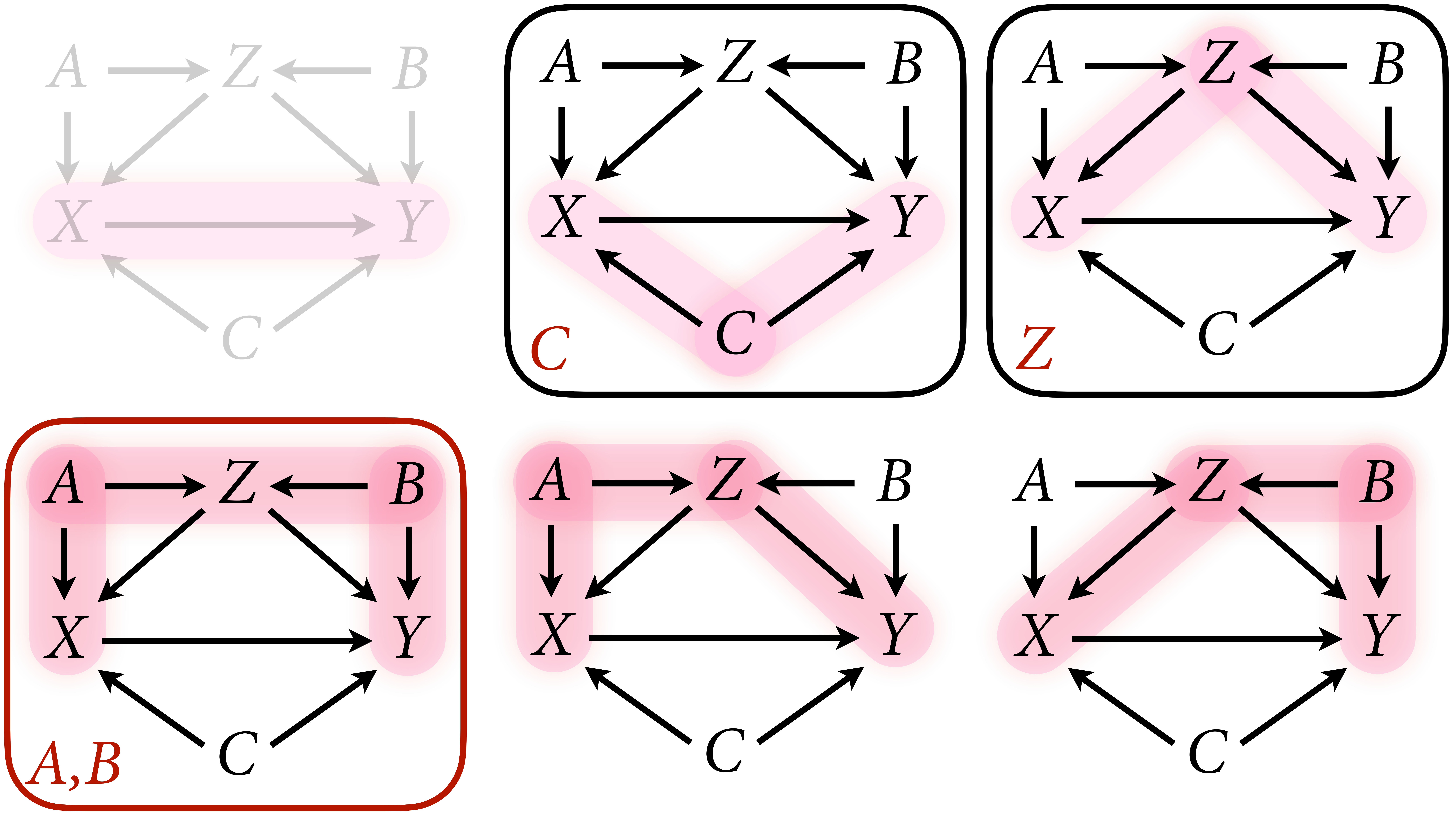


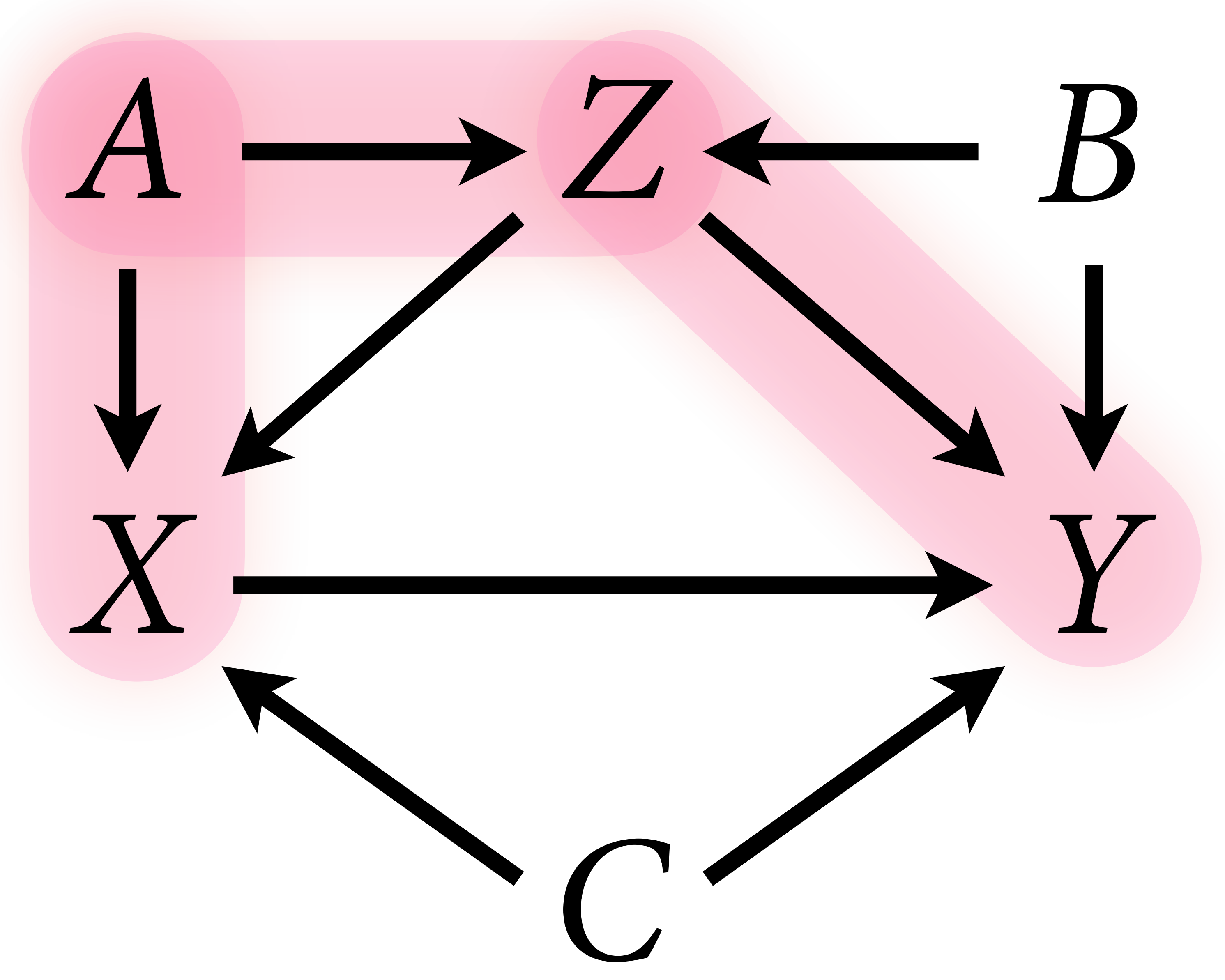
Backdoor path, opened by Z

A or B to close



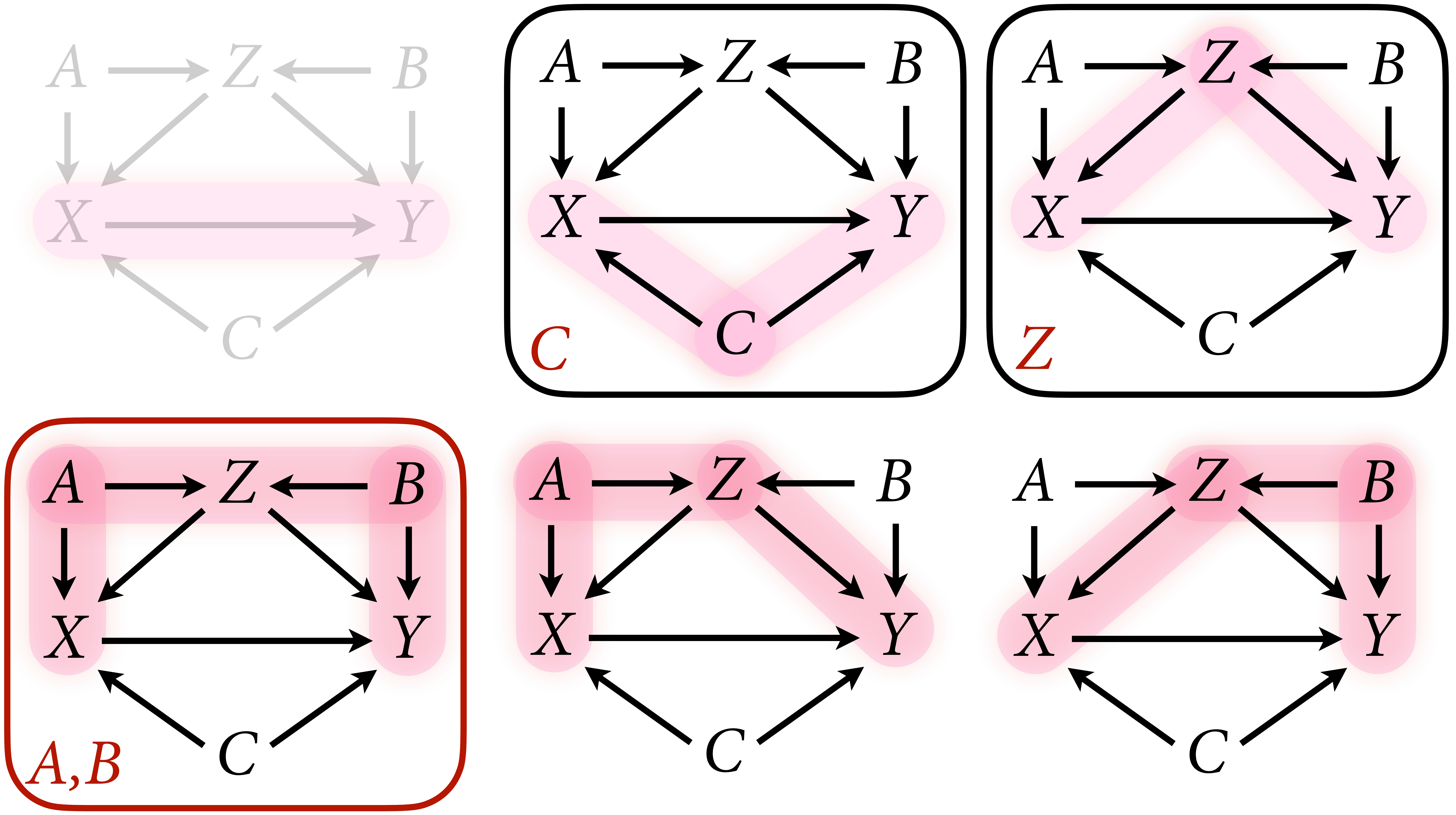
A
Z
B
X
Y
C
A
Z
B
X
Y
C
C
A
Z
B
X
Y
C
Z
A
Z
B
X
Y
C
A,B
A
Z
B
X
Y
C
A
Z
B
X
Y
C

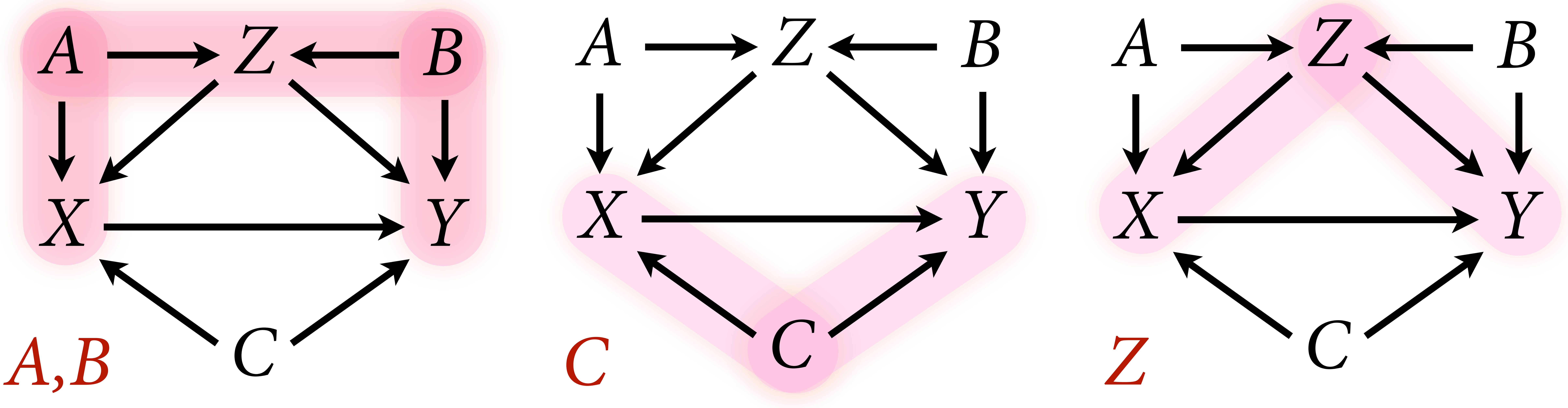

A Z B
X Y
C
A Z B
X Y
C
C
A Z B
X Y
C
Z
A Z B
X Y
C
A,B
A Z B
X Y
C
A Z B
X Y
C



Backdoor path, open

Close with A or Z


A
Z
B
X
Y
C
C
Z
A,B



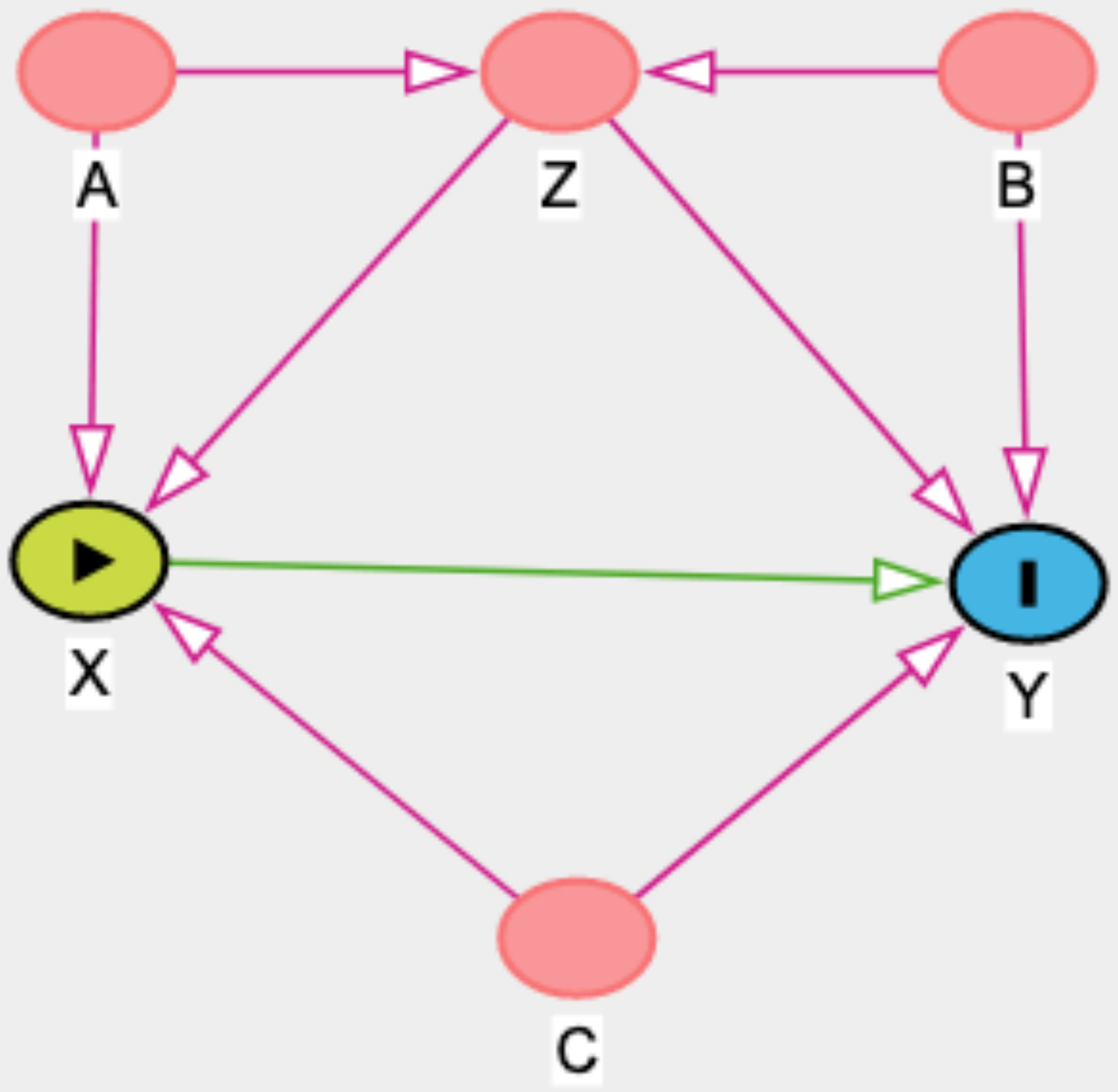
Minimum adjustment set:

C, Z, and either A or B

(*B* is better choice)

www.dagitty.net

Model | Examples | How to ... | Layout | Help



Causal effect identification

Adjustment (total effect)

Minimal sufficient adjustment sets for estimating the total effect of X on Y:

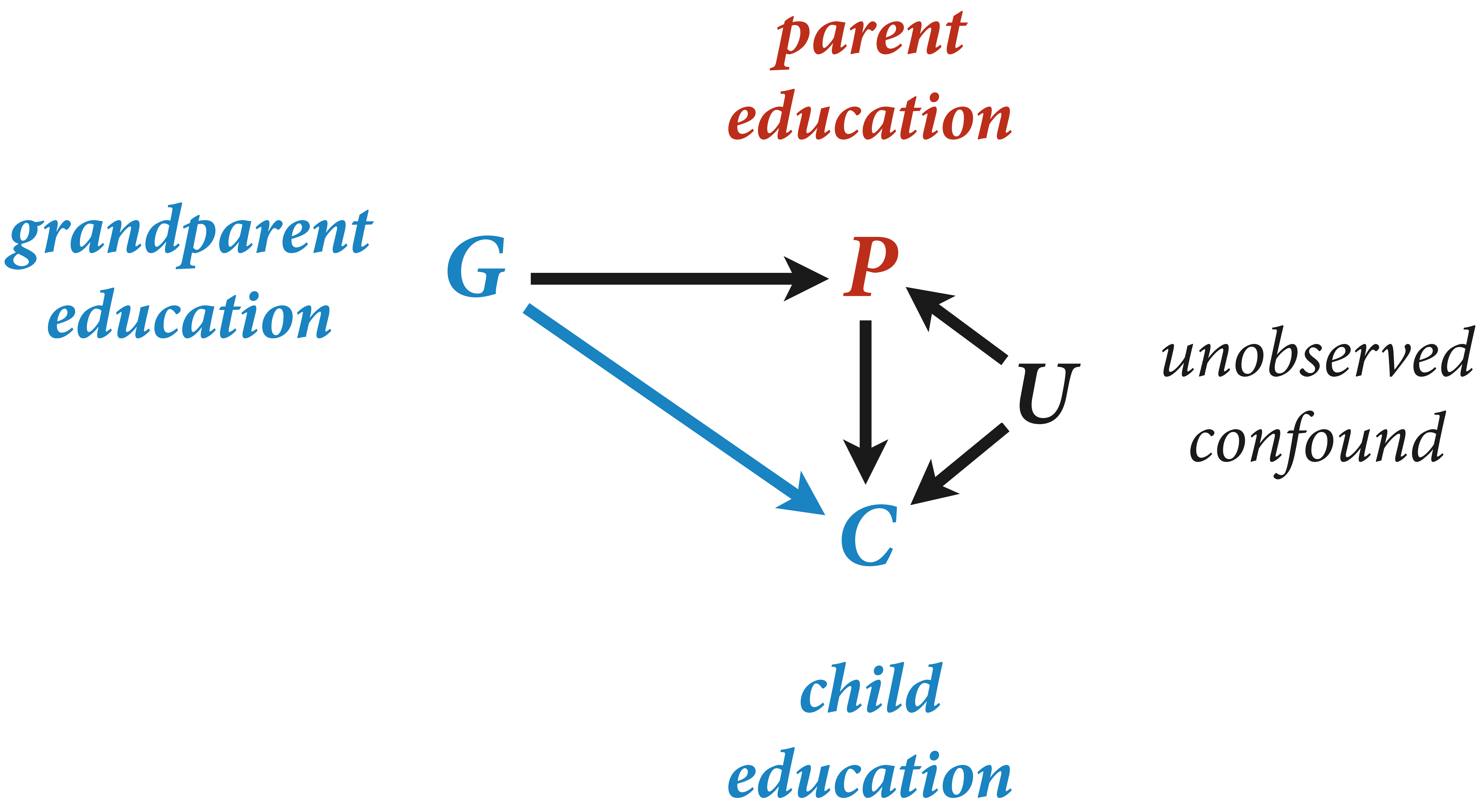
- A, C, Z
- B, C, Z

Testable implications

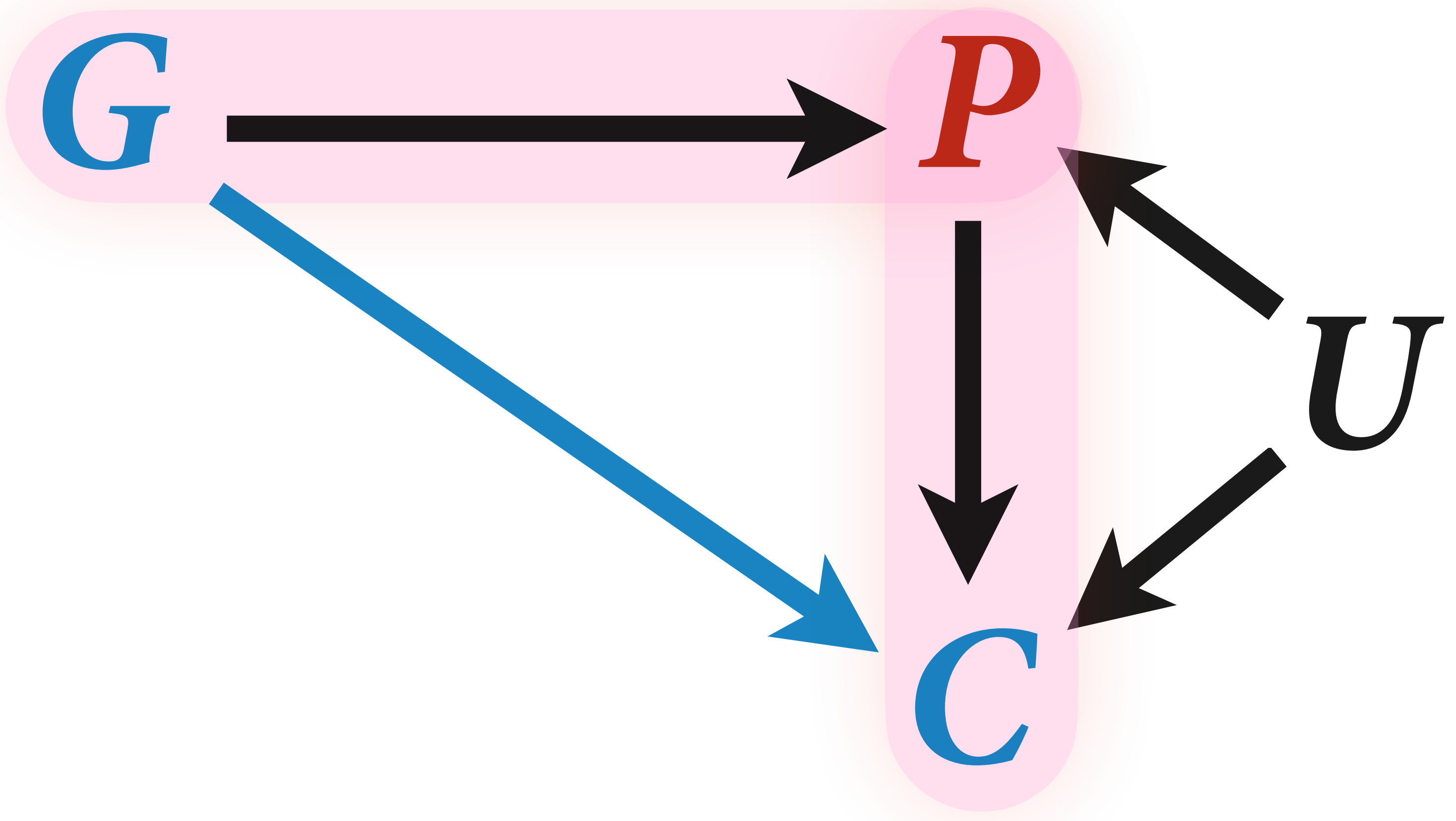
The model implies the following conditional independences:

- $X \perp B \mid A, Z$
- $Y \perp A \mid B, C, X, Z$
- $A \perp B$
- $A \perp C$
- $B \perp C$
- $Z \perp C$

Export R code

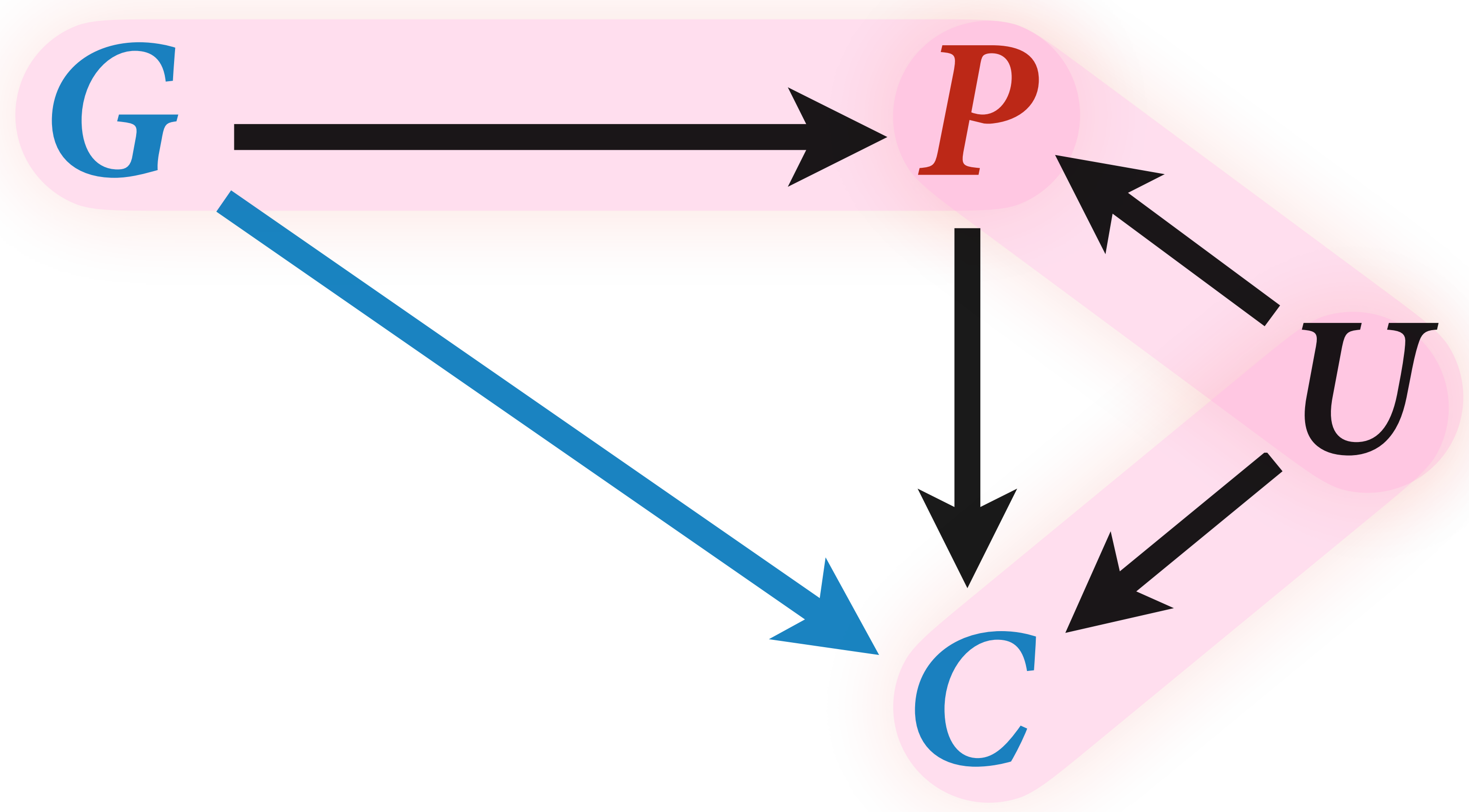

parent education
grandparent education
G
P
U
unobserved confound
C
child education

Pipe: G –> P –> C



P is a mediator

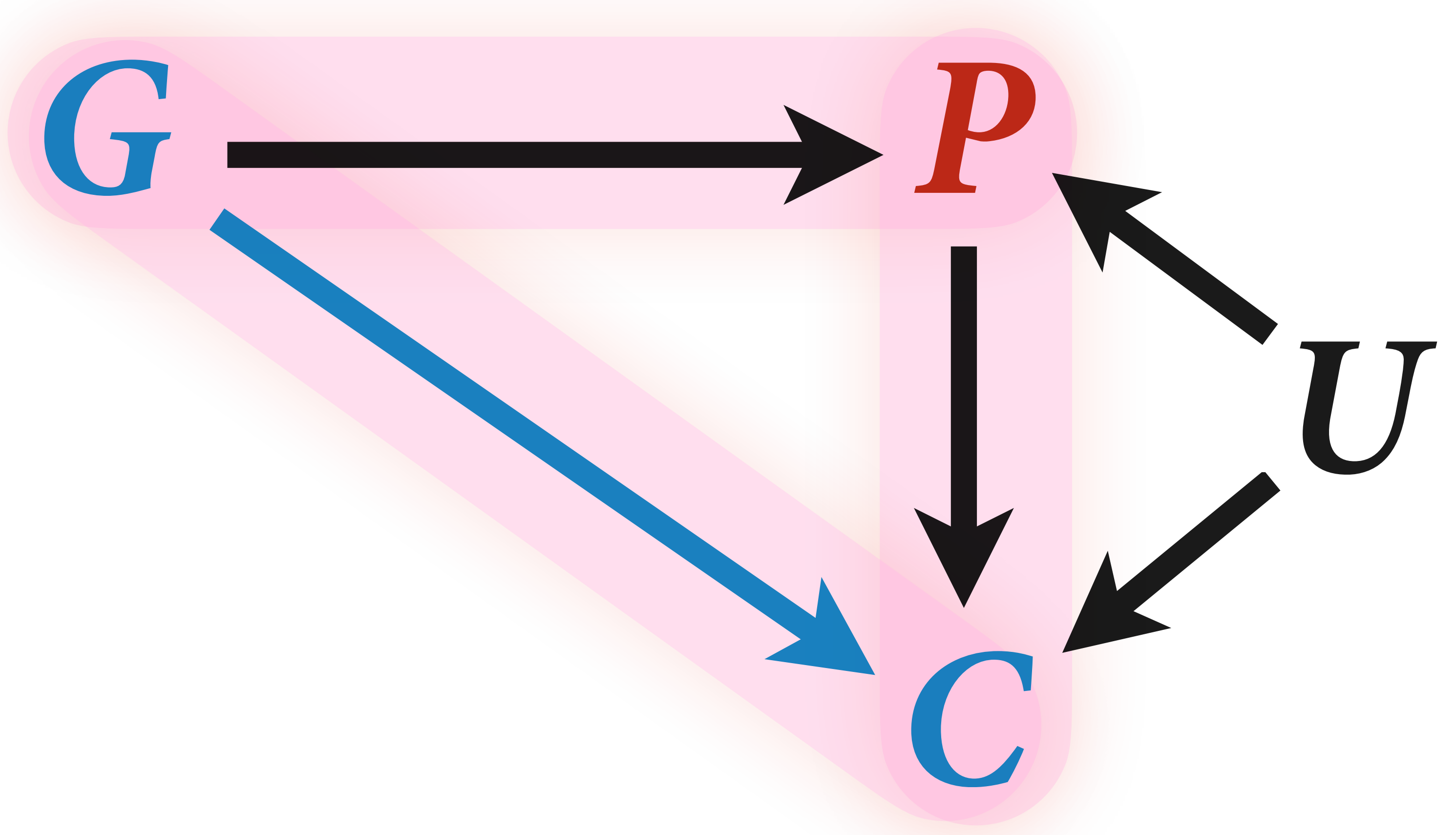
Pipe: G –> P –> C



P is a collider

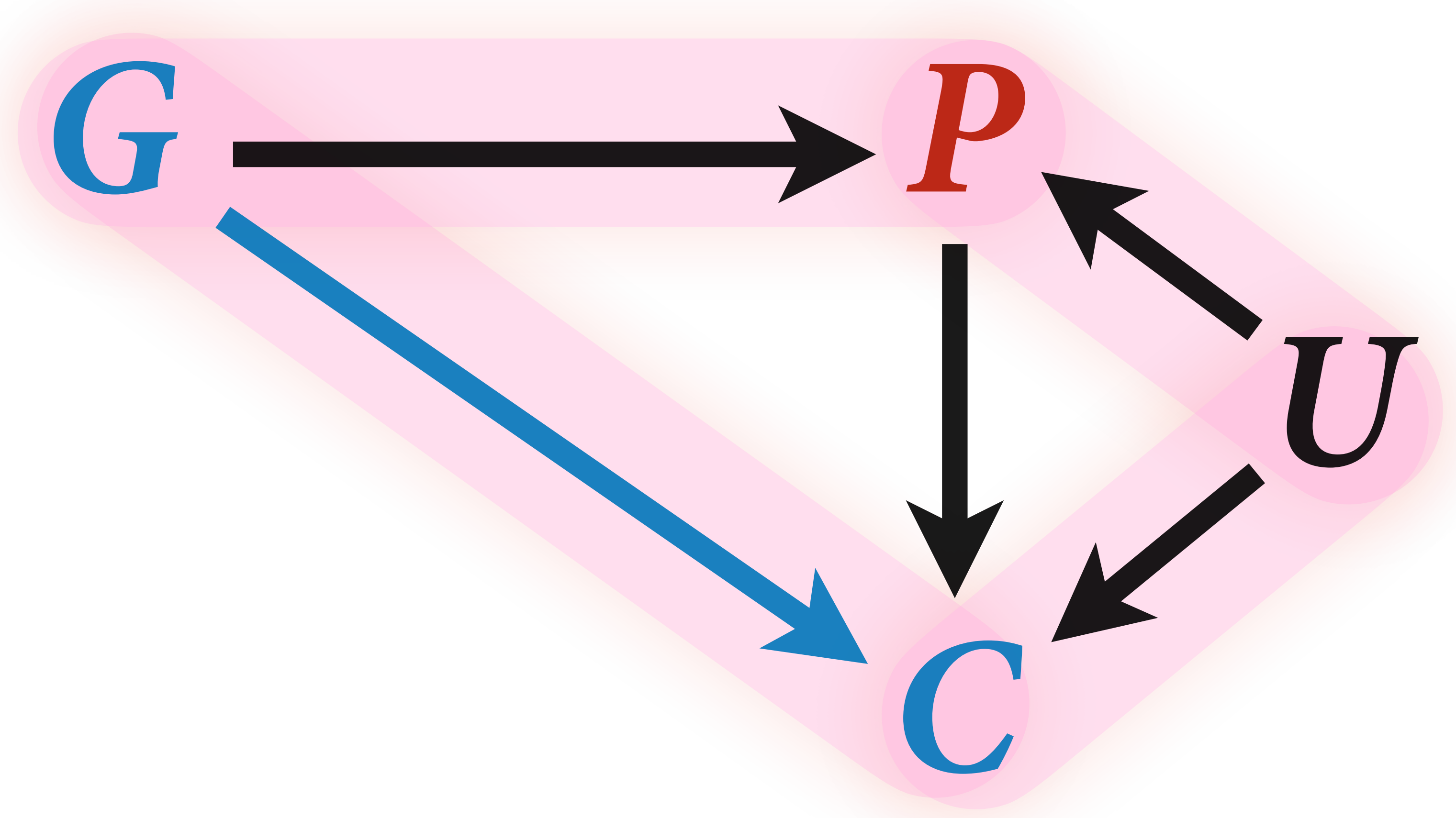
Fork: C <– U –> P

Can estimate **total** effect of G on C



$$C_i \sim \text{Normal}(\mu_i, \sigma)$$
$$\mu_i = \alpha + \beta_G G_i$$

Cannot estimate **direct** effect



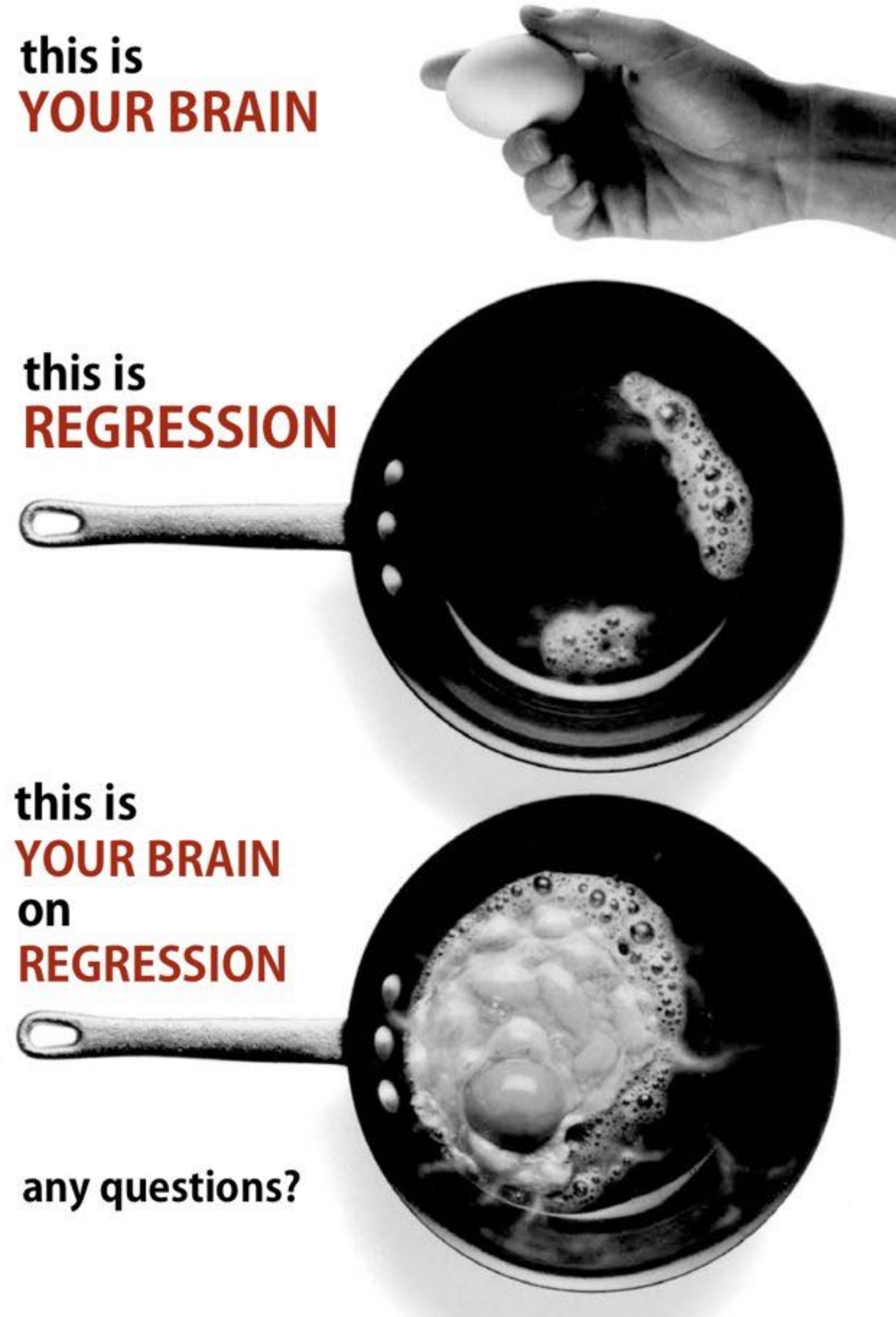
$$C_i \sim \text{Normal}(\mu_i, \sigma)$$
$$\mu_i = \alpha + \beta_G G_i + \beta_P P_i$$

Backdoor Criterion

do-calc more than backdoors & adjustment sets

Full Luxury Bayes: use all variables, but in separate sub-models instead of single regression

do-calc less demanding: finds relevant variables; saves us having to make some assumptions; not always a regression



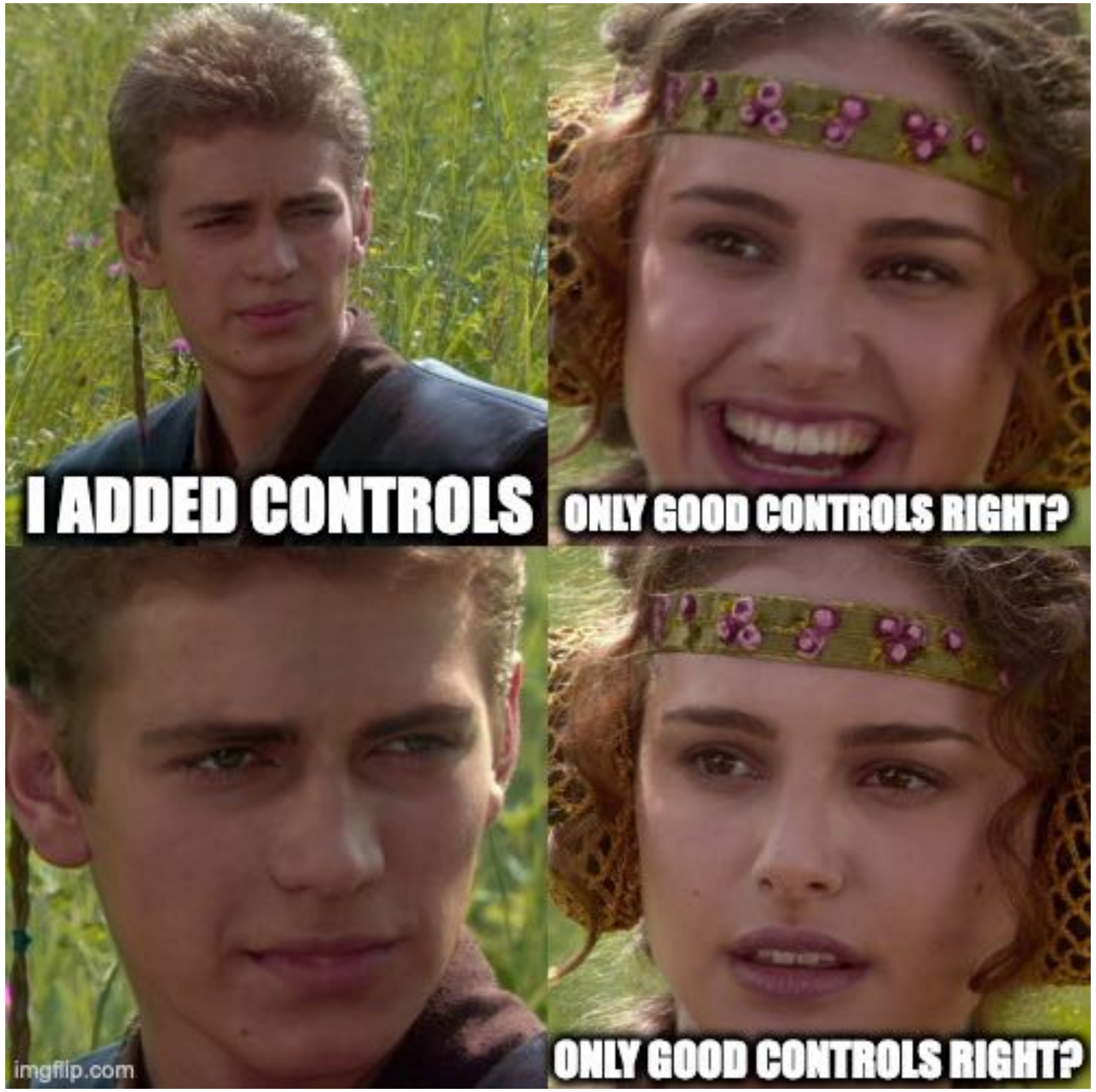
PAUSE

Good & Bad Controls

"Control" variable: Variable introduced to an analysis so that a causal estimate is possible

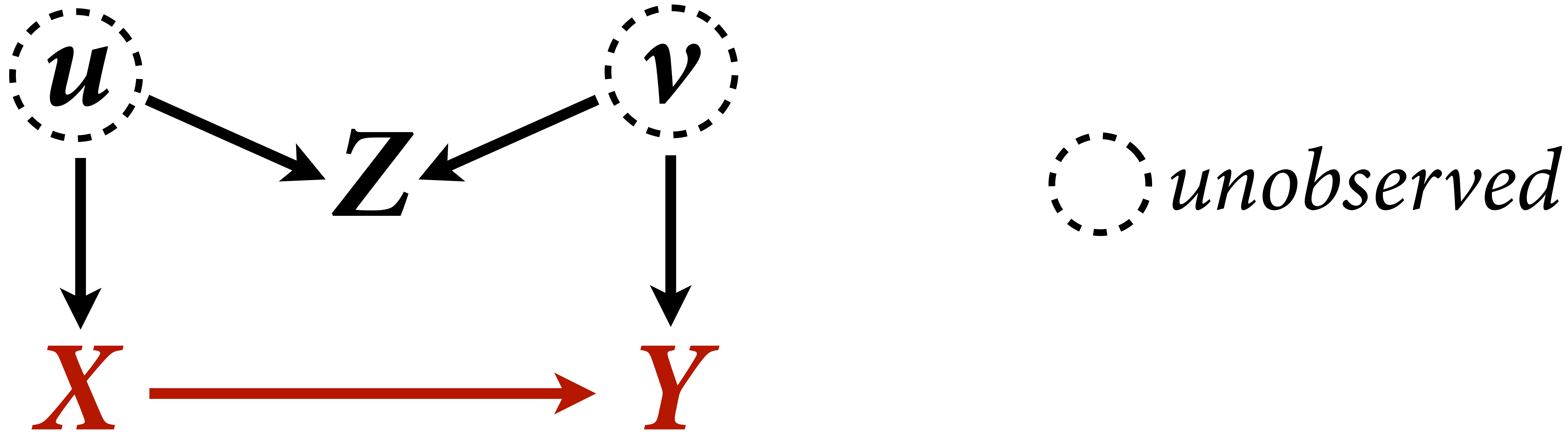
Common **wrong** heuristics for choosing control variables

- Anything in the spreadsheet **YOLO**!
- Any variables not highly **collinear**
- Any **pre-treatment** measurement (baseline)

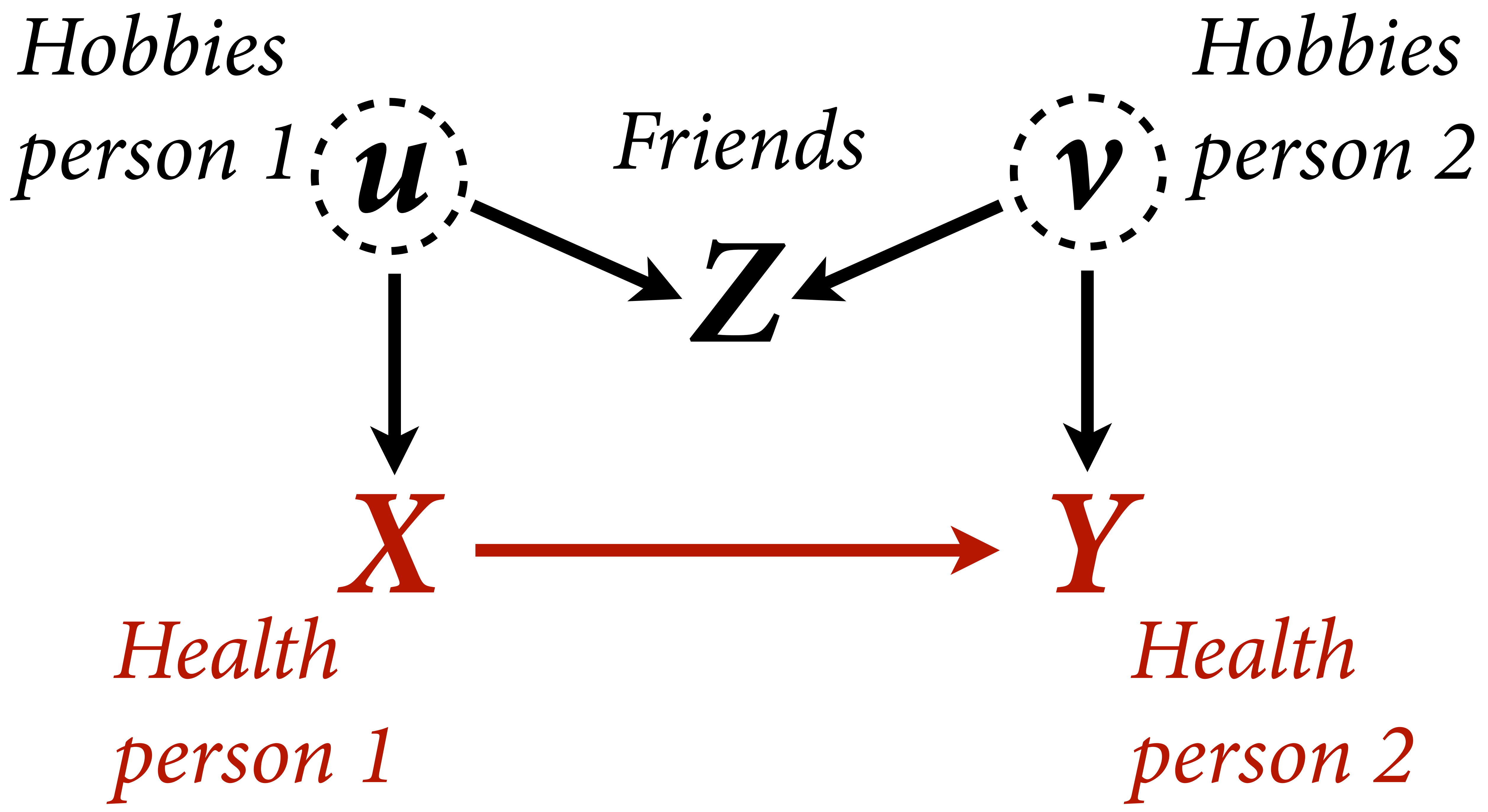


$$X \longrightarrow Y$$

Cinelli, Forney, Pearl 2021 A Crash Course in Good and Bad Controls

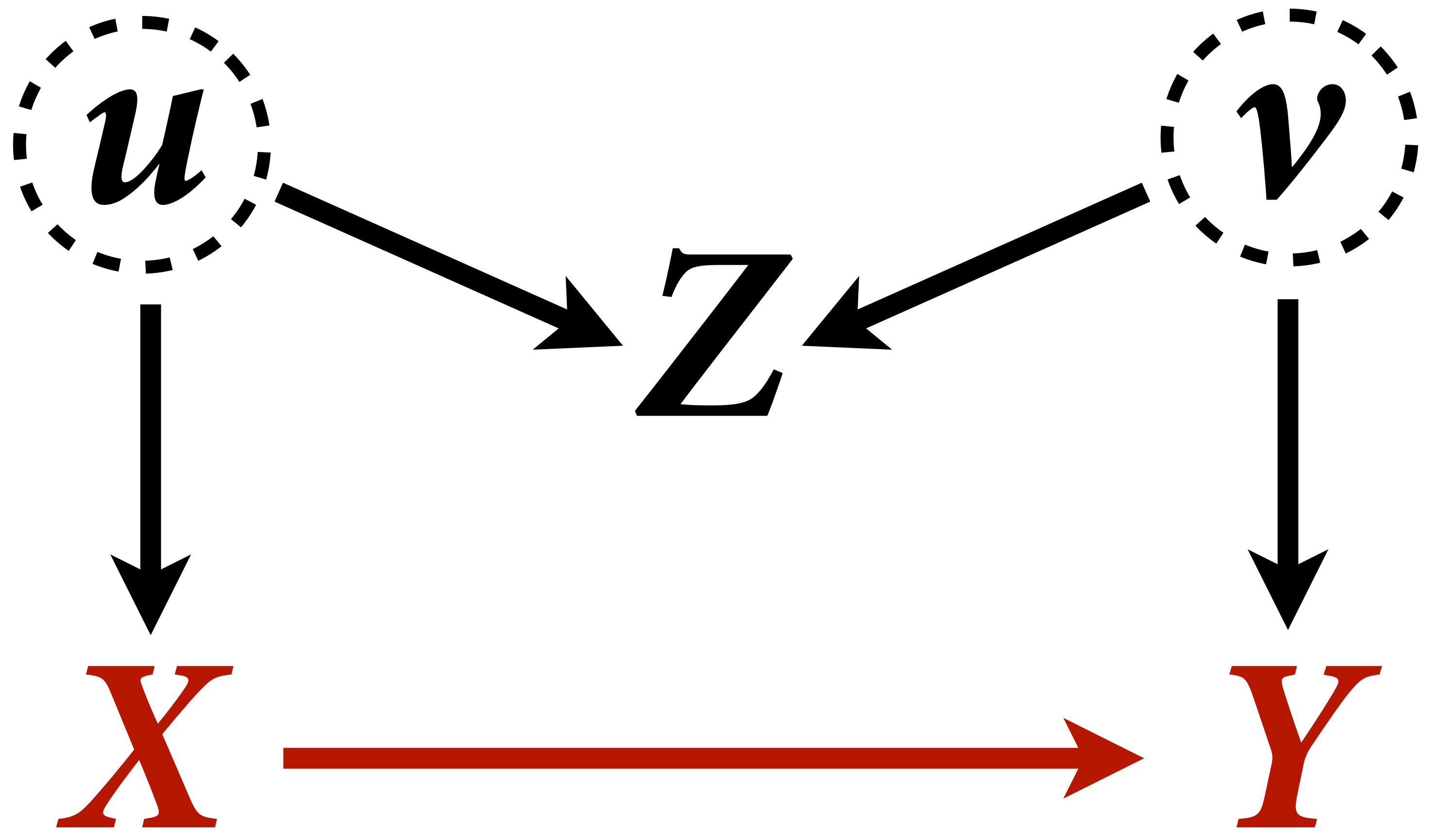


Cinelli, Forney, Pearl 2021 A Crash Course in Good and Bad Controls



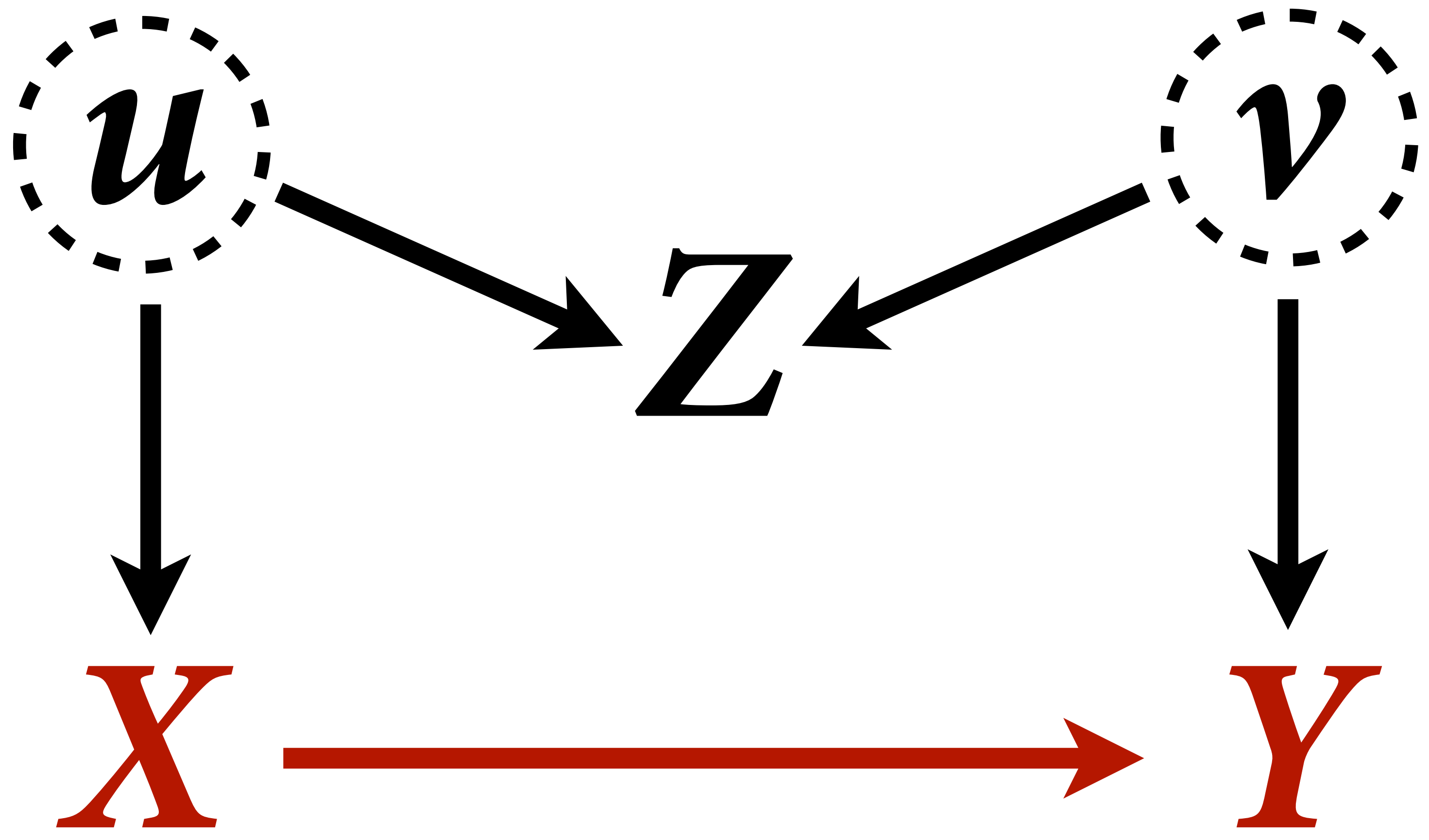
Cinelli, Forney, Pearl 2021 A Crash Course in Good and Bad Controls

(1) List the paths



(1) List the paths

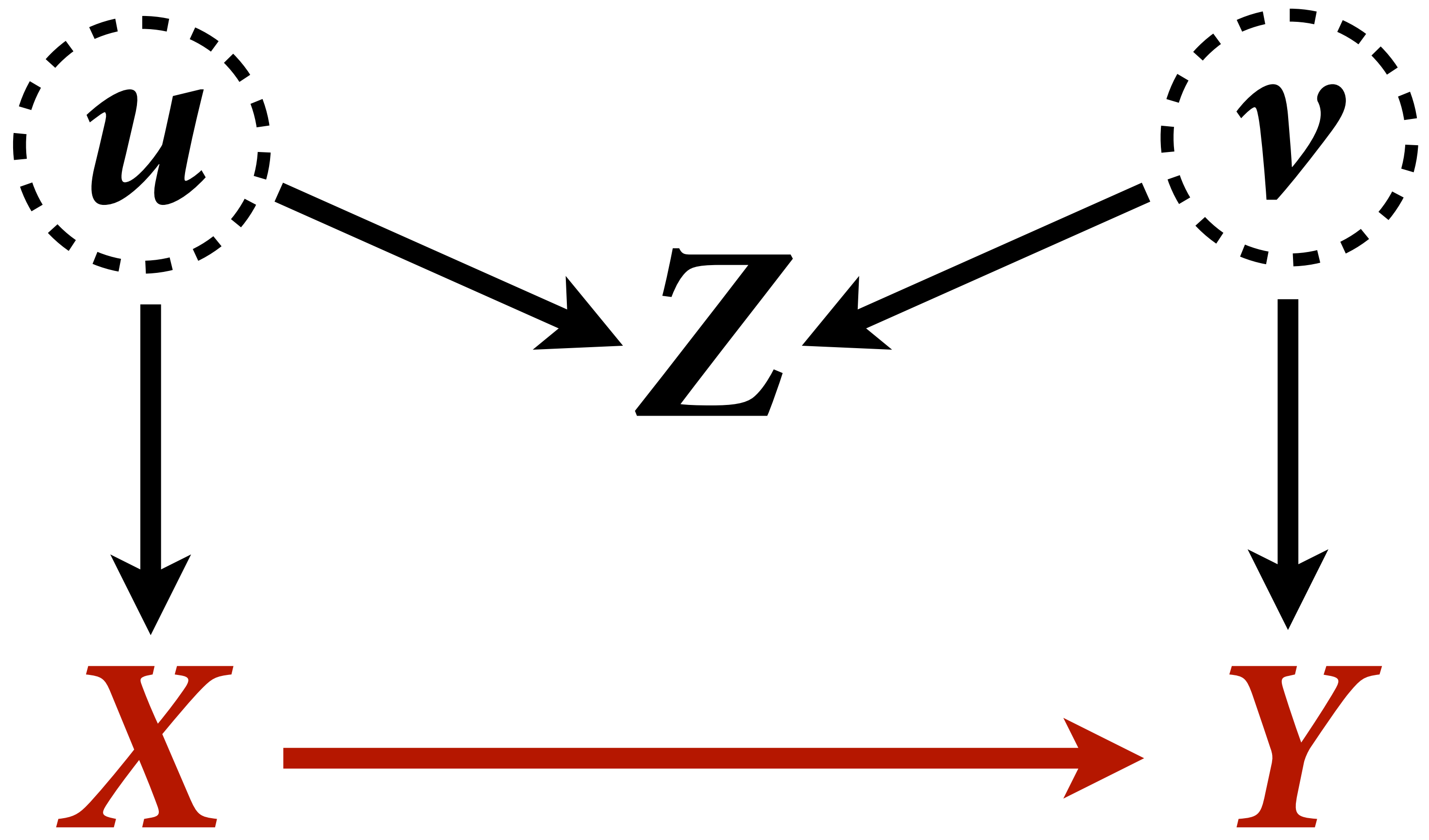
$$X \rightarrow Y$$



(1) List the paths

$$X \rightarrow Y$$

$$X \leftarrow u \rightarrow Z \leftarrow v \rightarrow Y$$



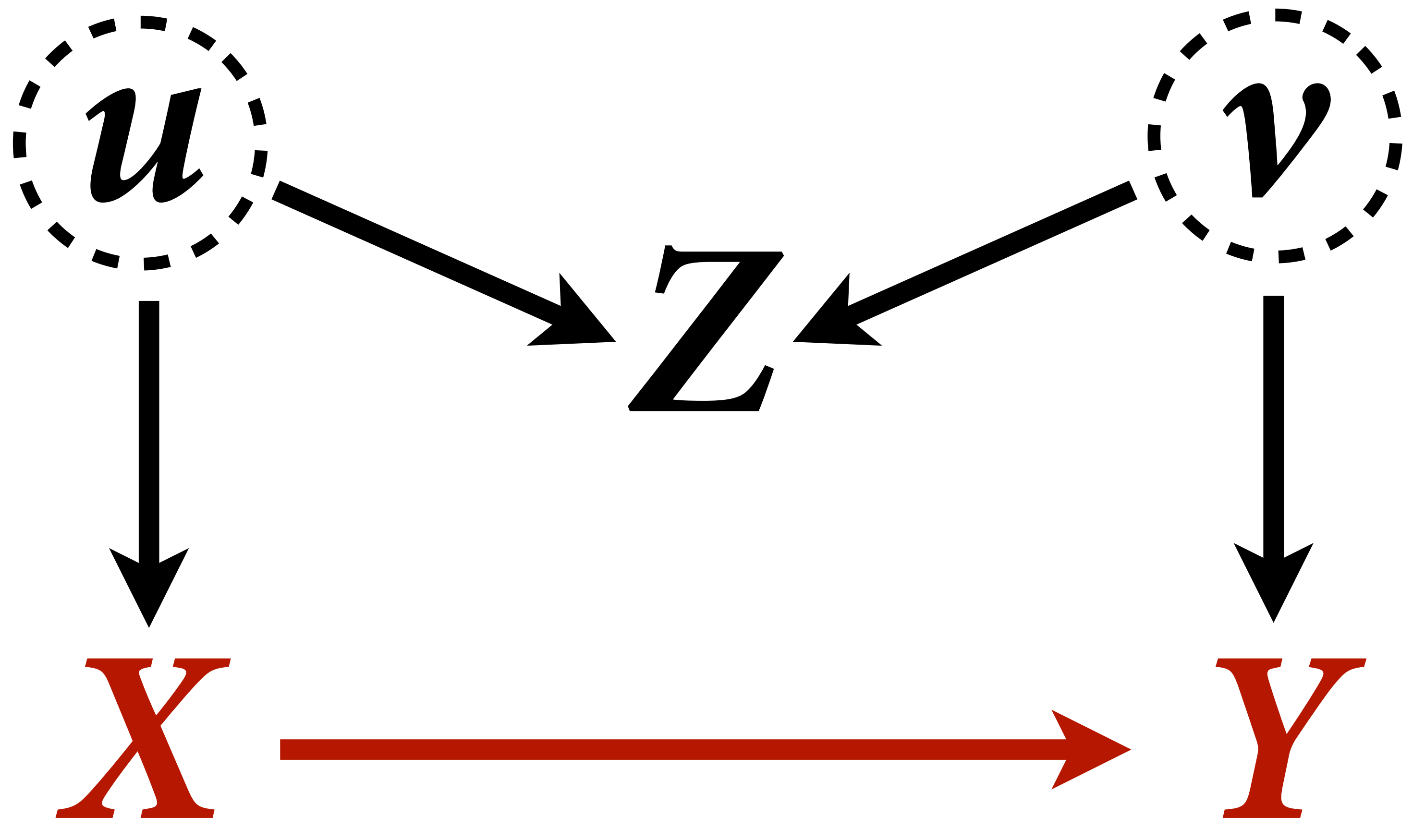
(1) List the paths (2) Find backdoors

$$X \rightarrow Y$$

frontdoor & open

$$X \leftarrow u \rightarrow Z \leftarrow v \rightarrow Y$$

backdoor & closed



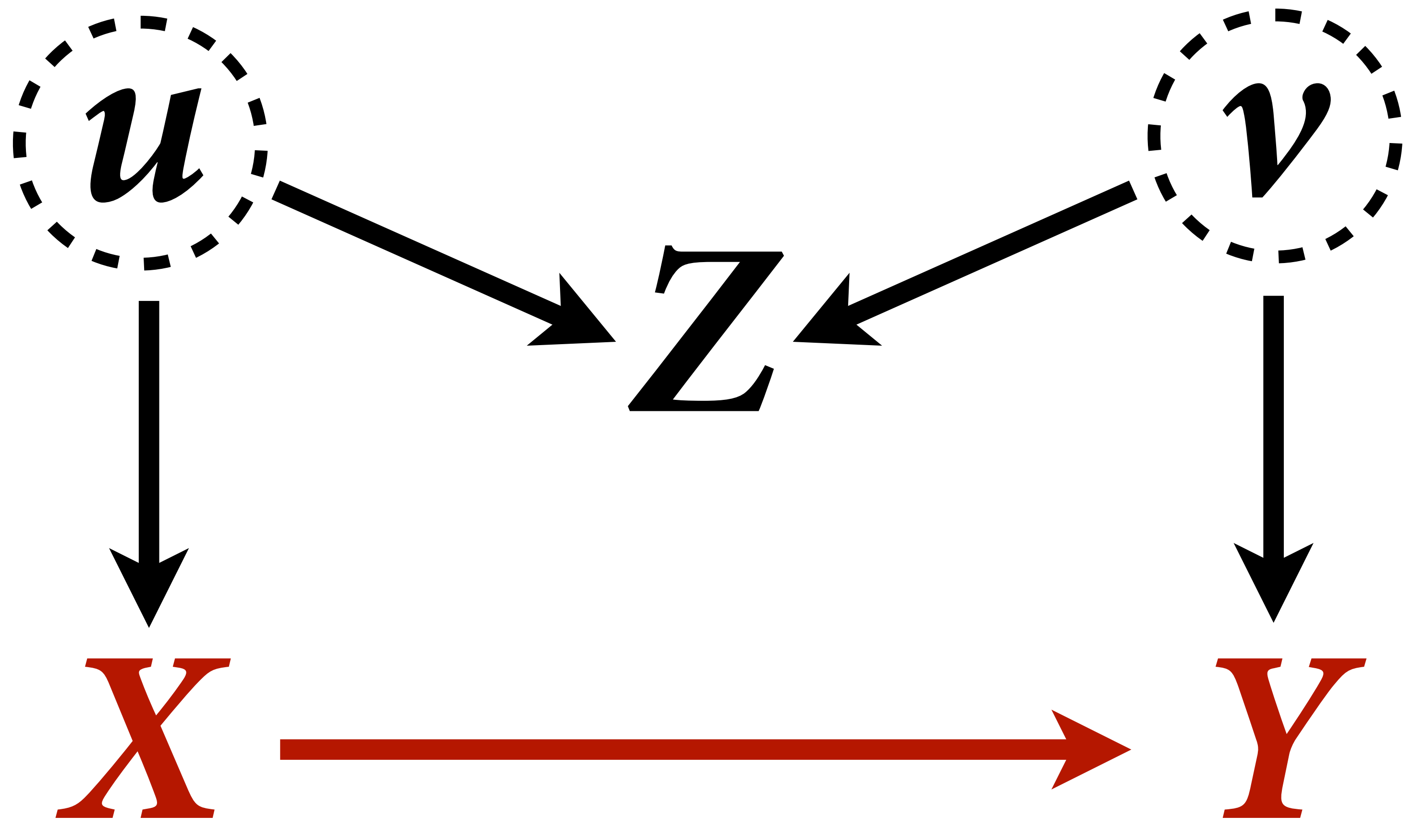
(1) List the paths (2) Find backdoors

$X \rightarrow Y$

frontdoor & open

$X \leftarrow u \rightarrow Z \leftarrow v \rightarrow Y$

backdoor & closed



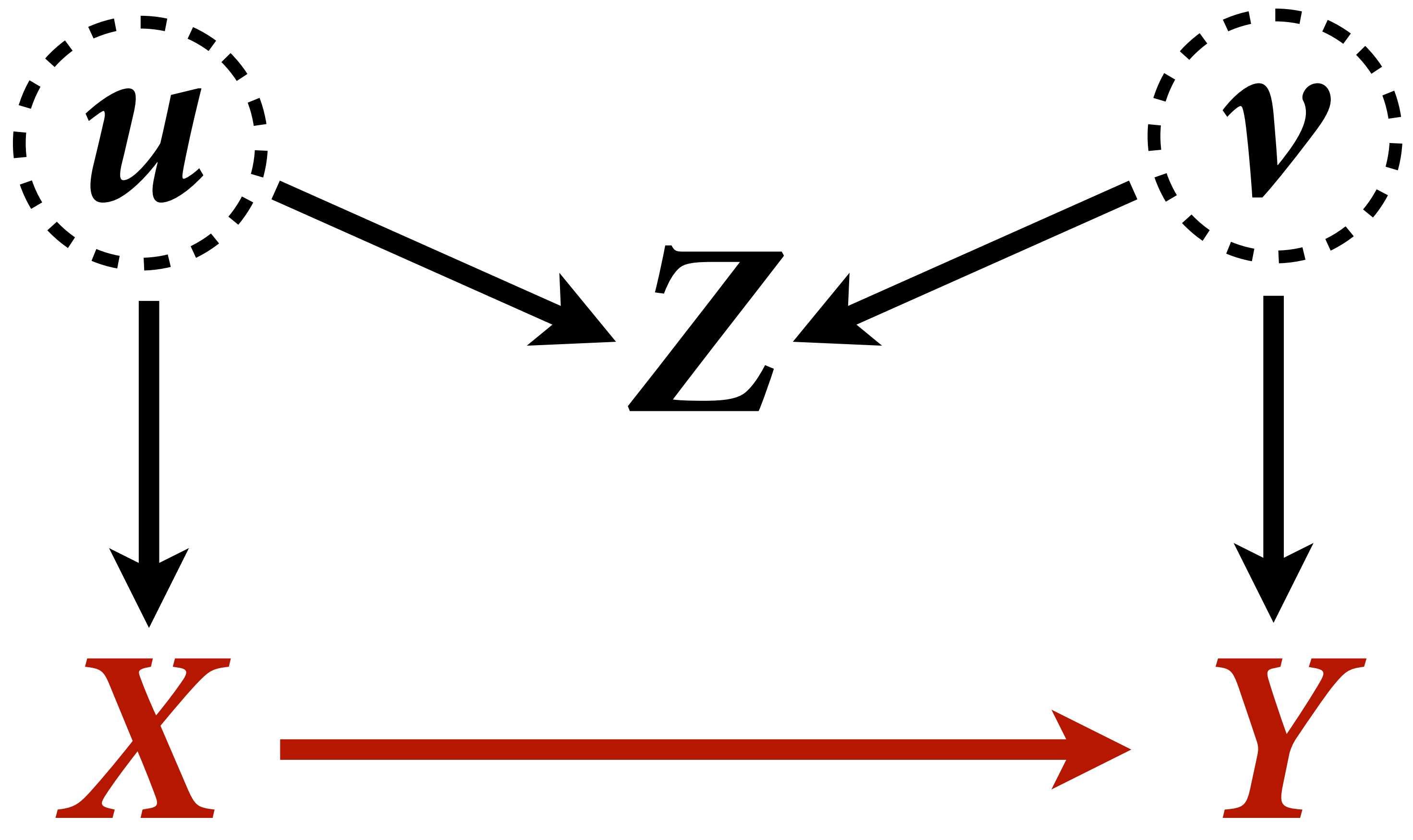
(1) List the paths (2) Find backdoors (3) Close backdoors

$X \rightarrow Y$

frontdoor & open

$X \leftarrow u \rightarrow Z \leftarrow v \rightarrow Y$

backdoor & closed

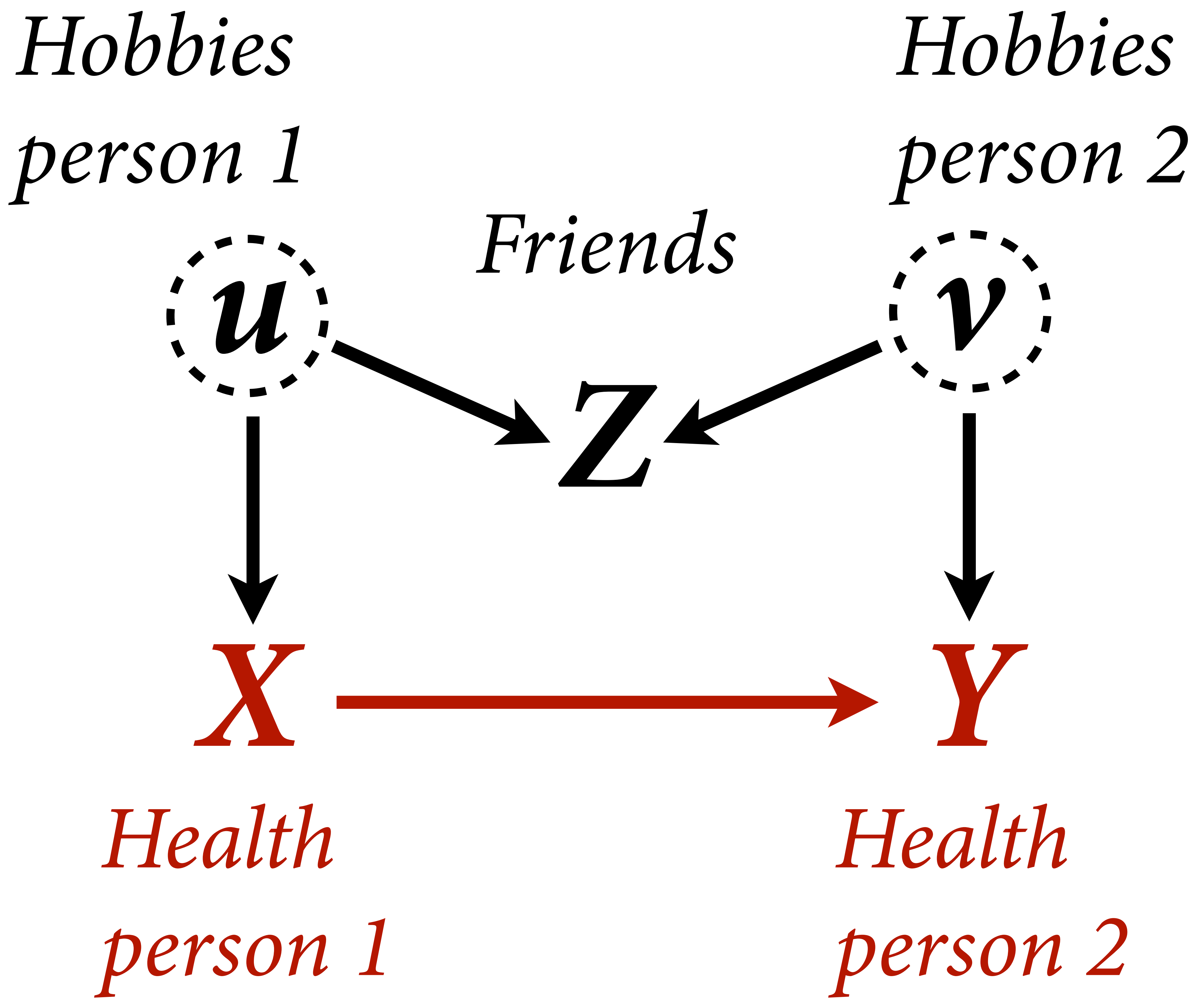


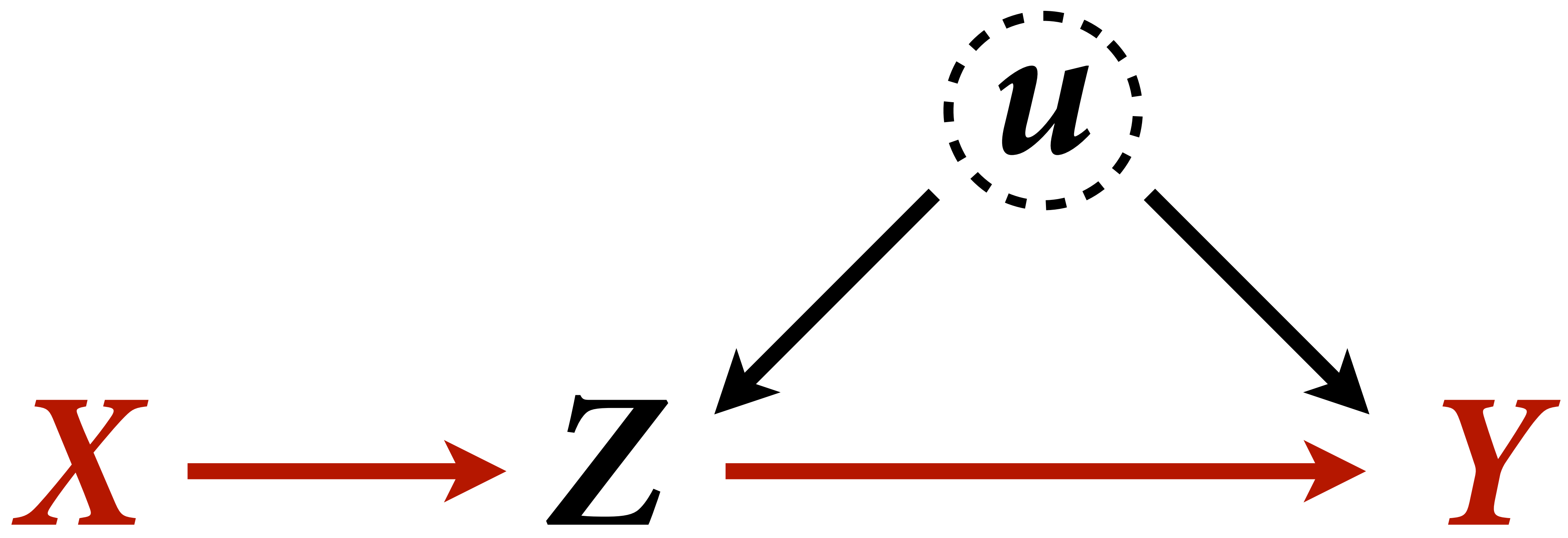
What happens if you stratify by Z?

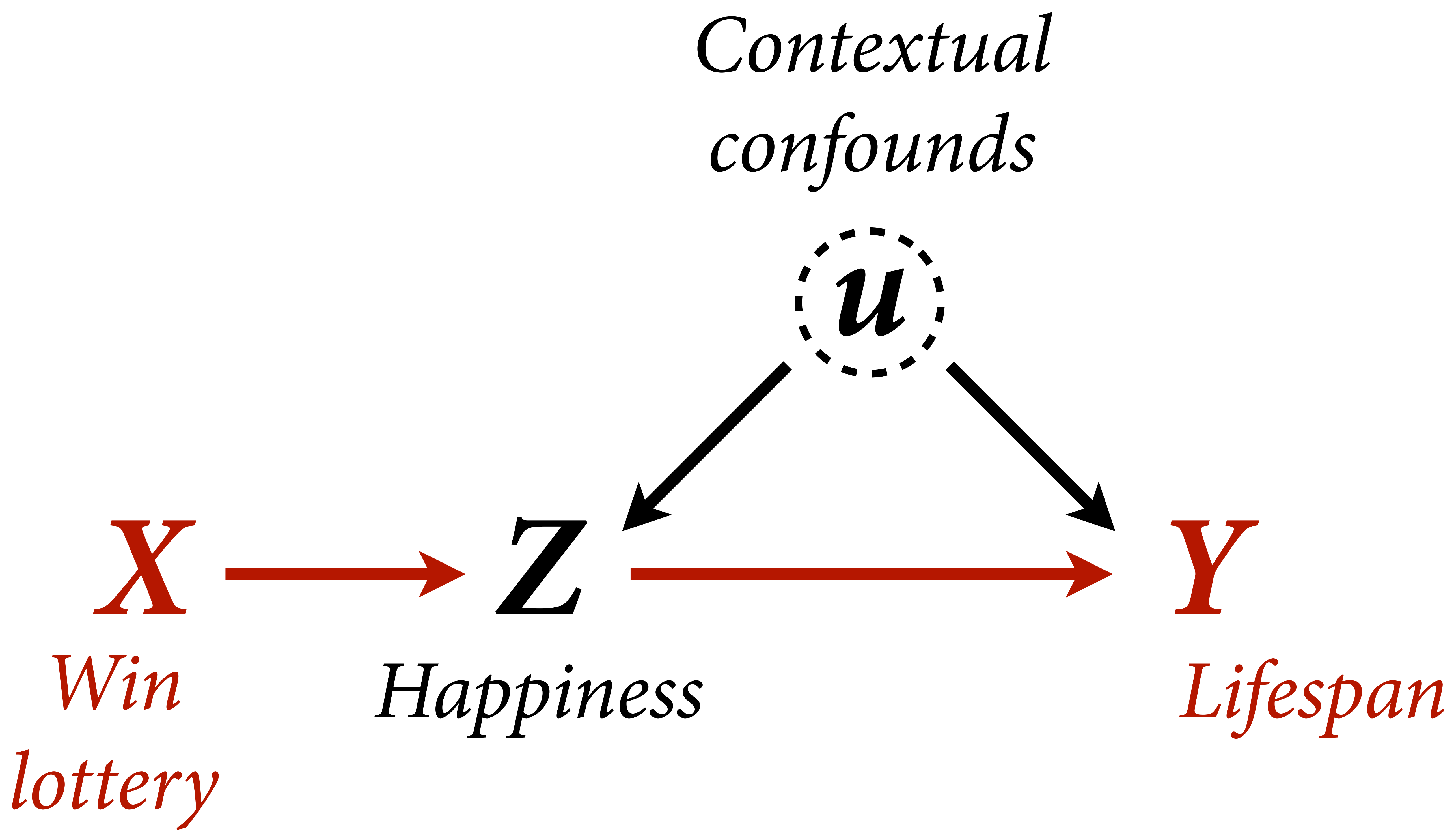
Opens the backdoor path

Z could be a **pre-treatment** variable

Not safe to always control pre-treatment measurements



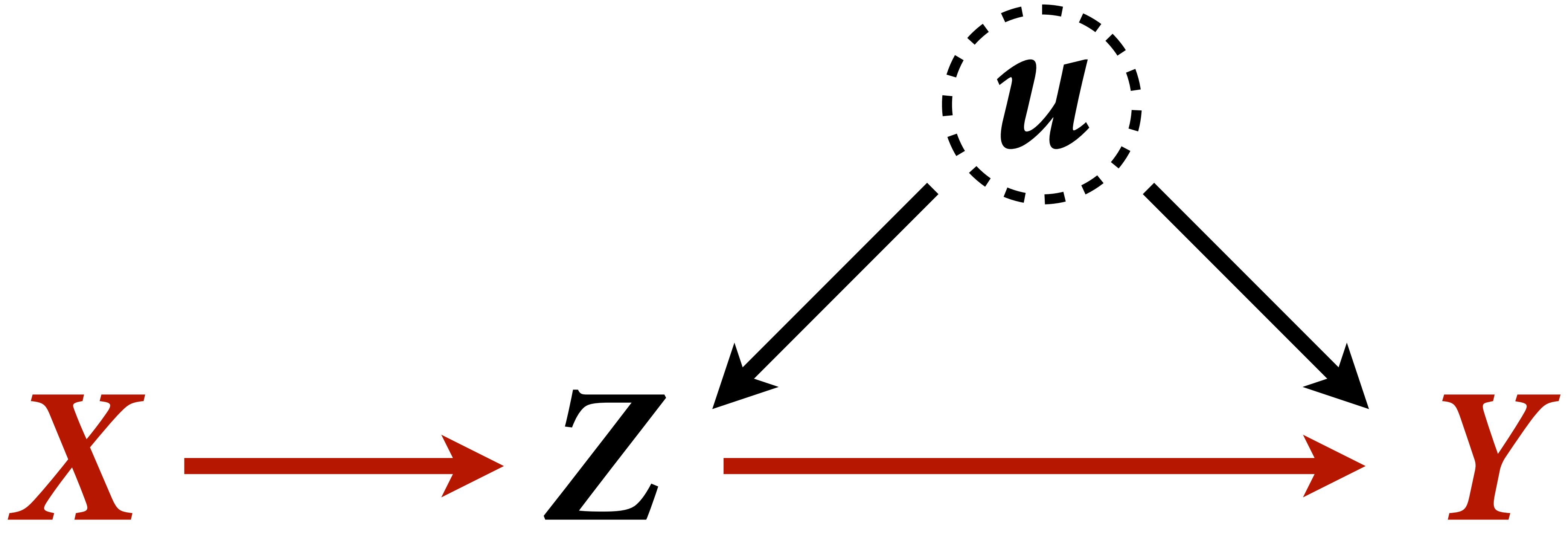

u
X
Z
Y


Contextual confounds
u
X
Win lottery
Z
Happiness
Y
Lifespan

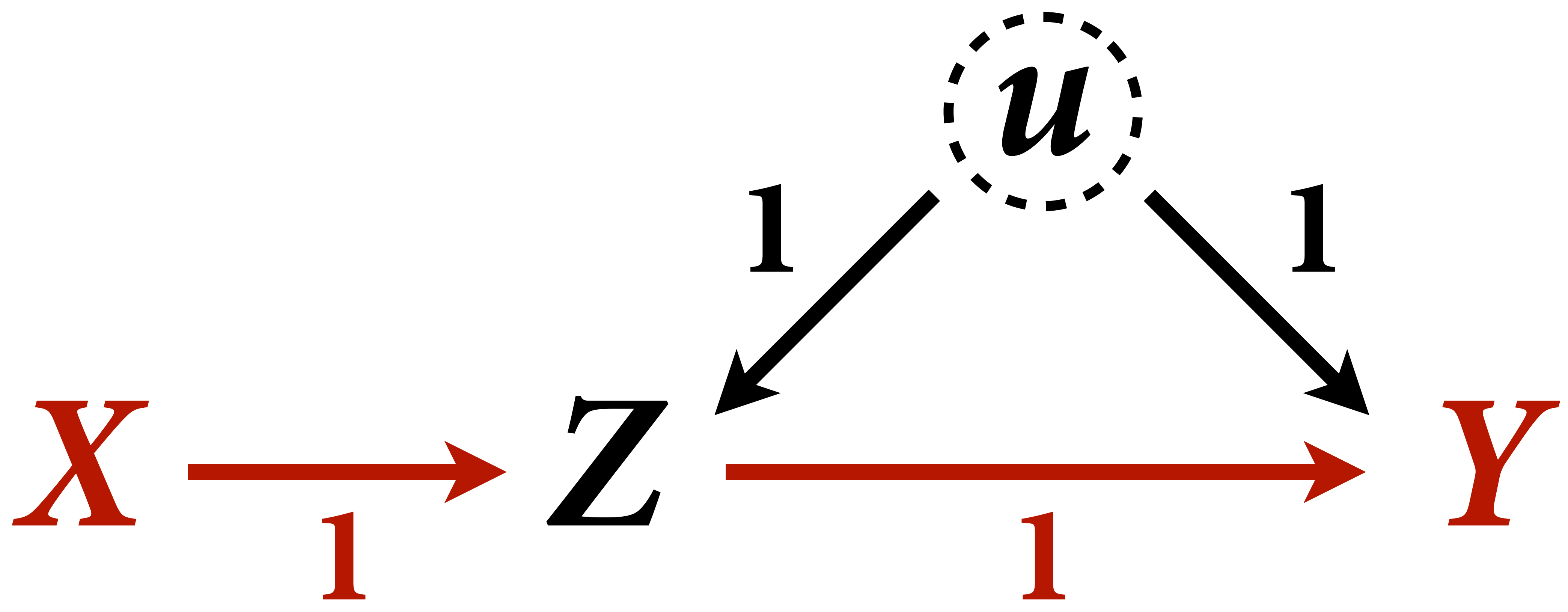
$$X \rightarrow Z \rightarrow Y$$

$$X \rightarrow Z \leftarrow u \rightarrow Y$$

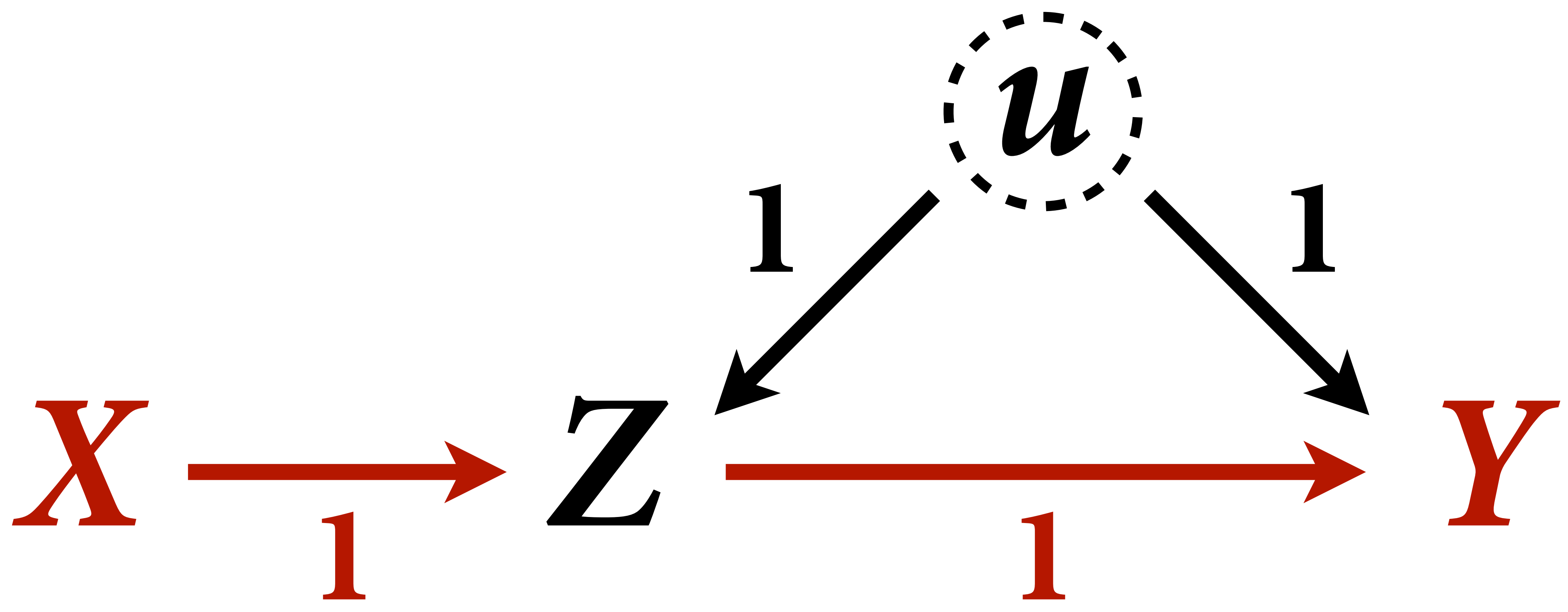
No backdoor, no need to control for Z

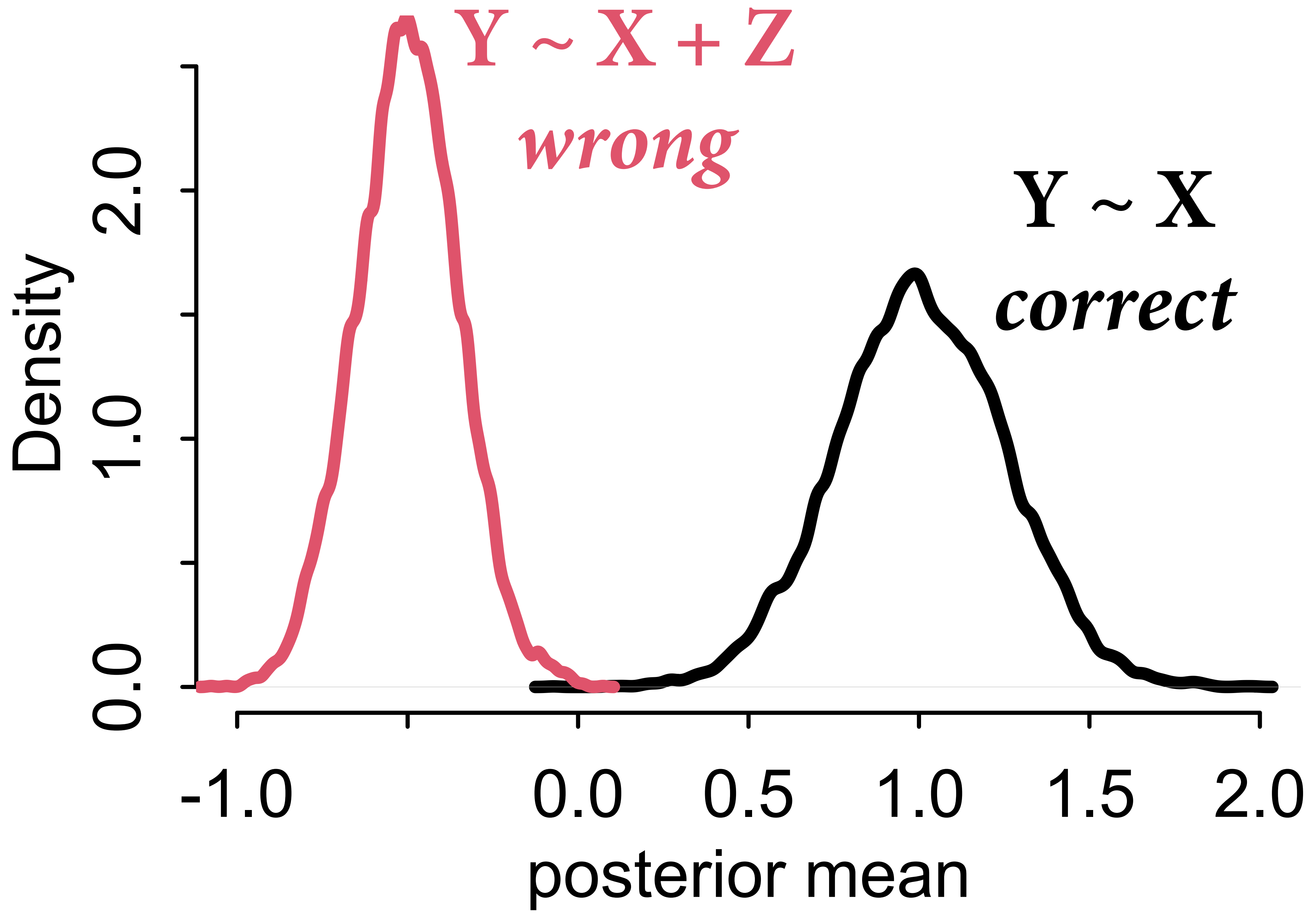


```
f <- function(n=100,bXZ=1,bZY=1) {
    X <- rnorm(n)
    u <- rnorm(n)
    Z <- rnorm(n, bXZ*X + u)
    Y <- rnorm(n, bZY*Z + u )
    bX <- coef( lm(Y ~ X) )['X']
    bXZ <- coef( lm(Y ~ X + Z) )['X']
    return( c(bX,bXZ) )
}

sim <- mcreplicate( 1e4 , f() , mc.cores=8 )

dens( sim[1,] , lwd=3 , xlab="posterior mean" )
dens( sim[2,] , lwd=3 , col=2 , add=TRUE )
```



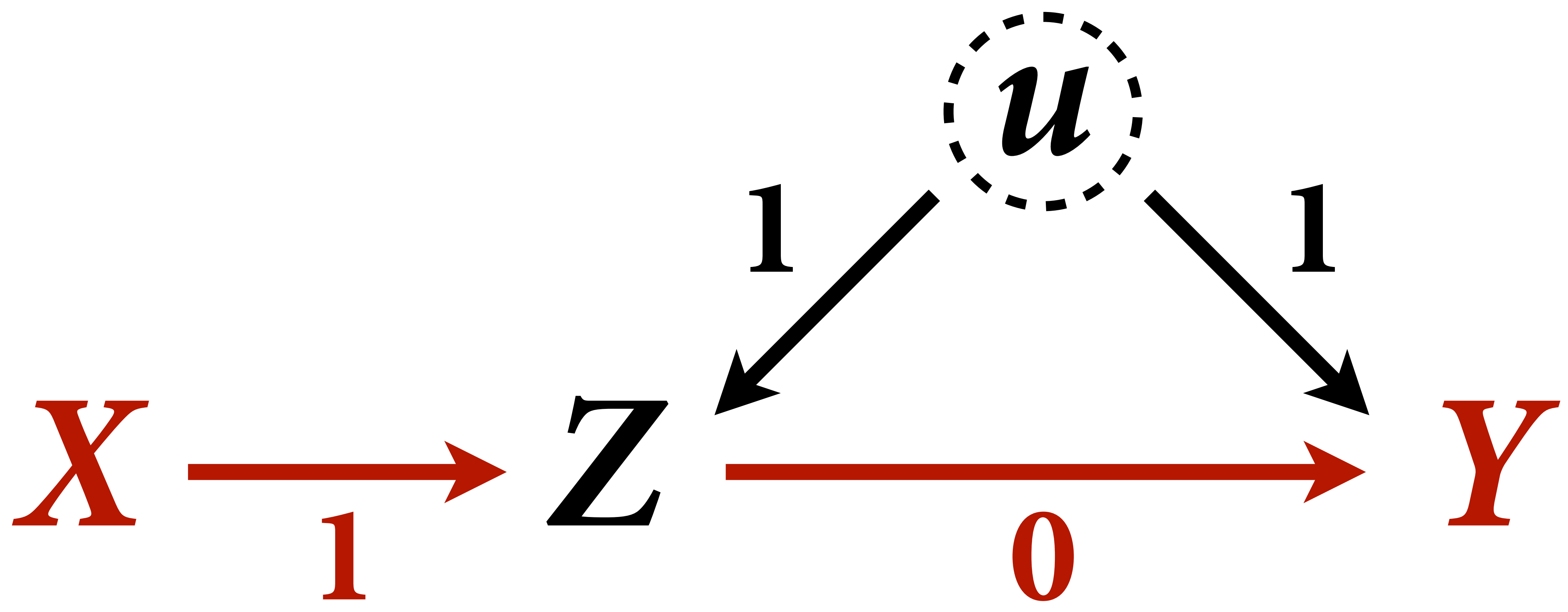
```
f <- function(n=100,bXZ=1,bZY=1) {
    X <- rnorm(n)
    u <- rnorm(n)
    Z <- rnorm(n, bXZ*X + u)
    Y <- rnorm(n, bZY*Z + u )
    bX <- coef( lm(Y ~ X) )['X']
    bXZ <- coef( lm(Y ~ X + Z) )['X']
    return( c(bX,bXZ) )
}

sim <- mcreplicate( 1e4 , f() , mc.cores=8 )

dens( sim[1,] , lwd=3 , xlab="posterior mean" )
dens( sim[2,] , lwd=3 , col=2 , add=TRUE )
```

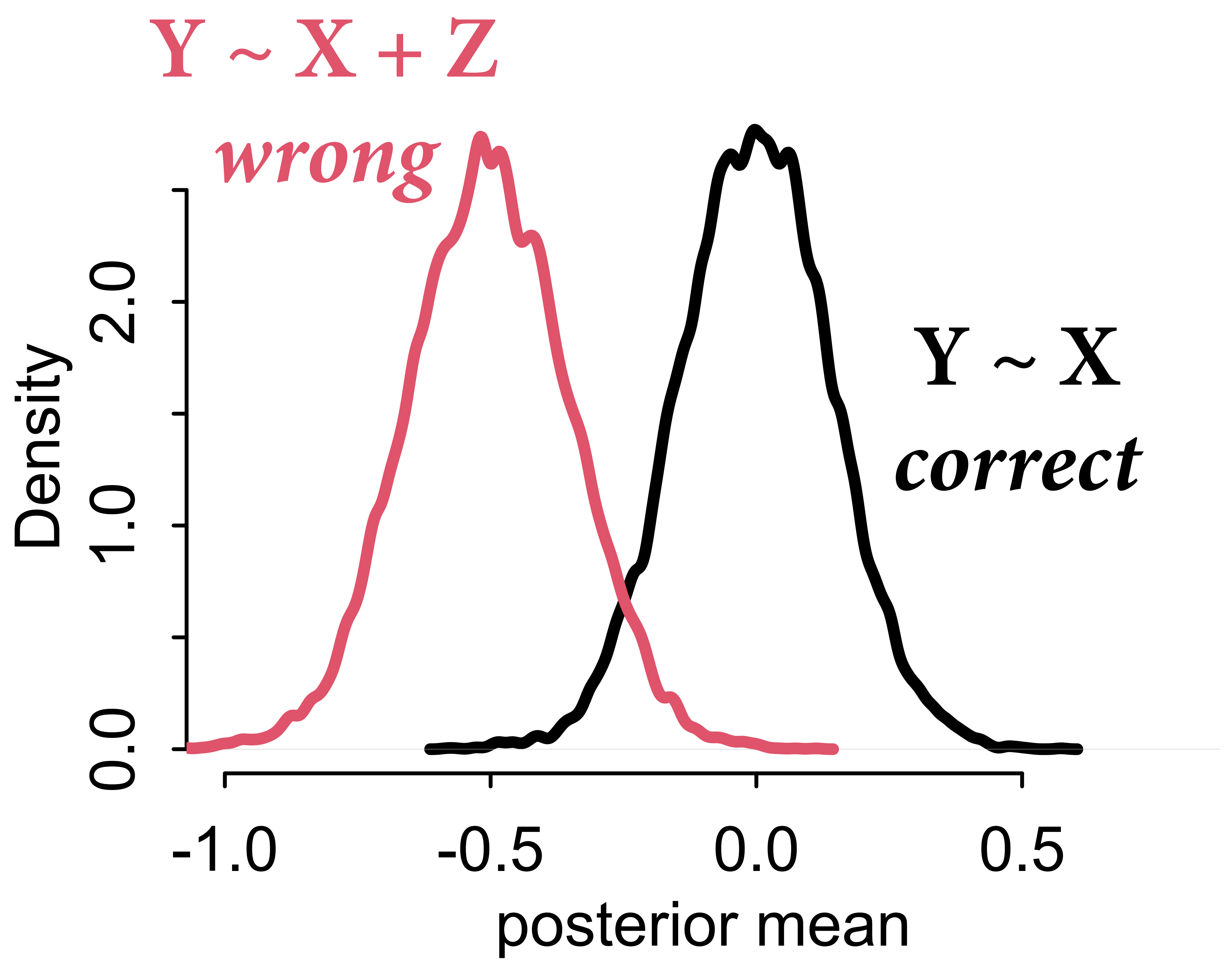




Change bZY to zero

```
f <- function(n=100,bXZ=1,bZY=1) {
    X <- rnorm(n)
    u <- rnorm(n)
    Z <- rnorm(n, bXZ*X + u)
    Y <- rnorm(n, bZY*Z + u )
    bX <- coef( lm(Y ~ X) )['X']
    bXZ <- coef( lm(Y ~ X + Z) )['X']
    return( c(bX,bXZ) )
}

sim <- mcreplicate( 1e4 , f(bZY=0) , mc.cores=8 )

dens( sim[1,] , lwd=3 , xlab="posterior mean" )
dens( sim[2,] , lwd=3 , col=2 , add=TRUE )
```





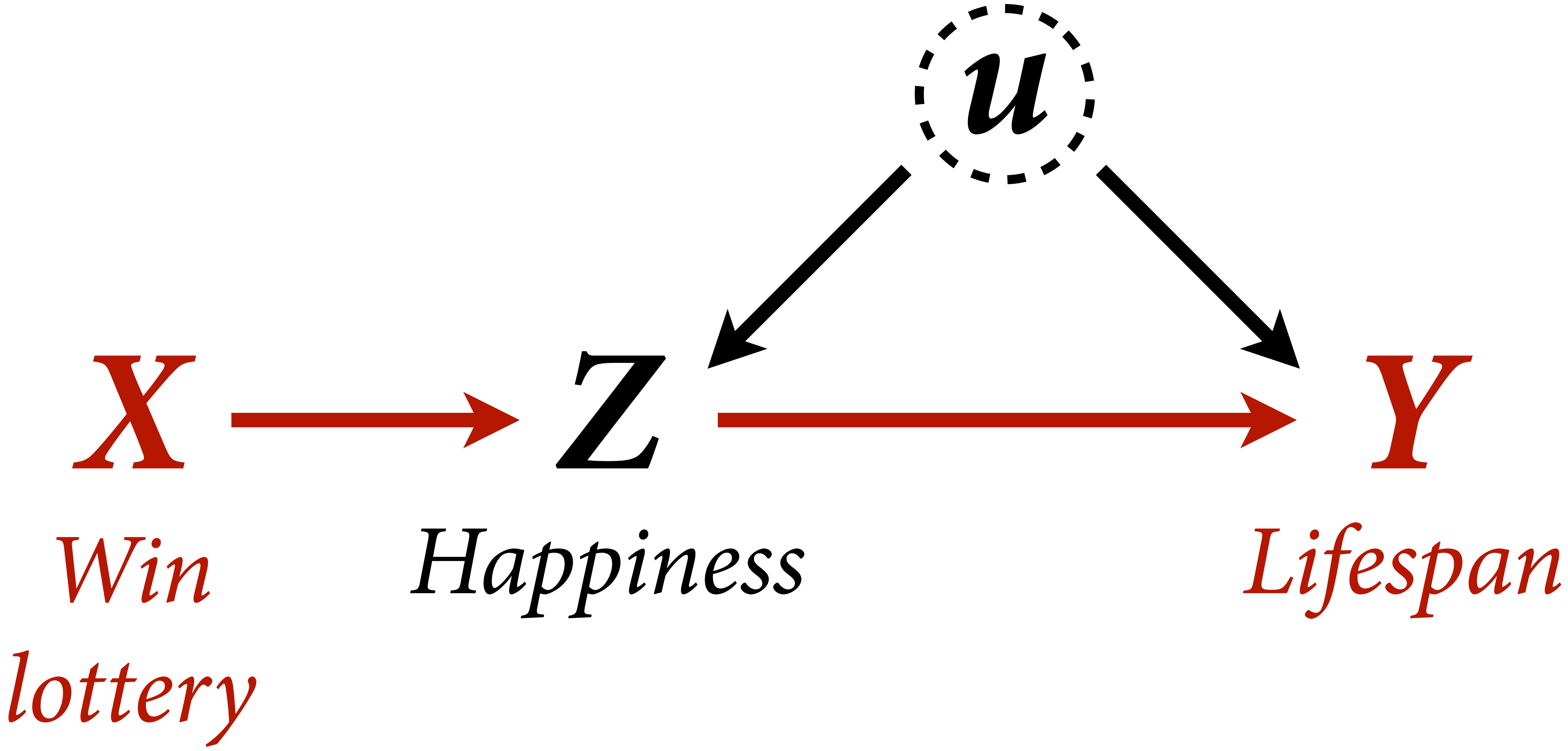
$X \to Z \to Y$
$X \to Z \leftarrow u \to Y$

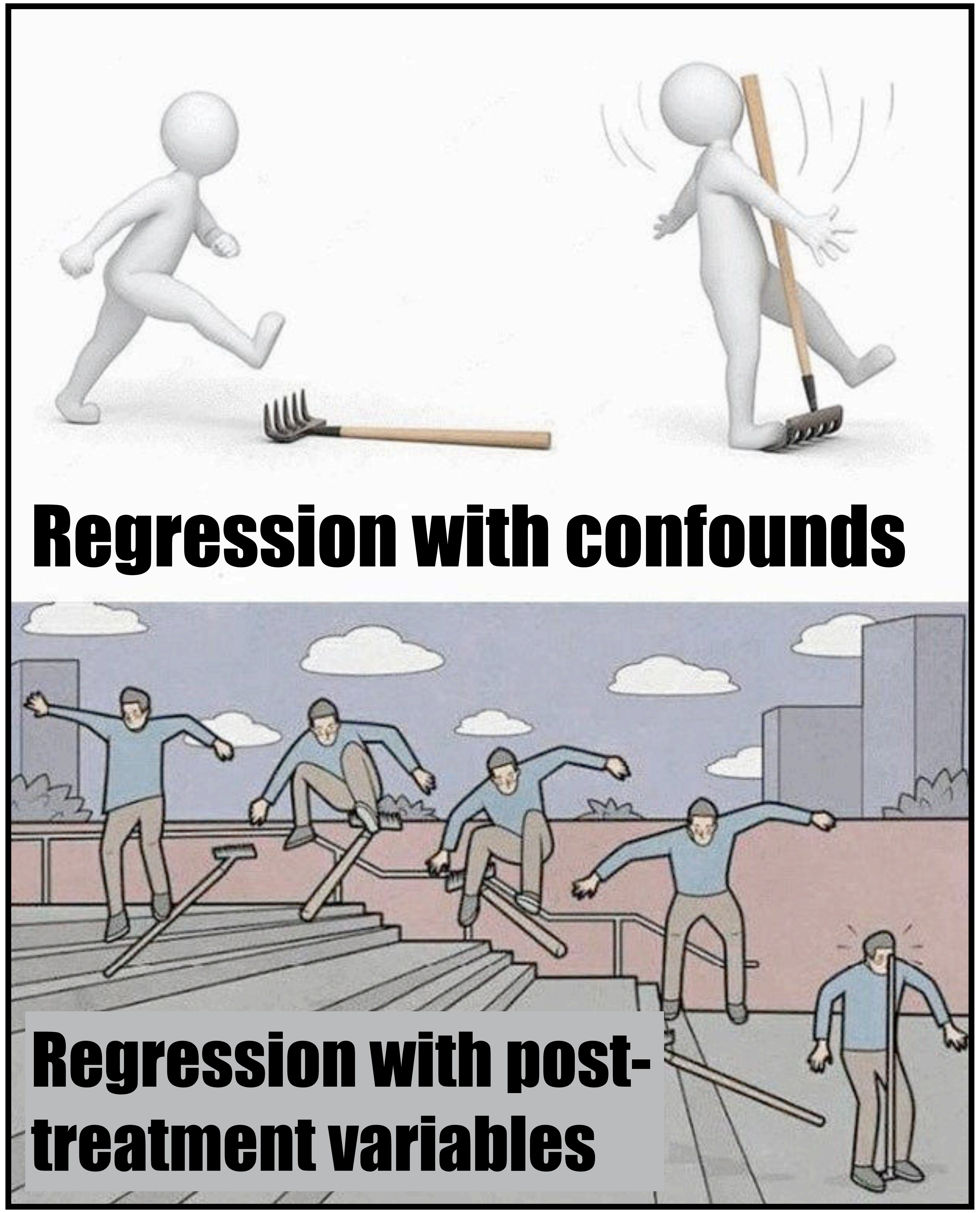
No backdoor, no need to control for Z

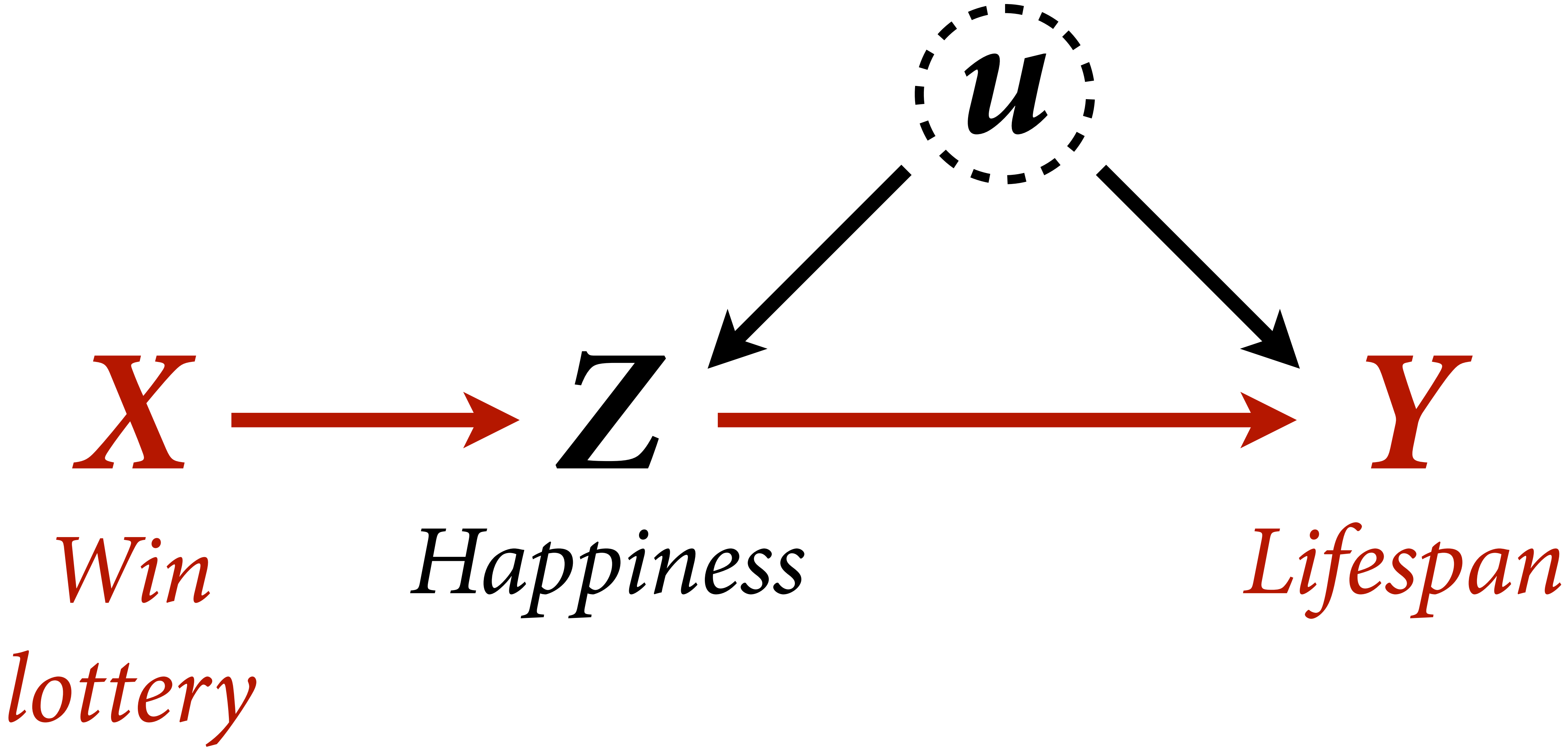
Controlling for Z biases treatment estimate X

Controlling for Z opens biasing path through u

Can estimate effect of X; Cannot estimate mediation effect Z

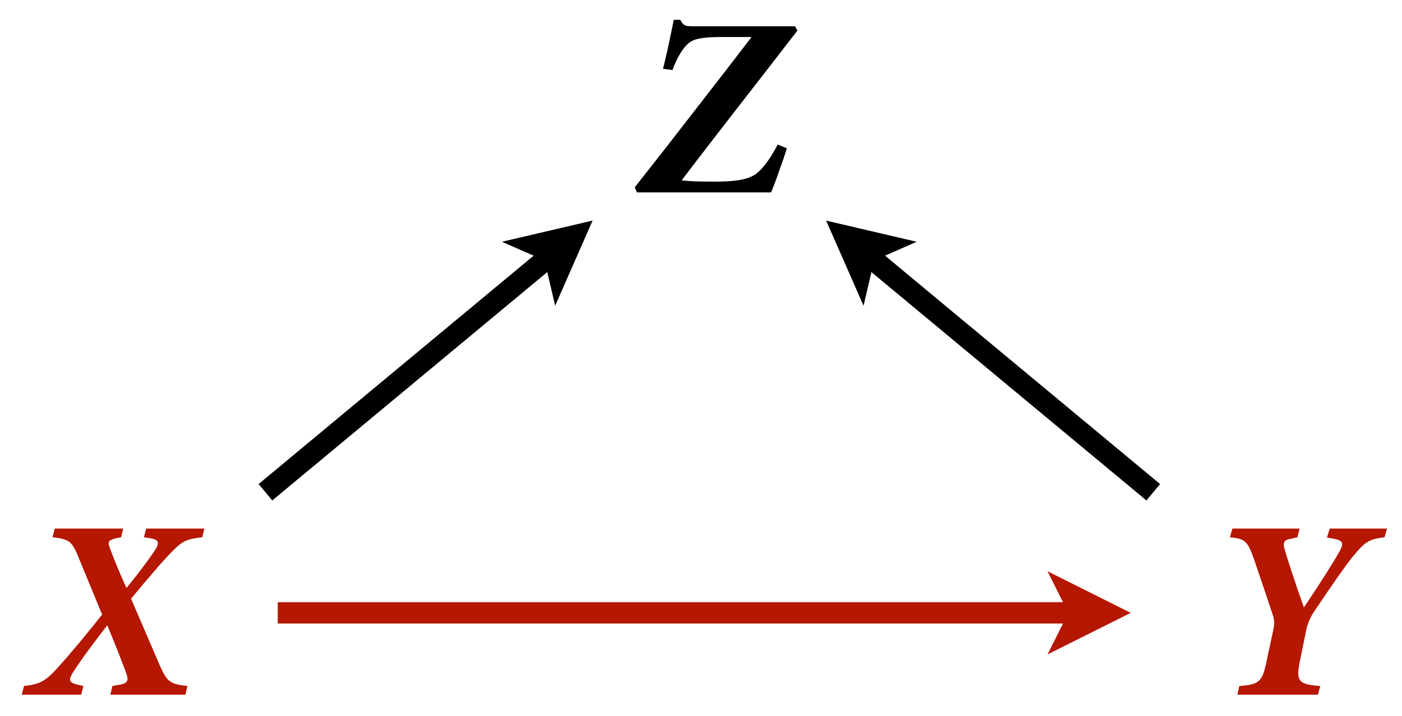




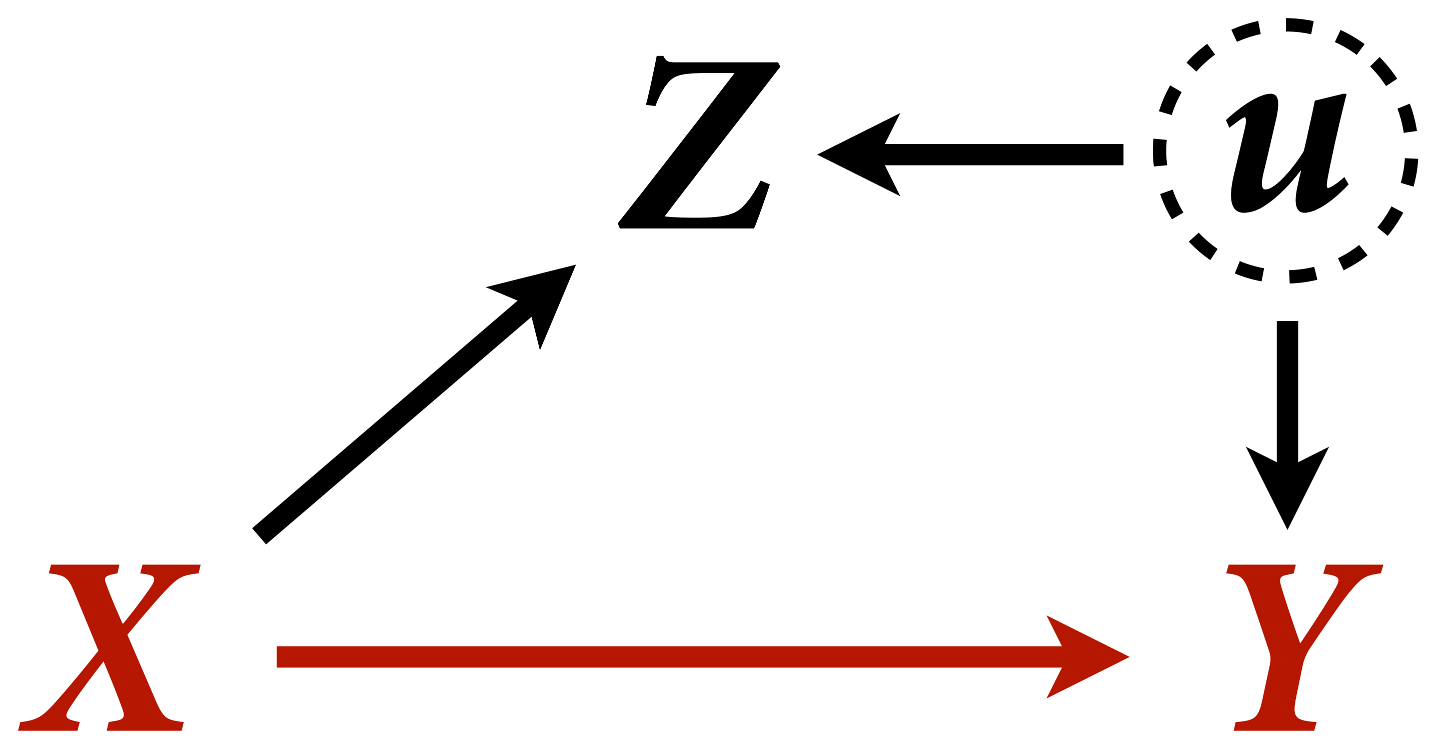


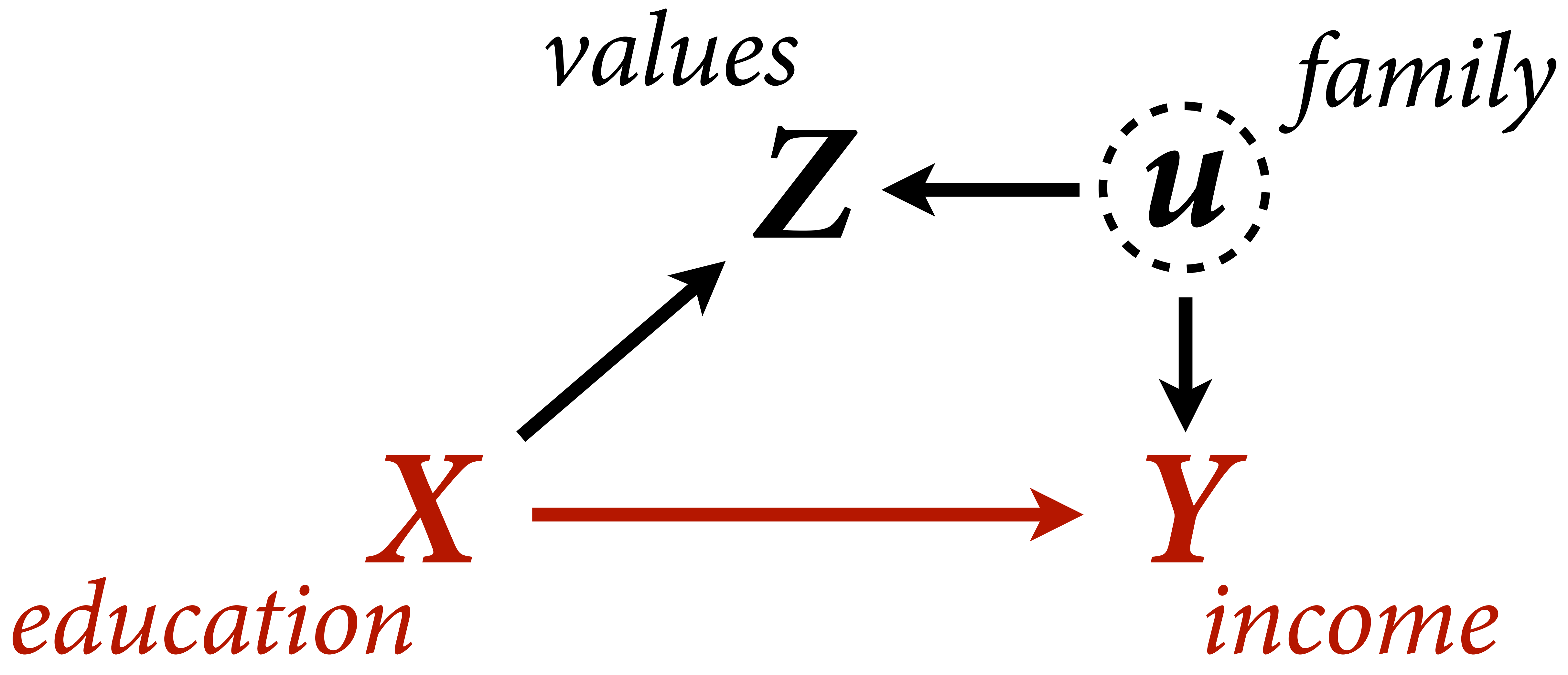
Montgomery et al 2018 How Conditioning on Posttreatment Variables Can Ruin Your Experiment

Do not touch the collider!

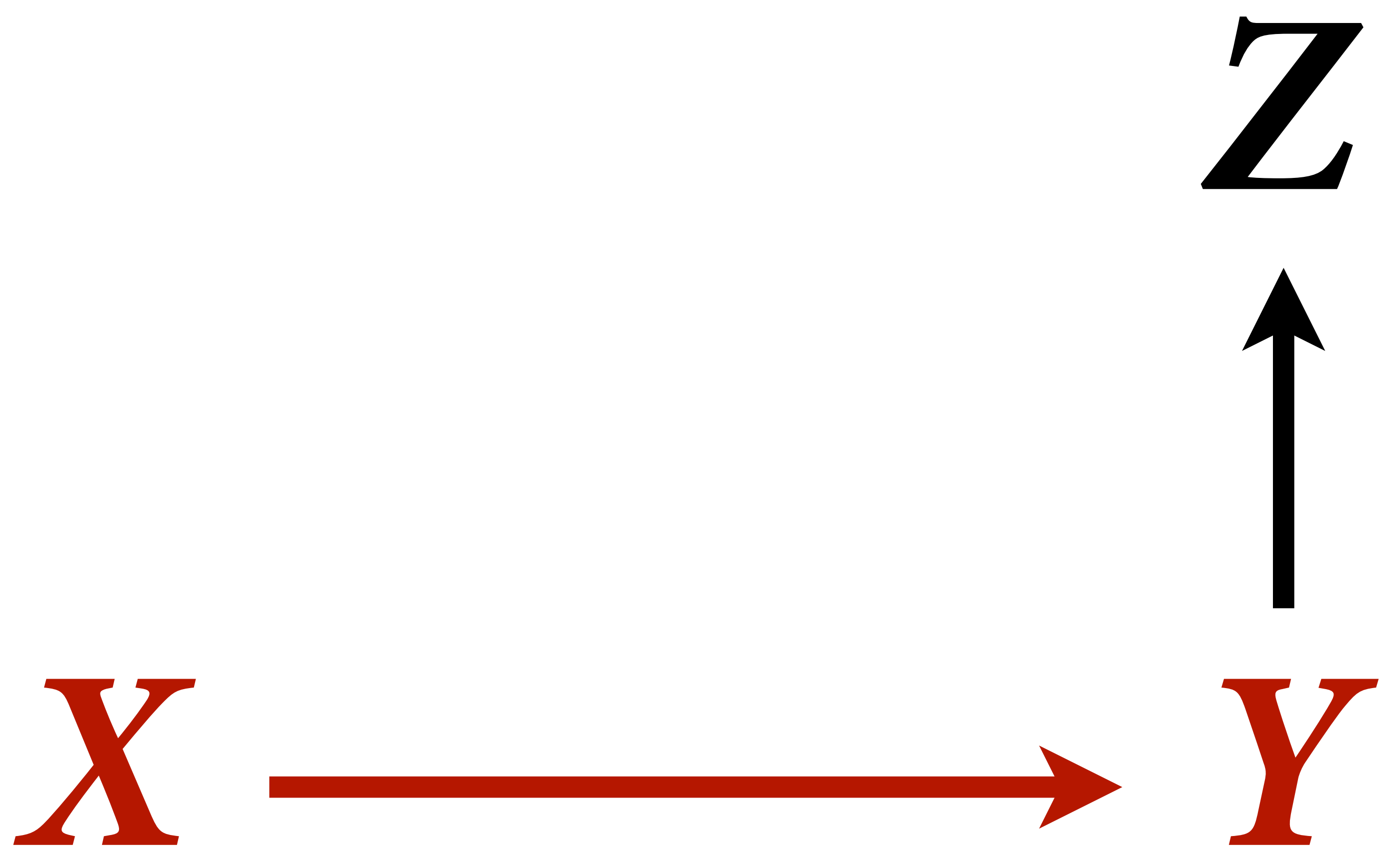


Colliders not always so obvious

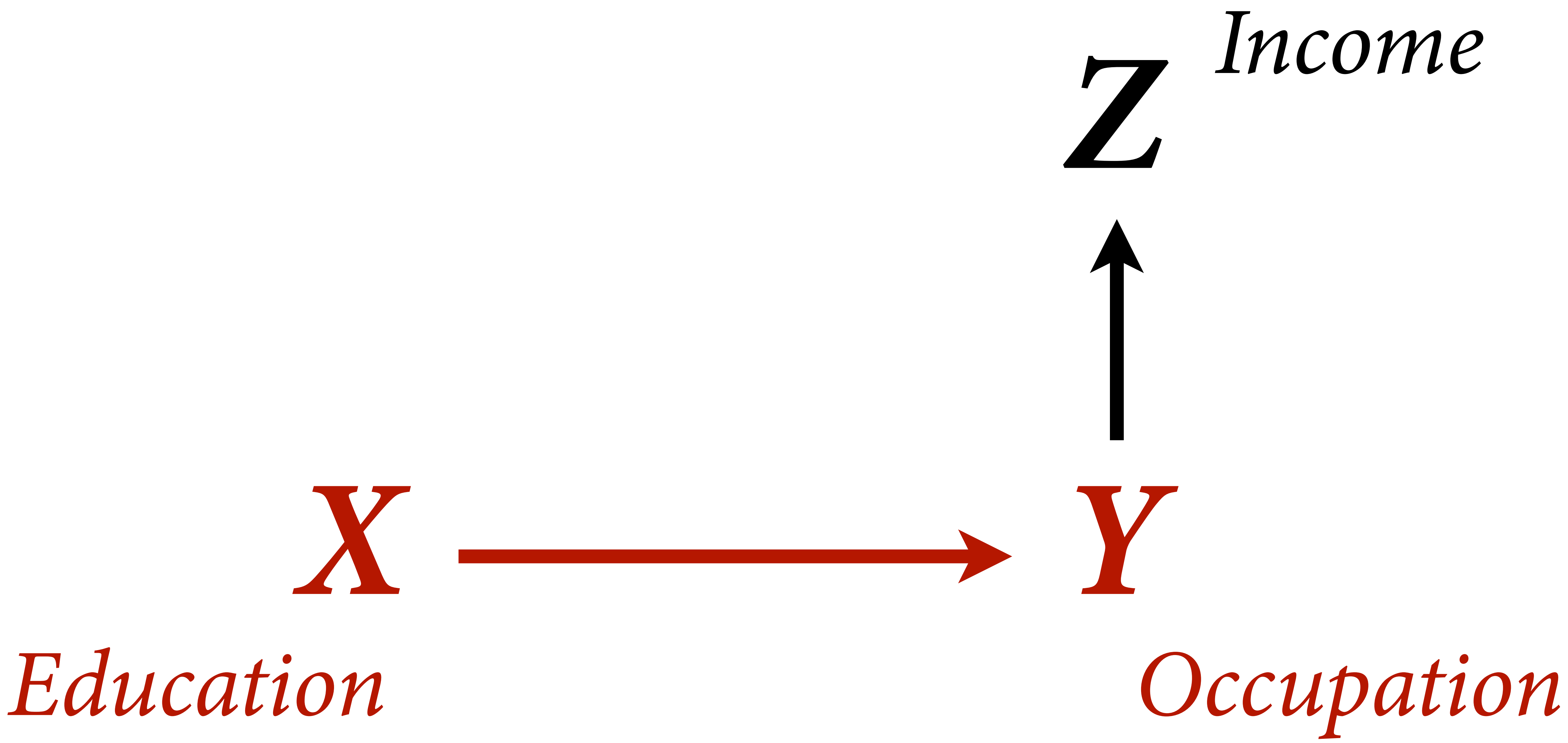



values
family
Z
u
X
Y
education
income

Case-control bias
(selection on outcome)

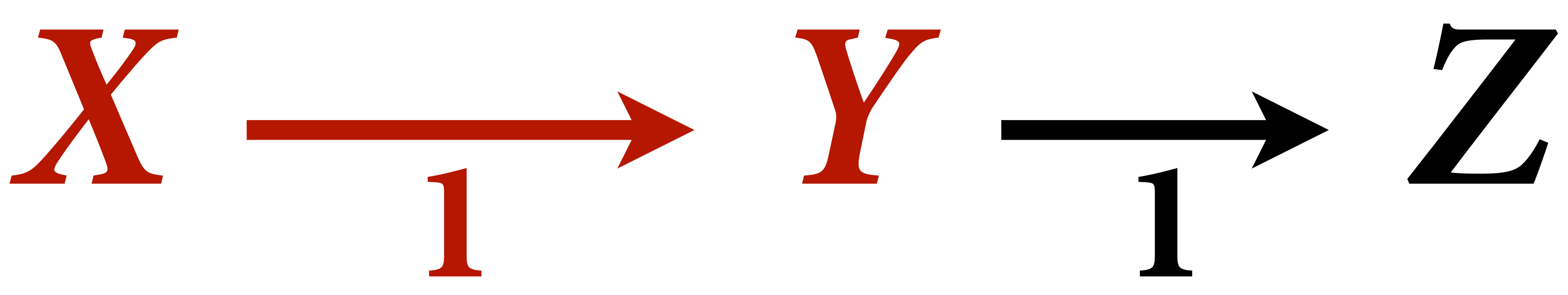


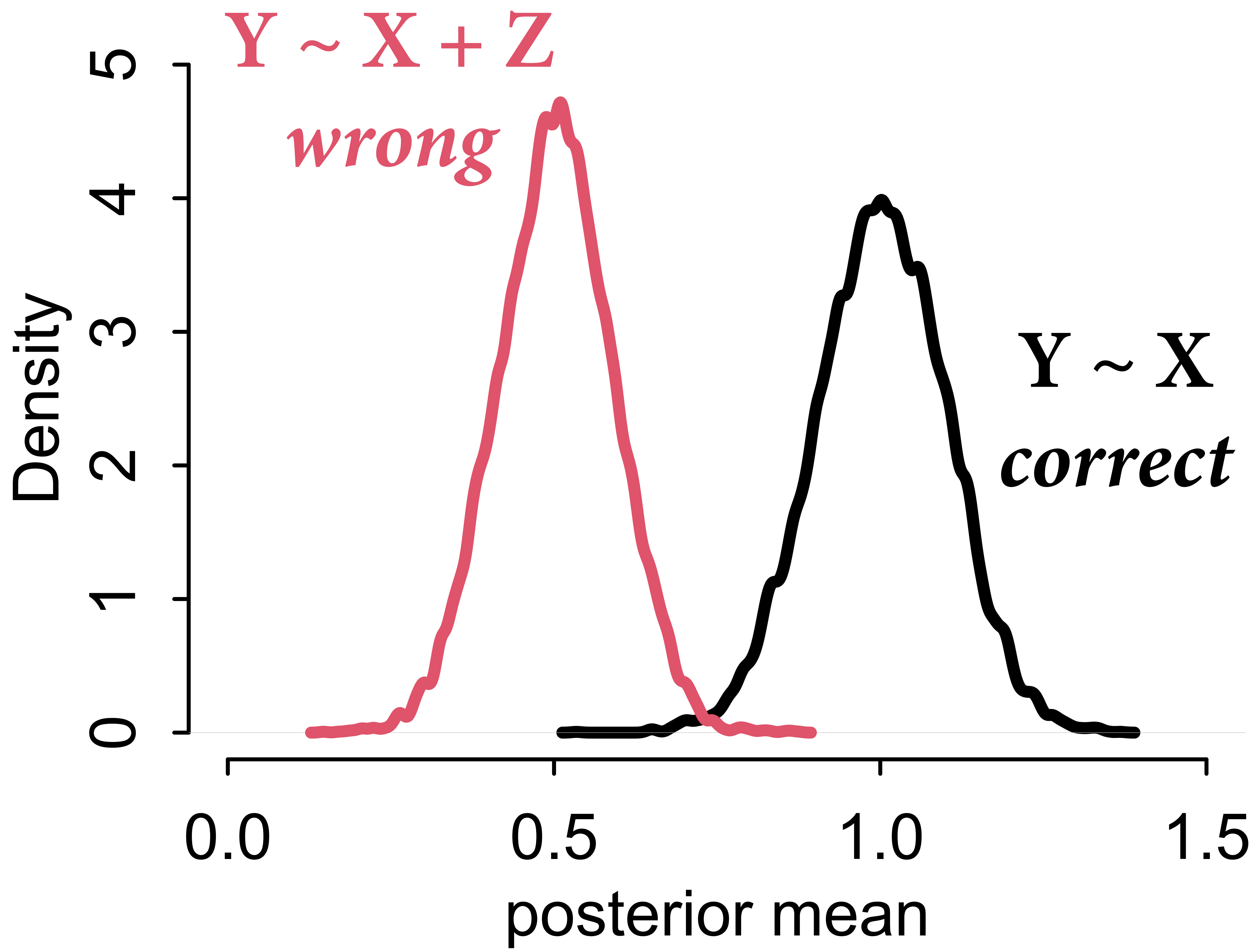
Case-control bias
(selection on outcome)



Case-control bias
(selection on outcome)

```
f <- function(n=100,bXY=1,bYZ=1) {
    X <- rnorm(n)
    Y <- rnorm(n, bXY*X )
    Z <- rnorm(n, bYZ*Y )
    bX <- coef( lm(Y ~ X) )['X']
    bXZ <- coef( lm(Y ~ X + Z) )['X']
    return( c(bX,bXZ) )
}

sim <- mcreplicate( 1e4 , f() , mc.cores=8 )

dens( sim[1,] , lwd=3 , xlab="posterior mean" )
dens( sim[2,] , lwd=3 , col=2 , add=TRUE )
```





"Precision parasite"

No backdoors

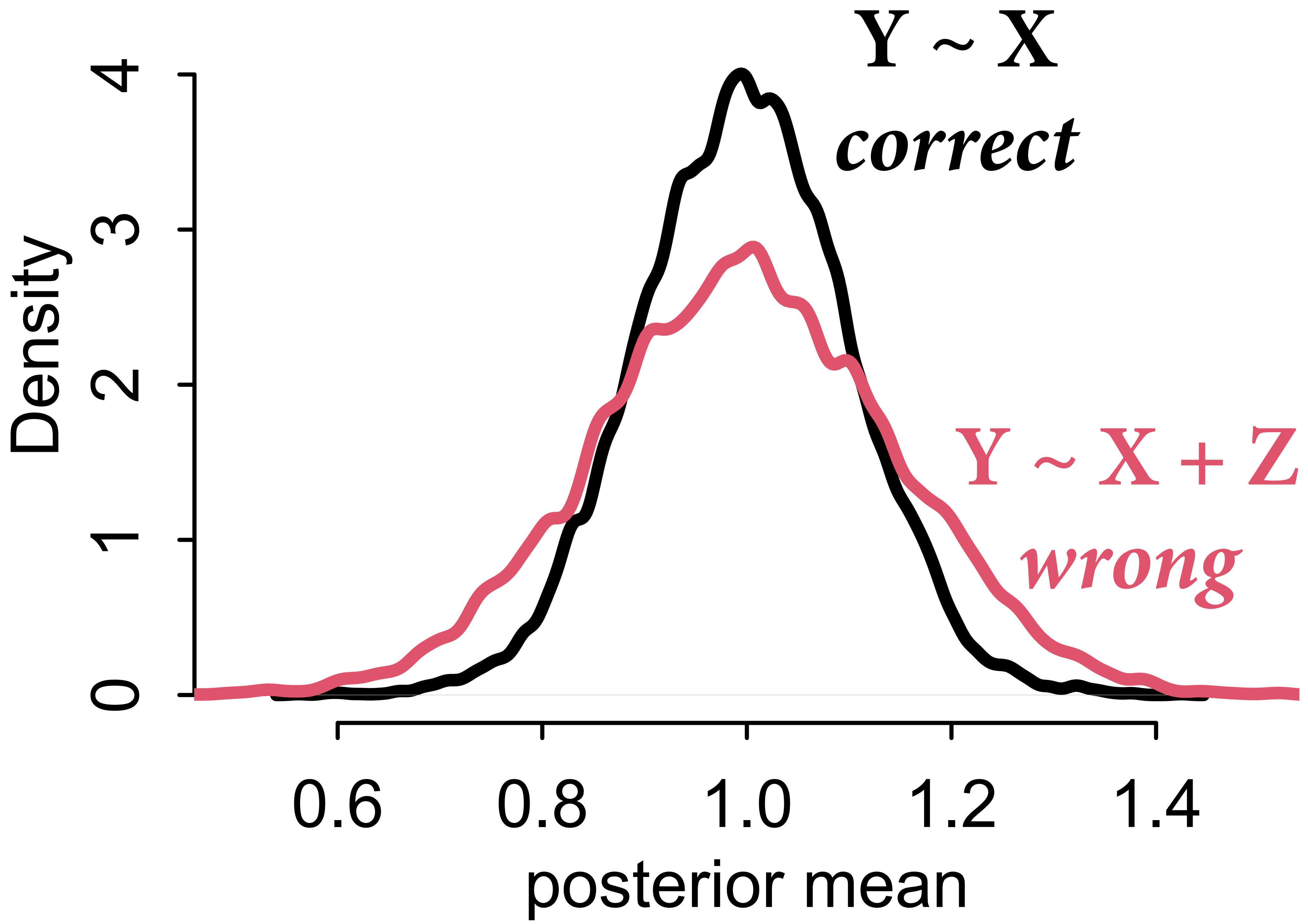
But still not good to condition on Z

$$\begin{tikzcd} Z \arrow[d] \\ X \arrow[r] & Y \end{tikzcd}$$

“Precision parasite”

$Z \longrightarrow X \longrightarrow Y$

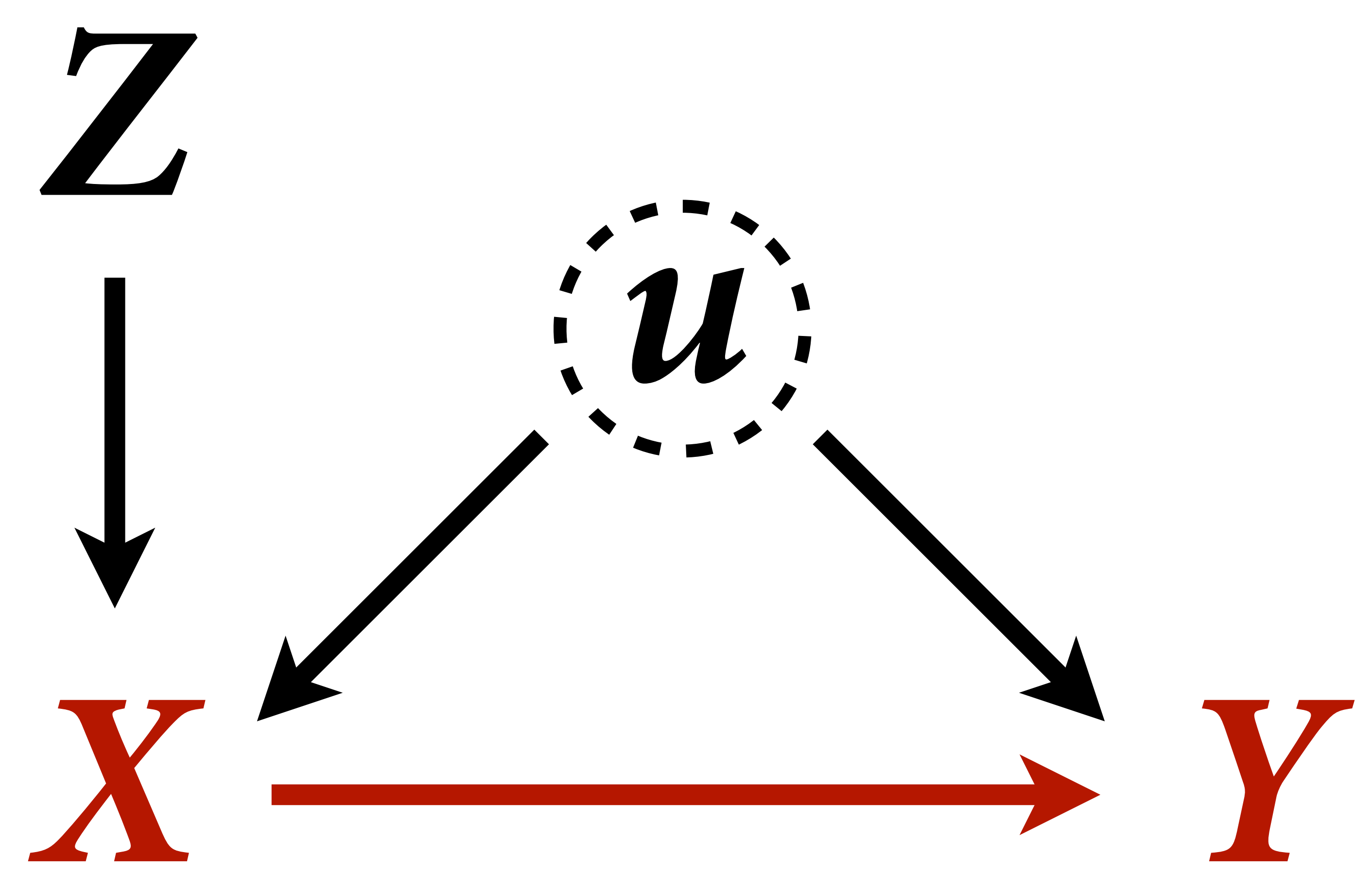
```
f <- function(n=100,bZX=1,bXY=1) {
    Z <- rnorm(n)
    X <- rnorm(n, bZX*Z )
    Y <- rnorm(n, bXY*X )
    bX <- coef( lm(Y ~ X) )['X']
    bXZ <- coef( lm(Y ~ X + Z) )['X']
    return( c(bX,bXZ) )
}

sim <- mcreplicate( 1e4 , f(n=50) , mc.cores=8 )

dens( sim[1,] , lwd=3 , xlab="posterior mean" )
dens( sim[2,] , lwd=3 , col=2 , add=TRUE )
```



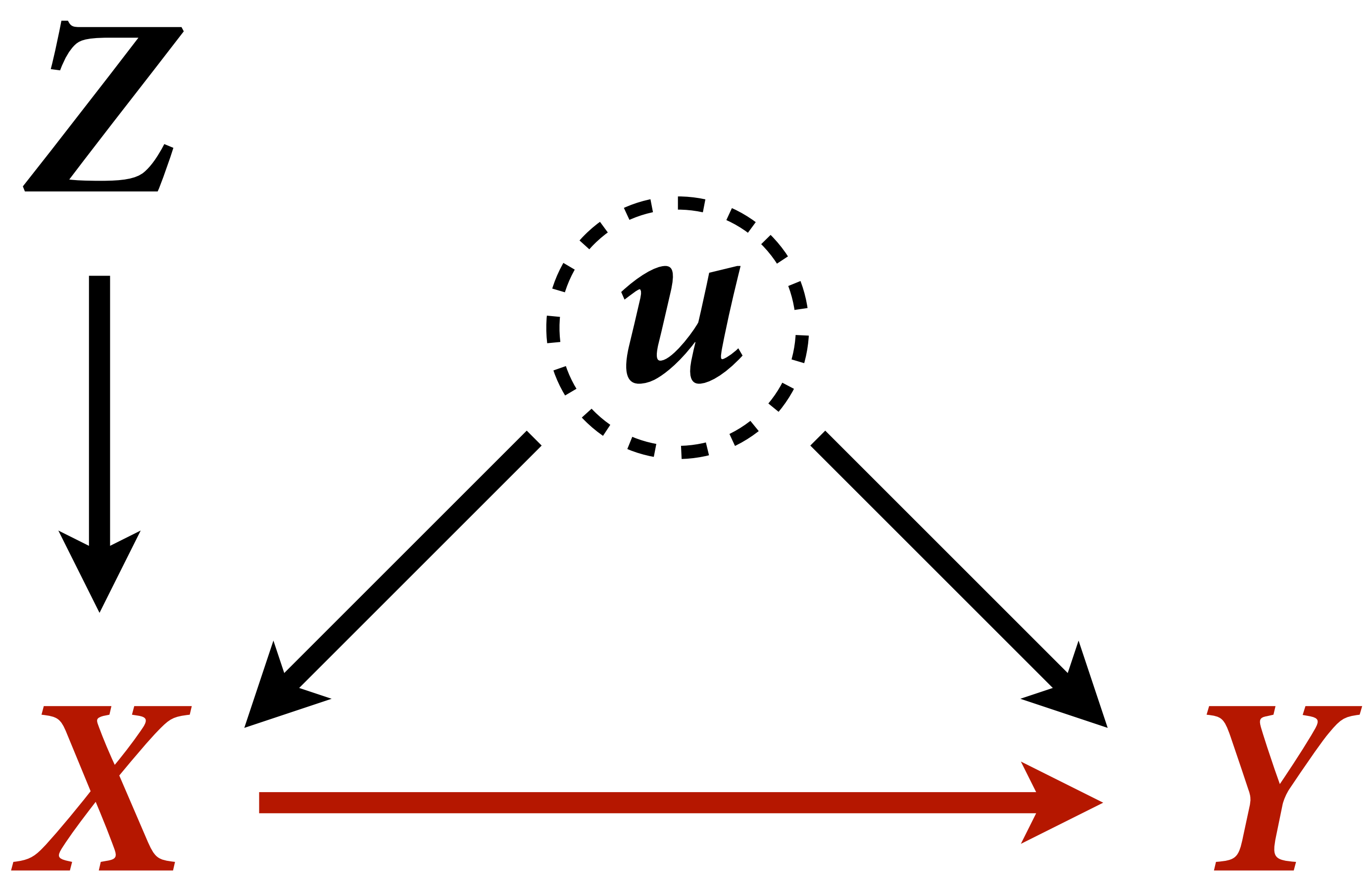
"Bias amplification"

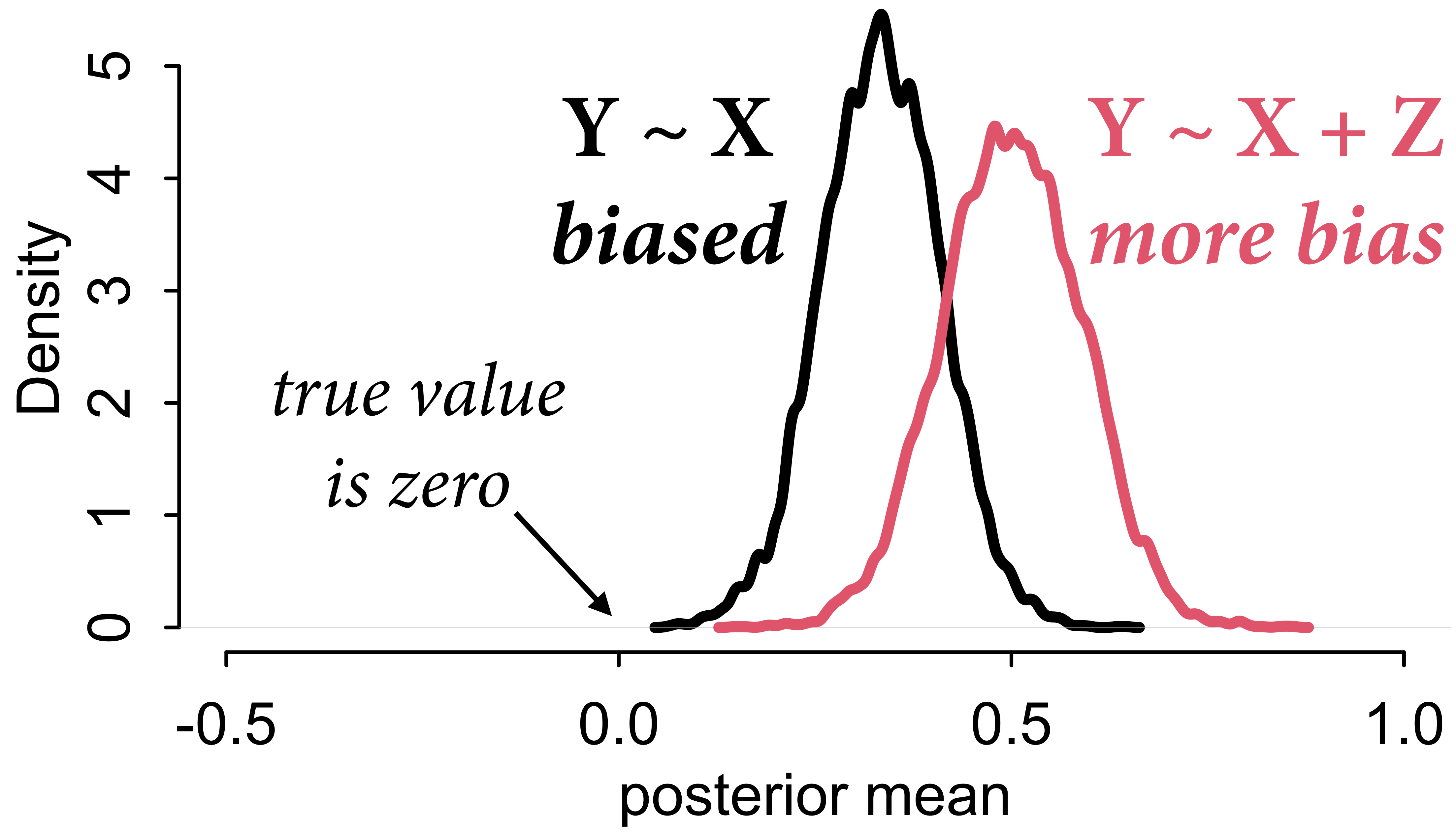
X and ***Y*** confounded by ***u***

Something **truly awful** happens when we add ***Z***



```
f <- function(n=100,bZX=1,bXY=1) {
    Z <- rnorm(n)
    u <- rnorm(n)
    X <- rnorm(n, bZX*Z + u )
    Y <- rnorm(n, bXY*X + u )
    bX <- coef( lm(Y ~ X) )['X']
    bXZ <- coef( lm(Y ~ X + Z) )['X']
    return( c(bX,bXZ) )
}

sim <- mcreplicate( 1e4 , f(bXY=0) , mc.cores=8 )

dens( sim[1,] , lwd=3 , xlab="posterior mean" )
dens( sim[2,] , lwd=3 , col=2 , add=TRUE )
```



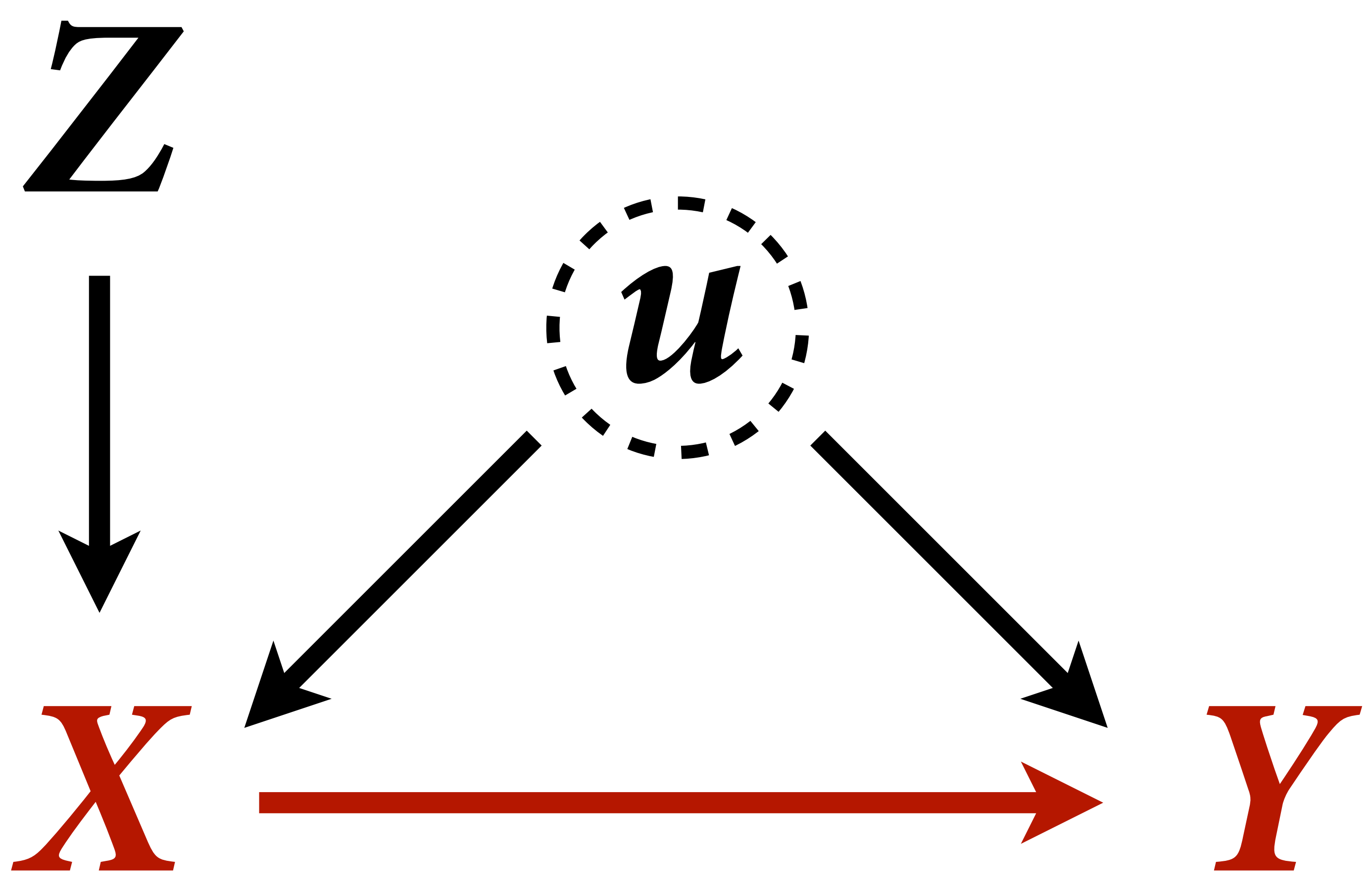


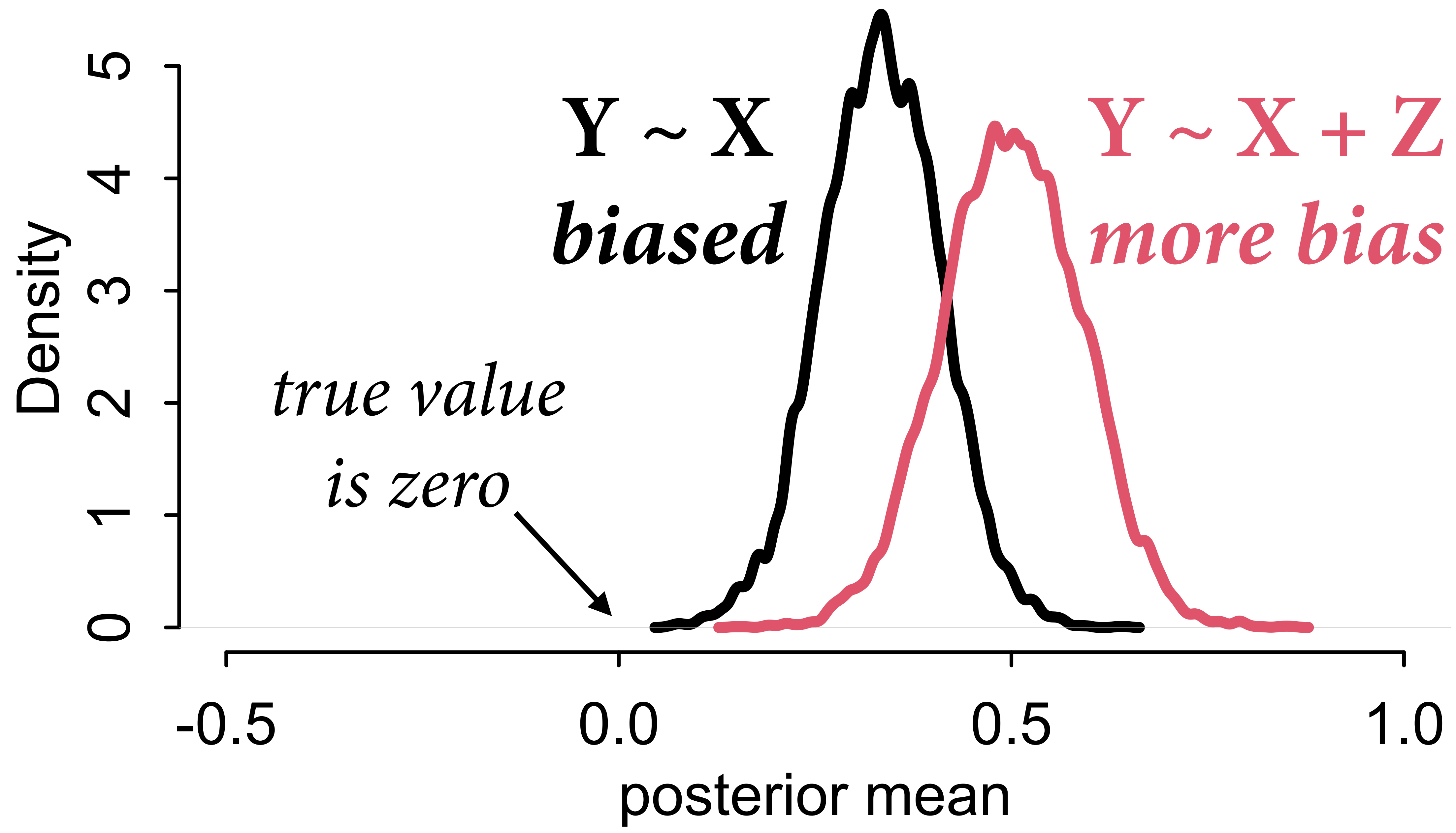
WHY?

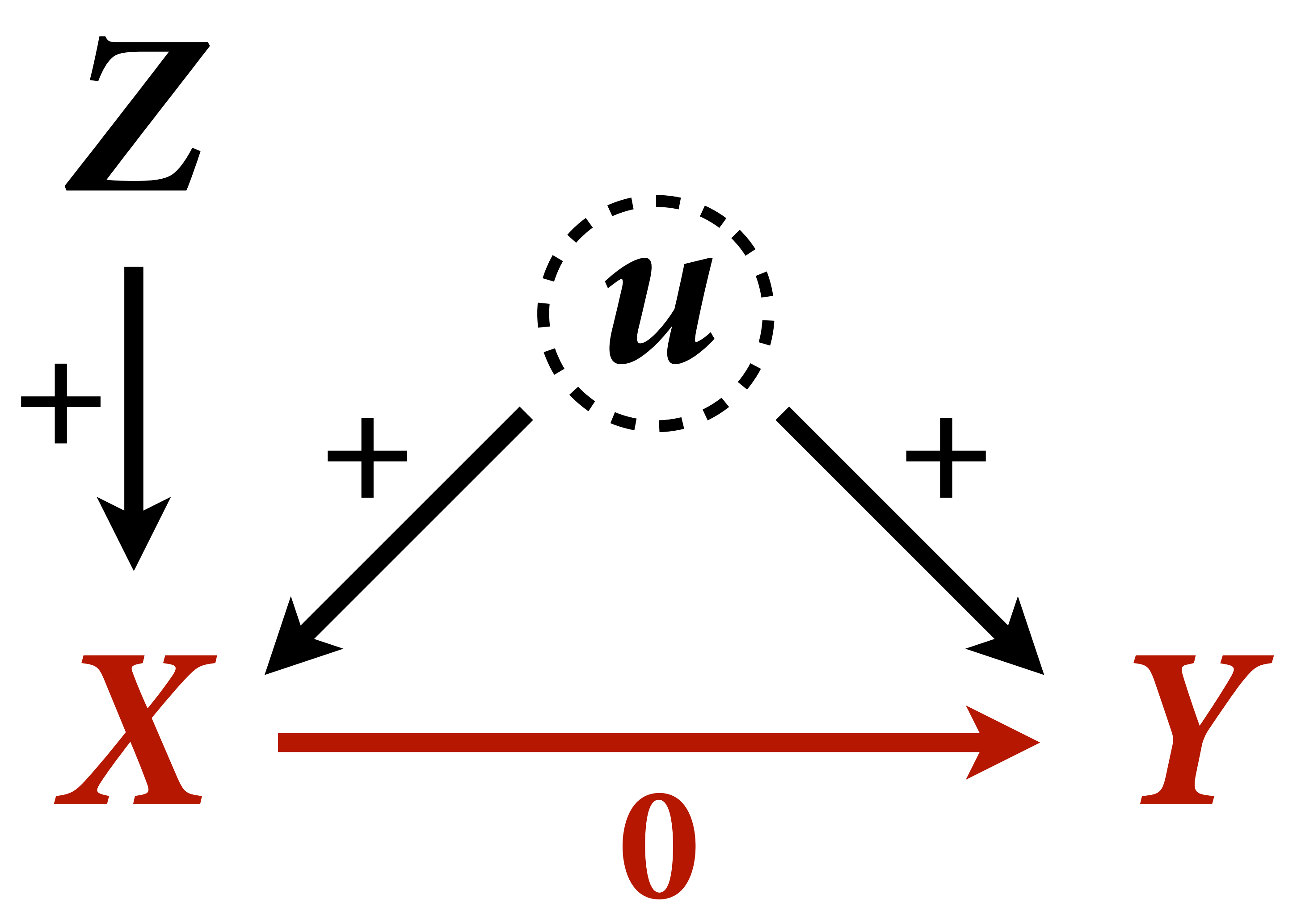
Covariation ***X*** & ***Y*** requires variation in their causes

Within each level of ***Z***, less variation in ***X***

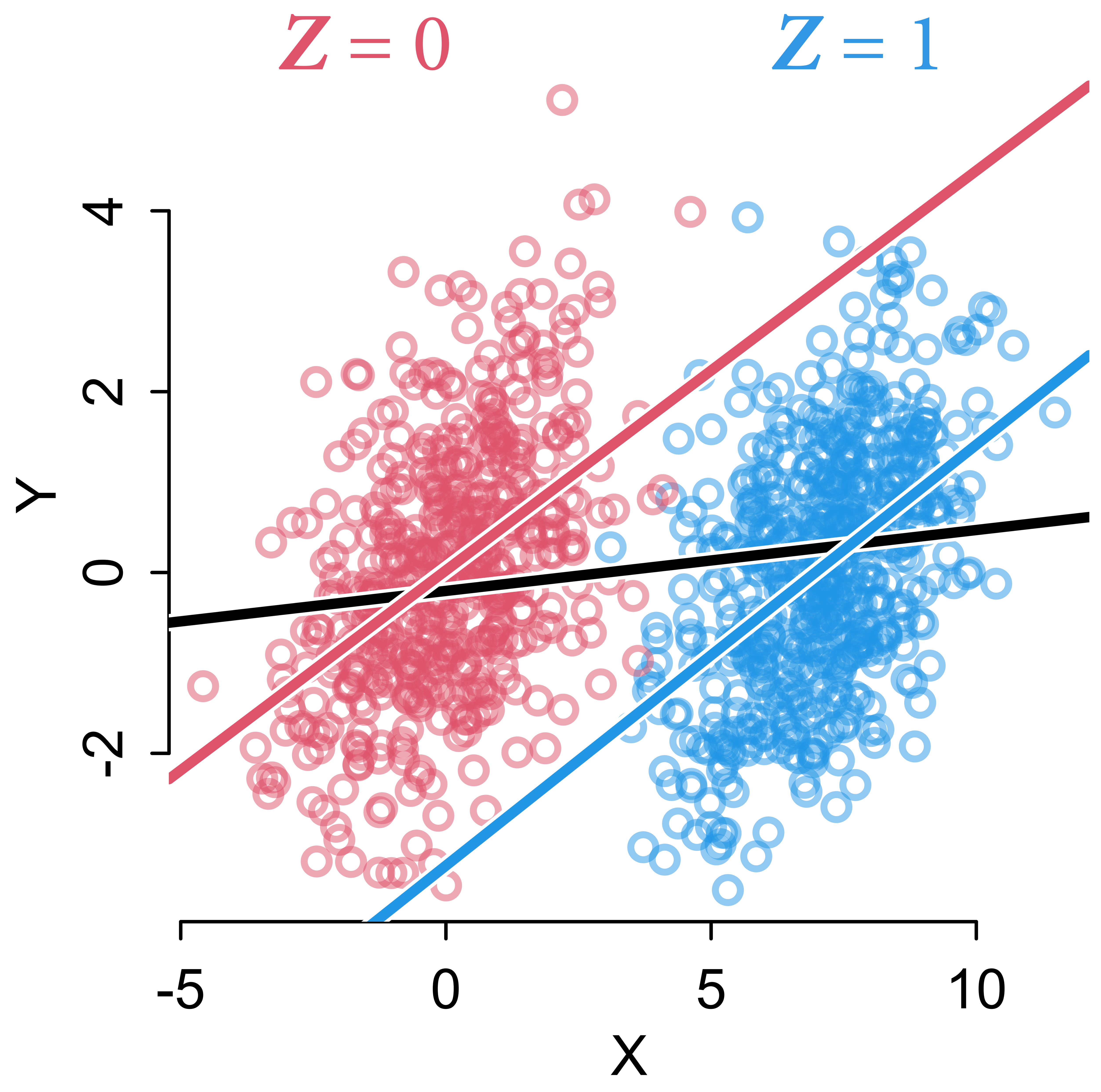
Confound ***u*** relatively more important within each ***Z***

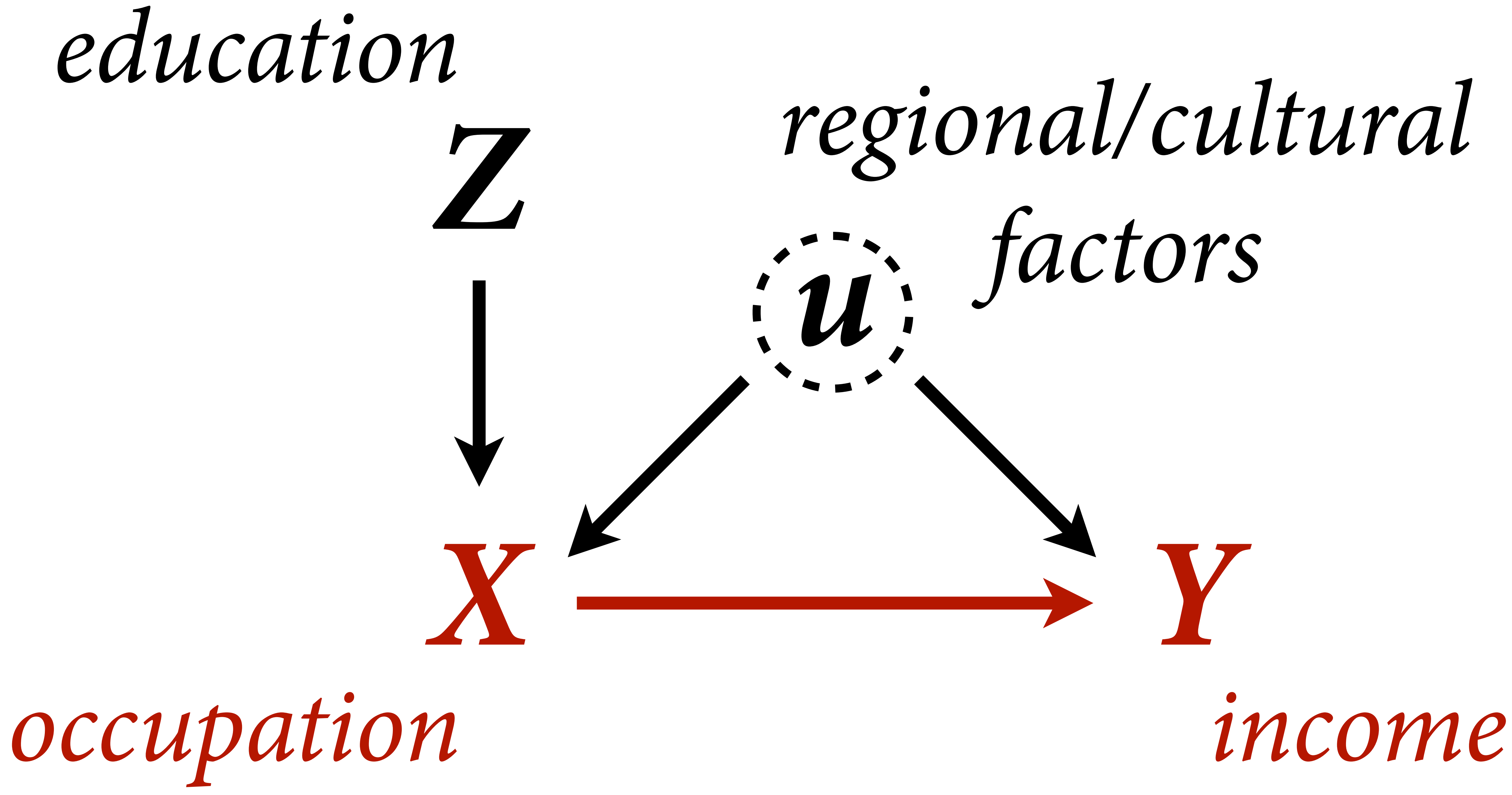







```
n <- 1000
Z <- rbern(n)
u <- rnorm(n)
X <- rnorm(n, 7*Z + u )
Y <- rnorm(n, 0*X + u )
```




education
Z
regional/cultural
factors
u
X
Y
occupation
income

Good & Bad Controls

"Control" variable: Variable introduced to an analysis so that a causal estimate is possible

Heuristics fail — adding control variables can be worse than omitting

Make assumptions explicit

Course Schedule

Week 1	Bayesian inference	Chapters 1, 2, 3
Week 2	Linear models & Causal Inference	Chapter 4
Week 3	Causes, Confounds & Colliders	Chapters 5 & 6
Week 4	Overfitting / MCMC	Chapters 7, 8, 9
Week 5	Generalized Linear Models	Chapters 10, 11
Week 6	Integers & Other Monsters	Chapters 11 & 12
Week 7	Multilevel models I	Chapter 13
Week 8	Multilevel models II	Chapter 14
Week 9	Measurement & Missingness	Chapter 15
Week 10	Generalized Linear Madness	Chapter 16

`https://github.com/rmcelreath/stat_rethinking_2023`

BONUS

TABLE 2—ESTIMATED PROBIT MODELS FOR THE USE OF A SCREEN

	Preliminaries blind		Finals blind
	(1)	(2)	(3)
(Proportion female)$_{t-1}$	2.744	3.120	0.490
	(3.265)	(3.271)	(1.163)
	[0.006]	[0.004]	[0.011]
(Proportion of orchestra personnel with <6 years tenure)$_{t-1}$	−26.46	−28.13	−9.467
	(7.314)	(8.459)	(2.787)
	[−0.058]	[−0.039]	[−0.207]
"Big Five" orchestra		0.367	
		(0.452)	
		[0.001]	
pseudo R^2	0.178	0.193	0.050
Number of observations	294	294	434

Table 2 Fallacy

Not all coefficients are causal effects

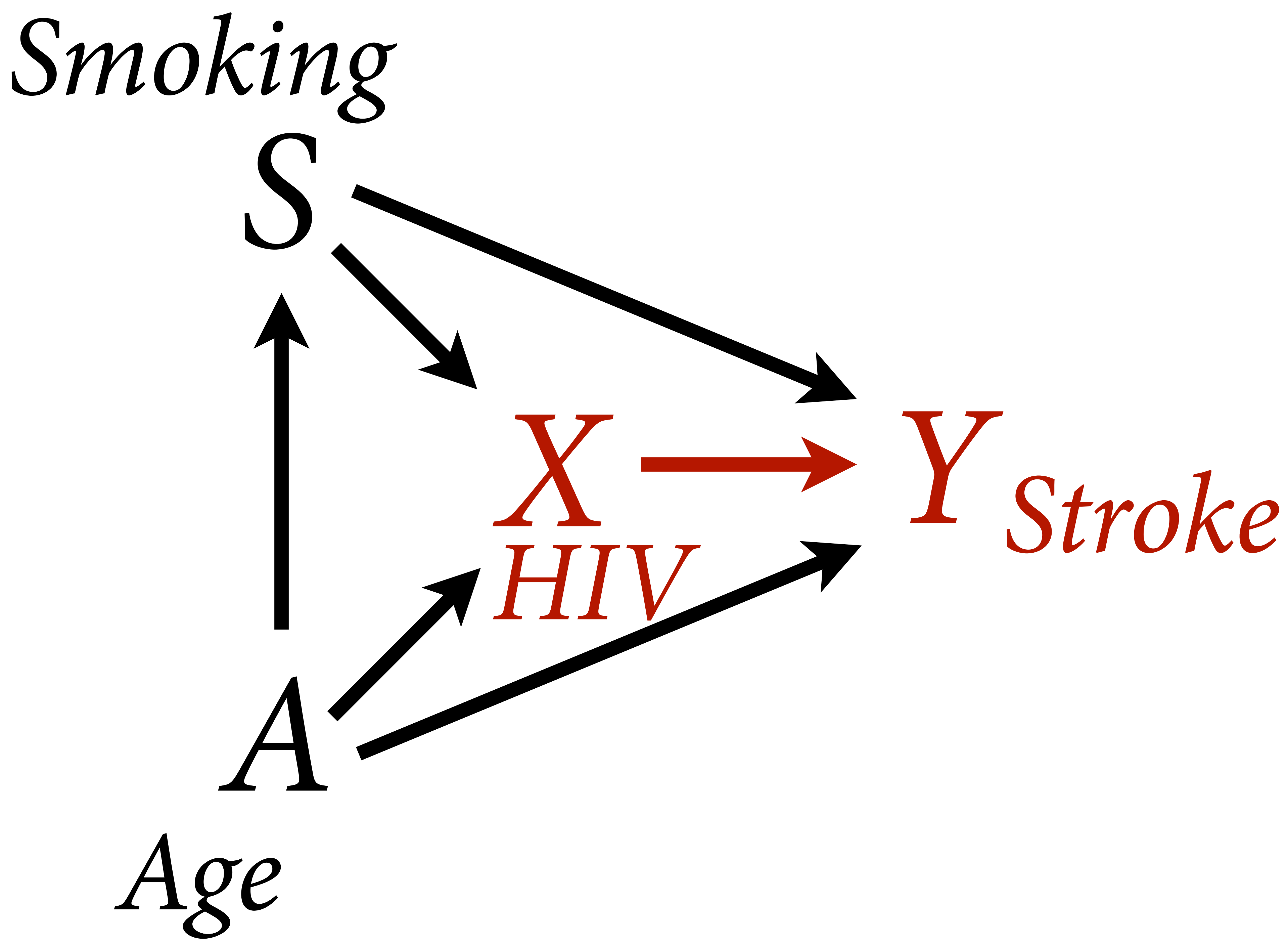
Statistical model designed to identify ***X*** **–>** ***Y*** will not also identify effects of control variables

Table 2 is dangerous

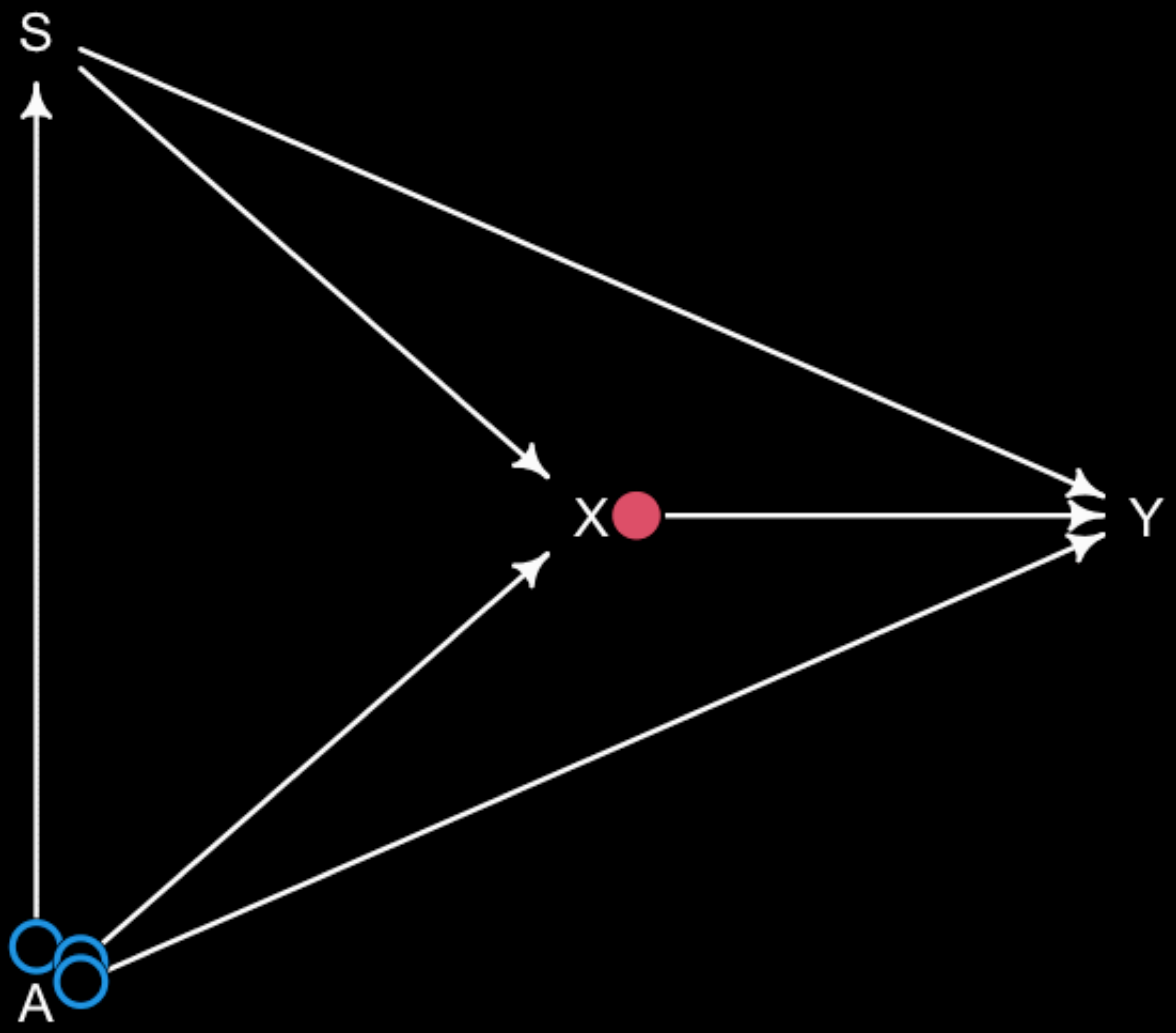
TABLE 2—ESTIMATED PROBIT MODELS FOR THE USE OF A SCREEN

	Preliminaries blind		Finals blind
	(1)	(2)	(3)
(Proportion female)$_{t-1}$	2.744	3.120	0.490
	(3.265)	(3.271)	(1.163)
	[0.006]	[0.004]	[0.011]
(Proportion of orchestra personnel with <6 years tenure)$_{t-1}$	−26.46	−28.13	−9.467
	(7.314)	(8.459)	(2.787)
	[−0.058]	[−0.039]	[−0.207]
"Big Five" orchestra		0.367	
		(0.452)	
		[0.001]	
pseudo R^2	0.178	0.193	0.050
Number of observations	294	294	434

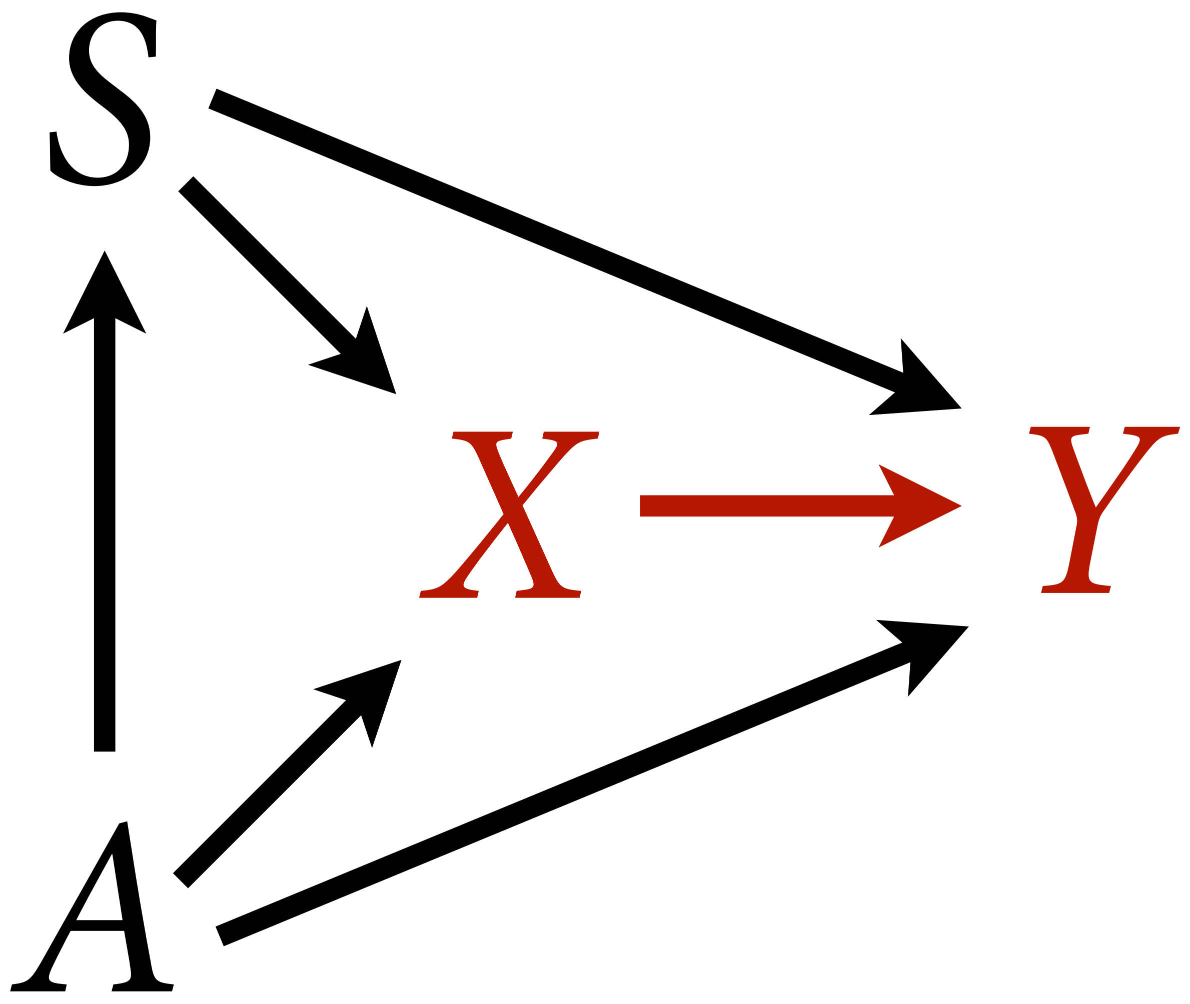
Westreich & Greenland 2013 The Table 2 Fallacy



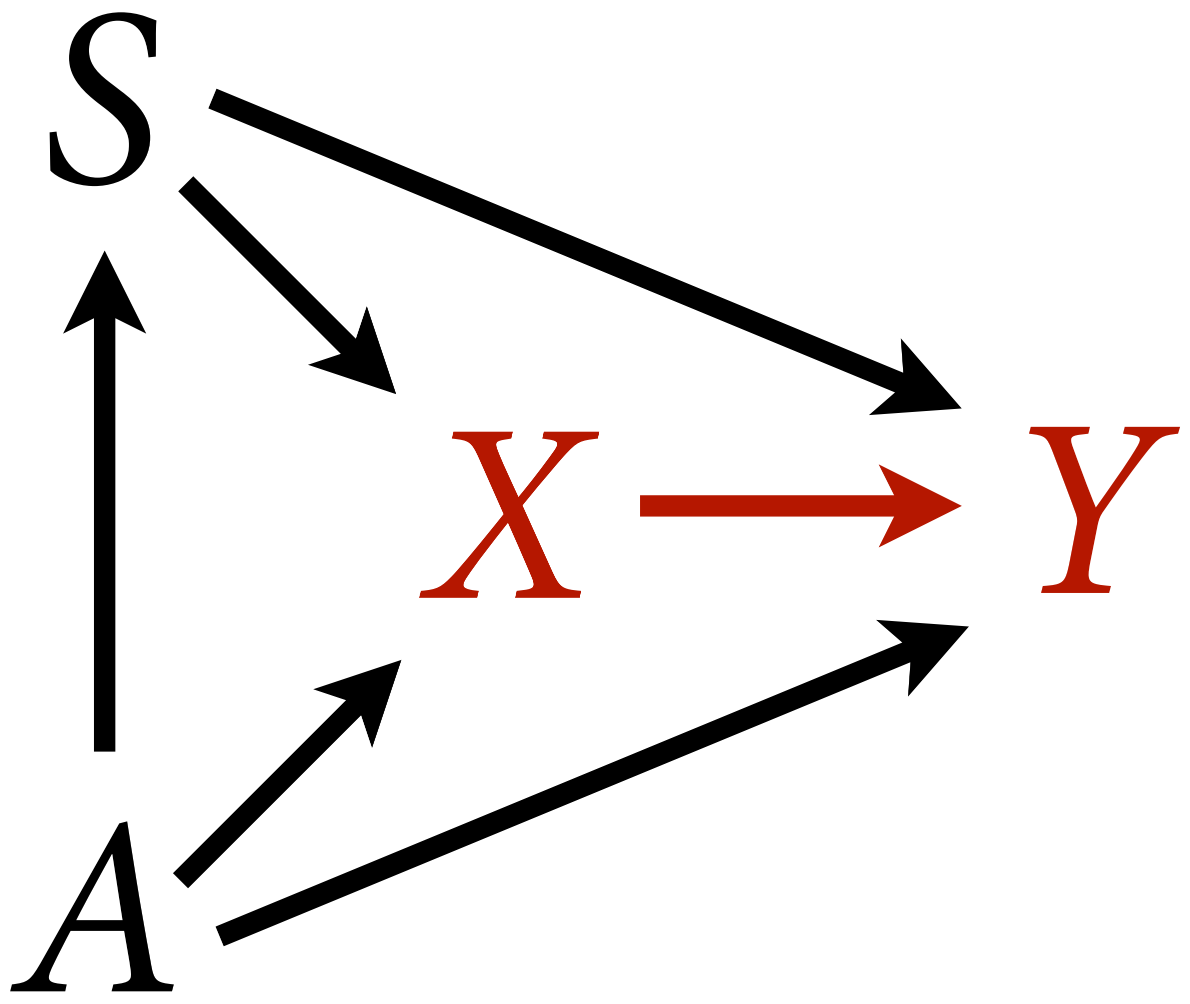
Westreich & Greenland 2013 The Table 2 Fallacy


S
X
Y
A

Use Backdoor Criterion

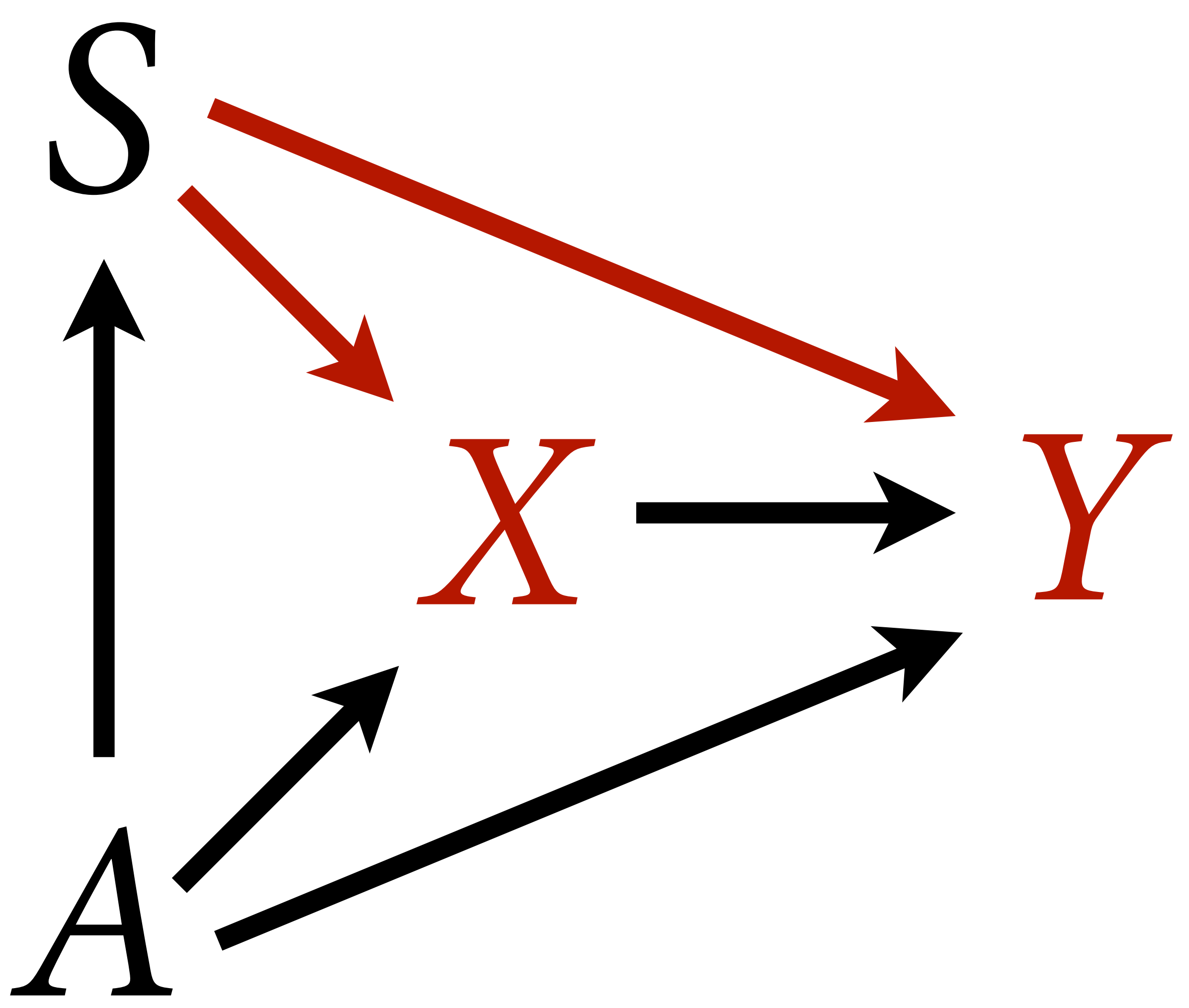


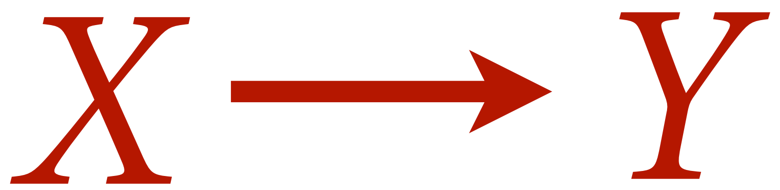
Use Backdoor Criterion

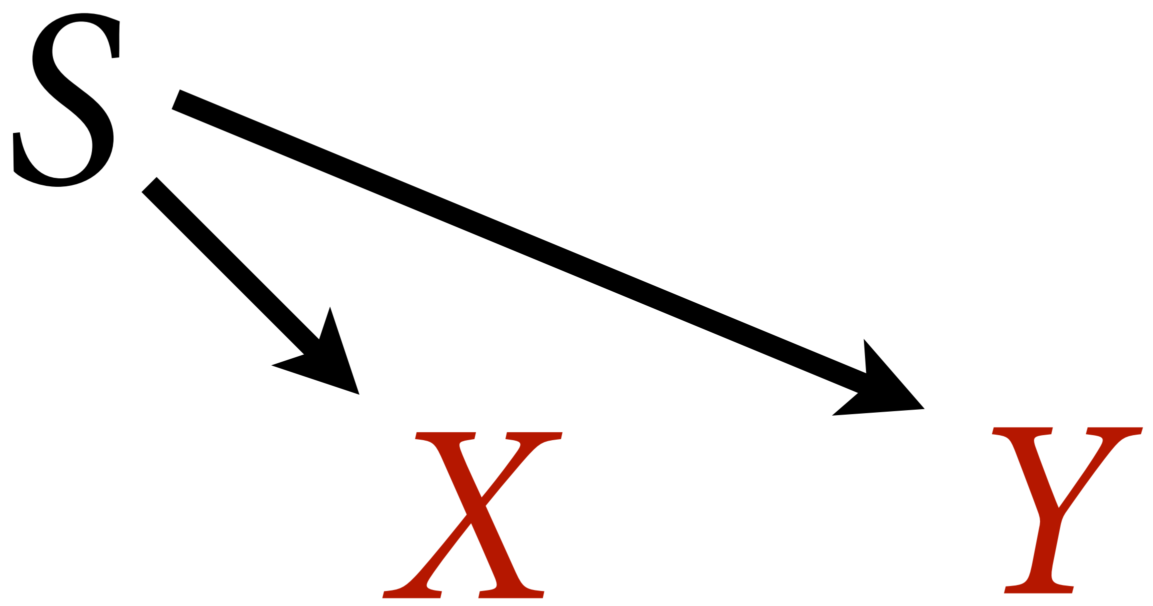


$X \longrightarrow Y$

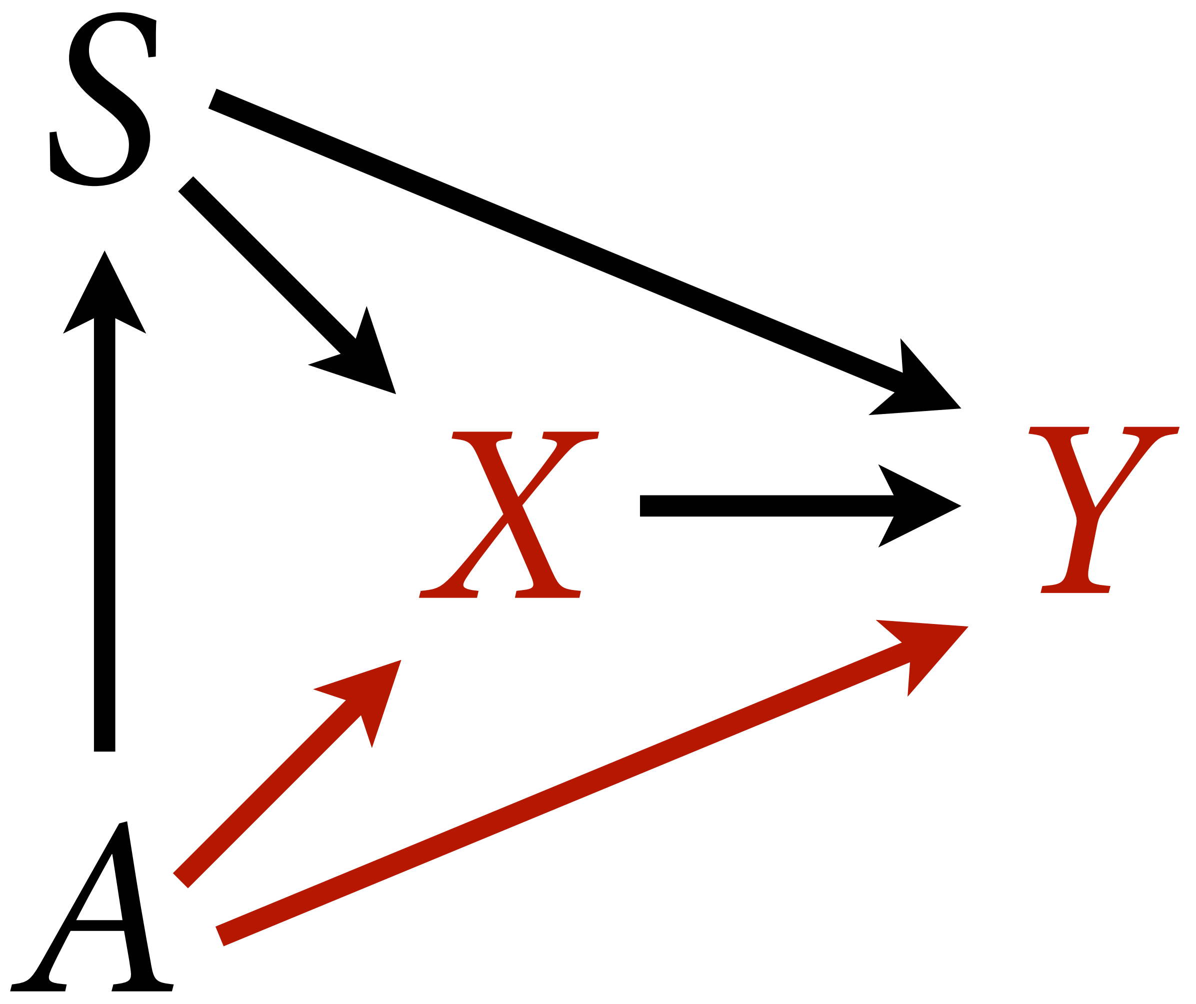
Use Backdoor Criterion

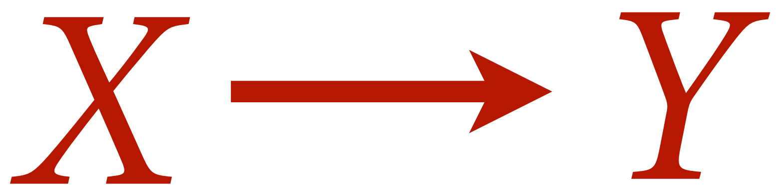


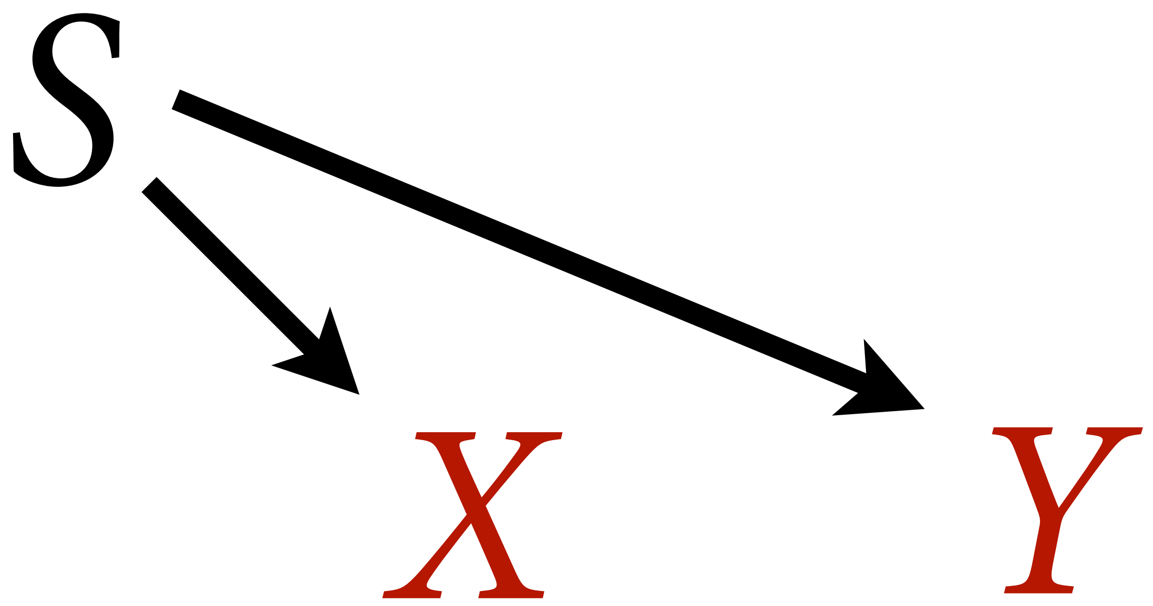


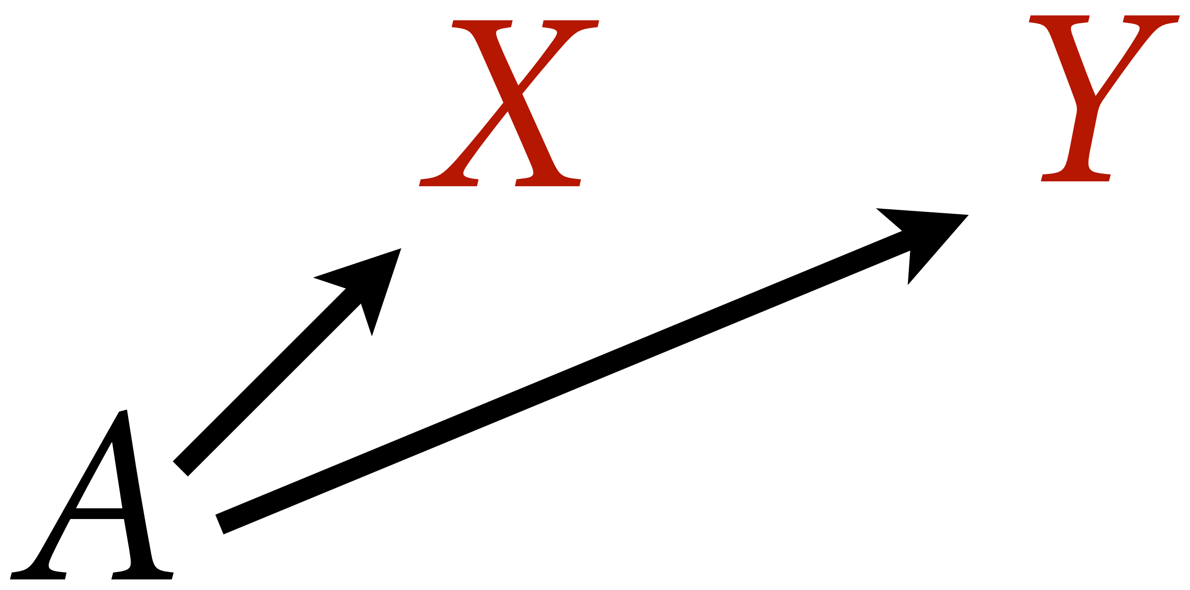


Use Backdoor Criterion

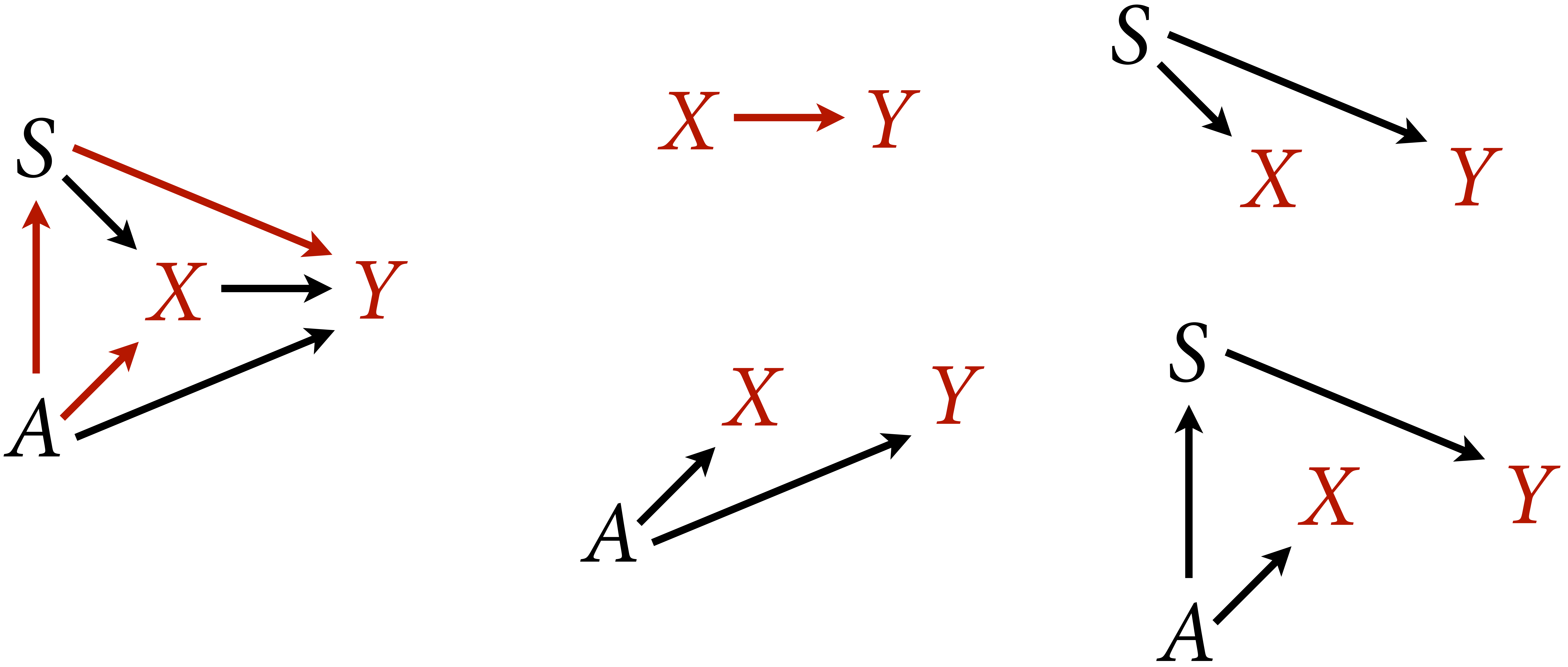




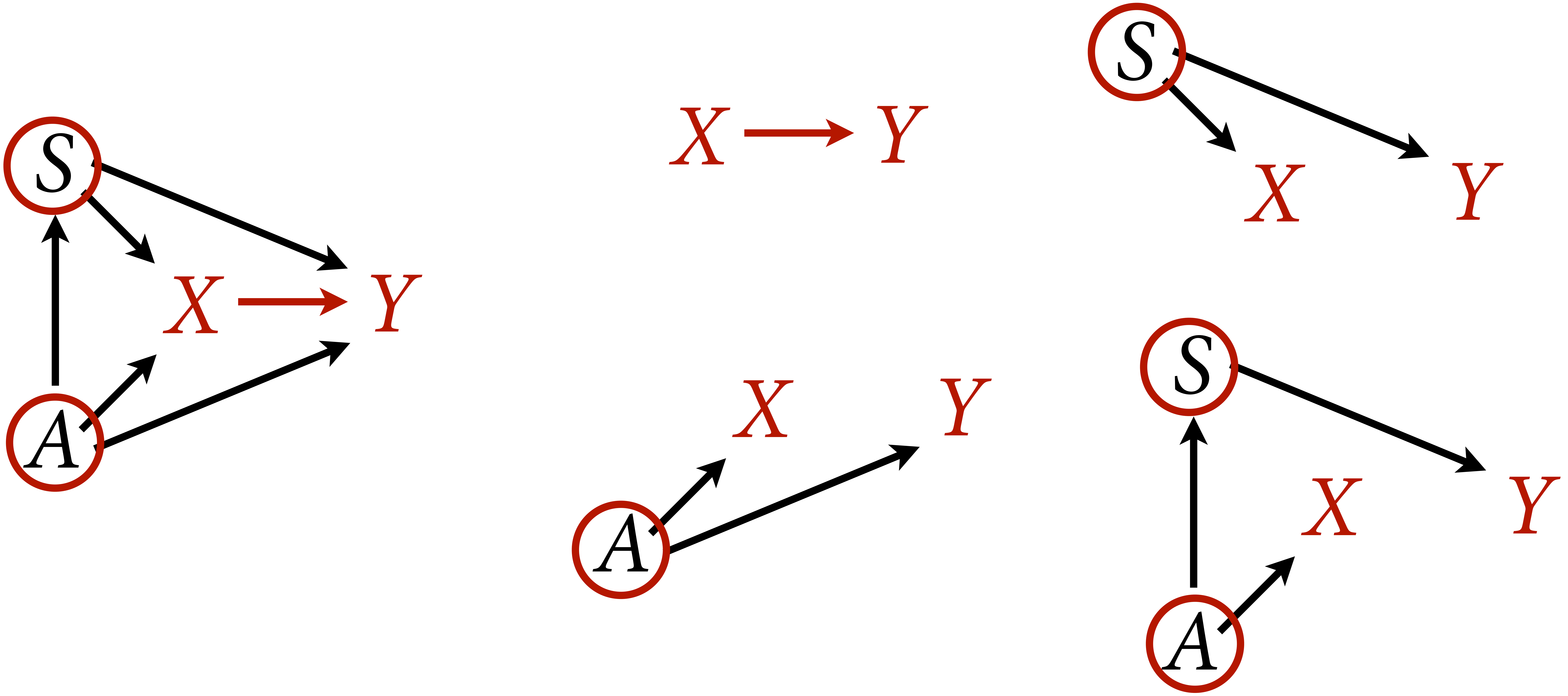


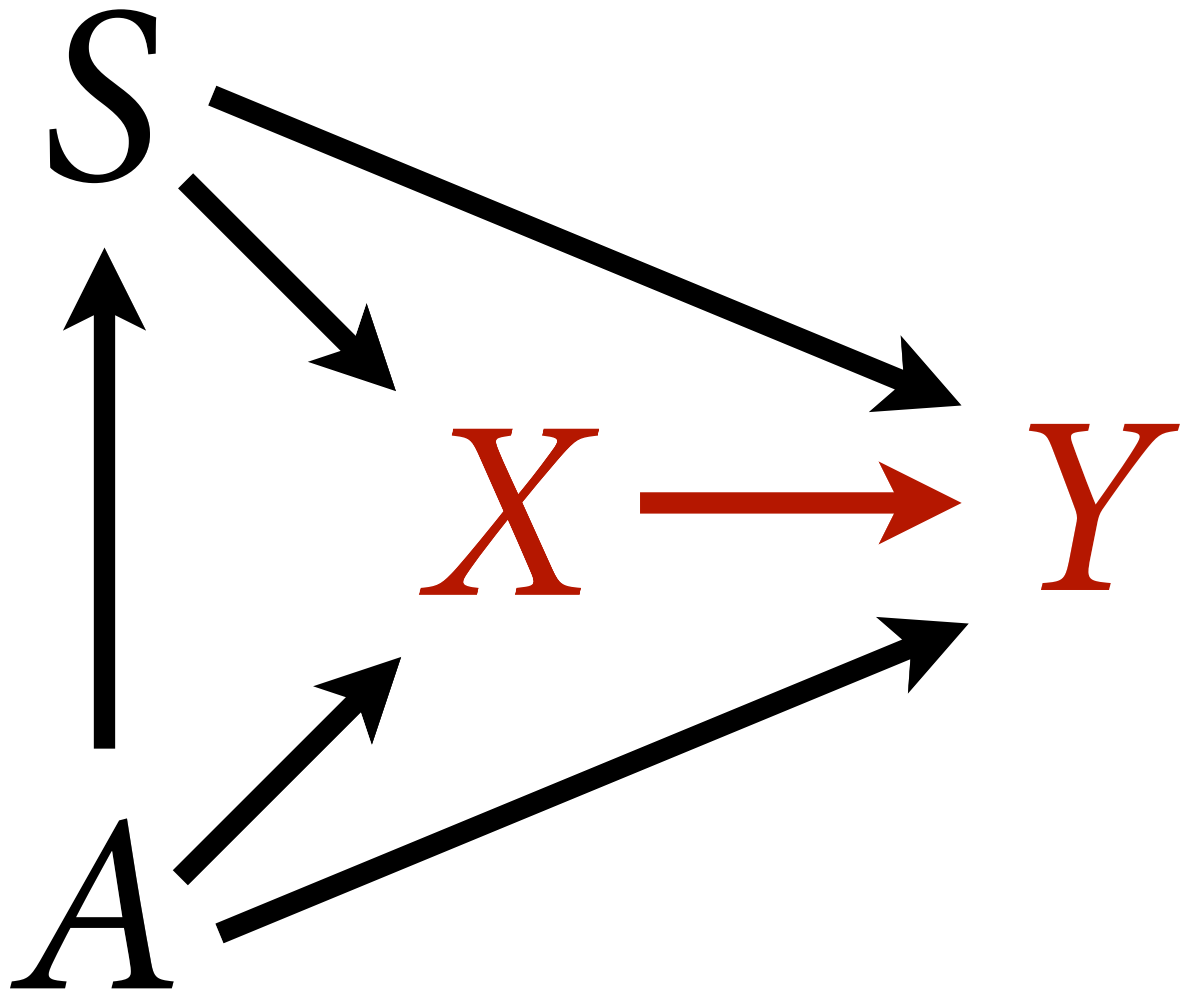


Use Backdoor Criterion



Use Backdoor Criterion

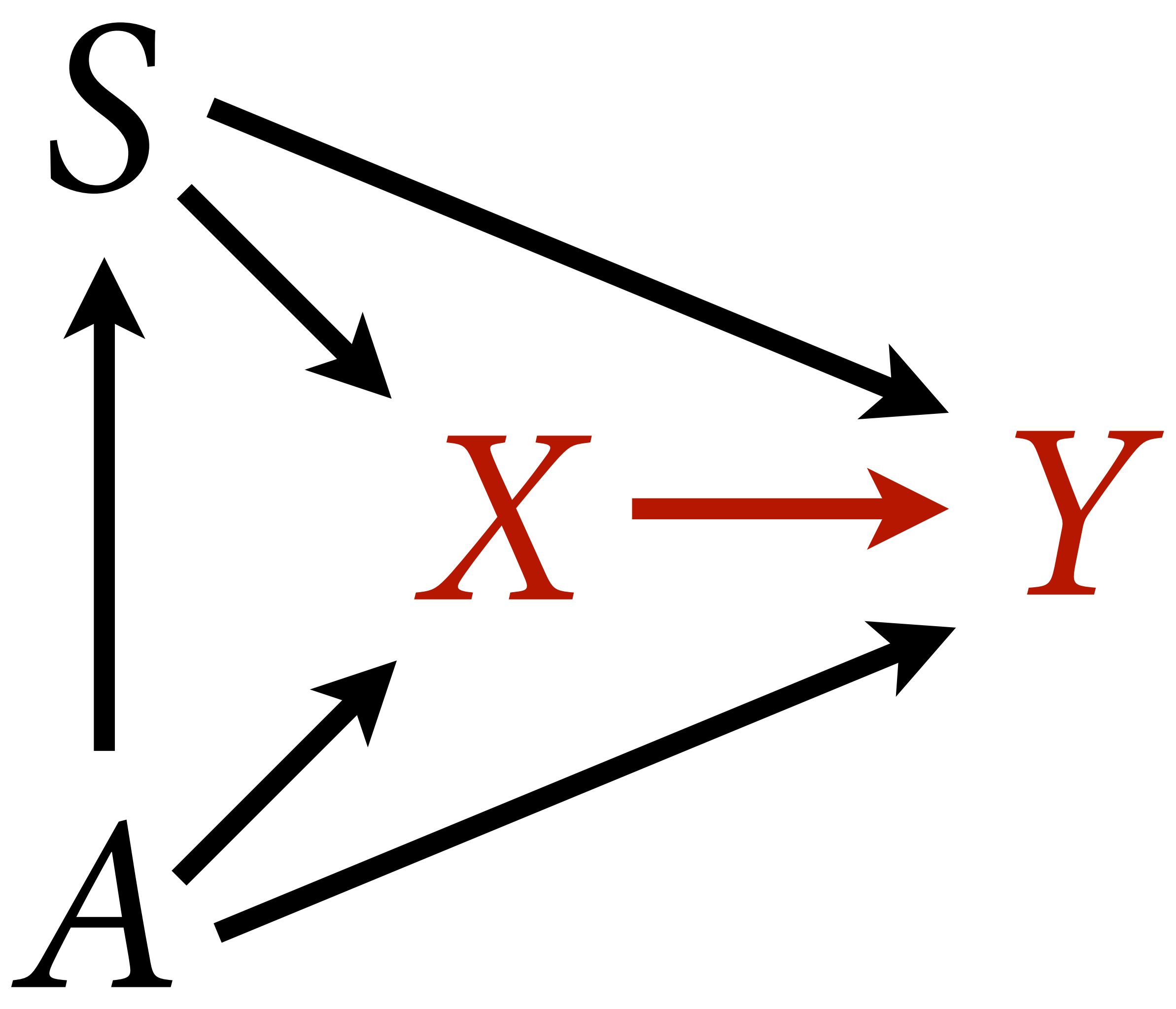




$$Y_i \sim \text{Normal}(\mu_i, \sigma)$$
$$\mu_i = \alpha + \beta_X X_i + \beta_S S_i + \beta_A A_i$$

X

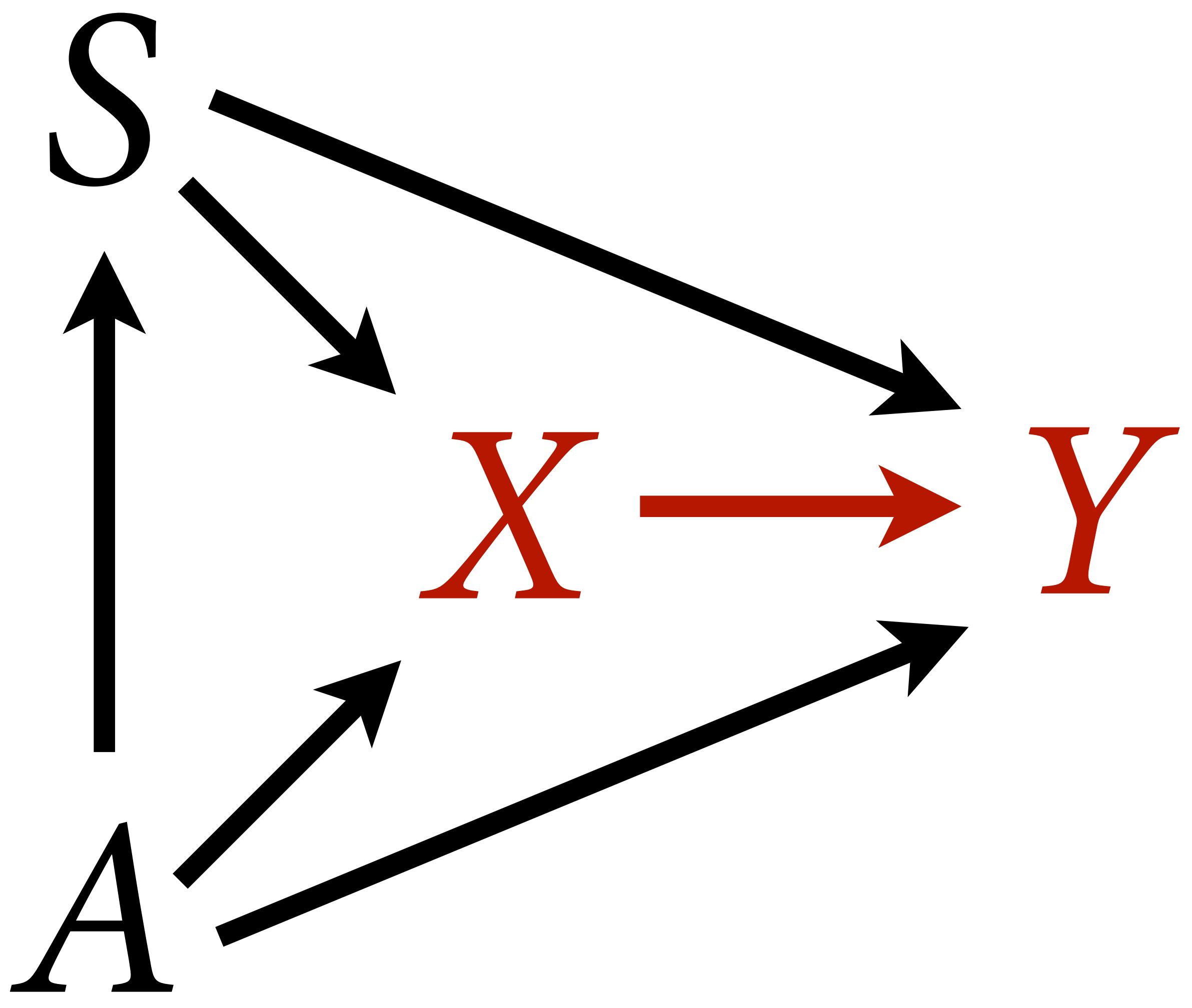
Unconditional



Confounded by A and S

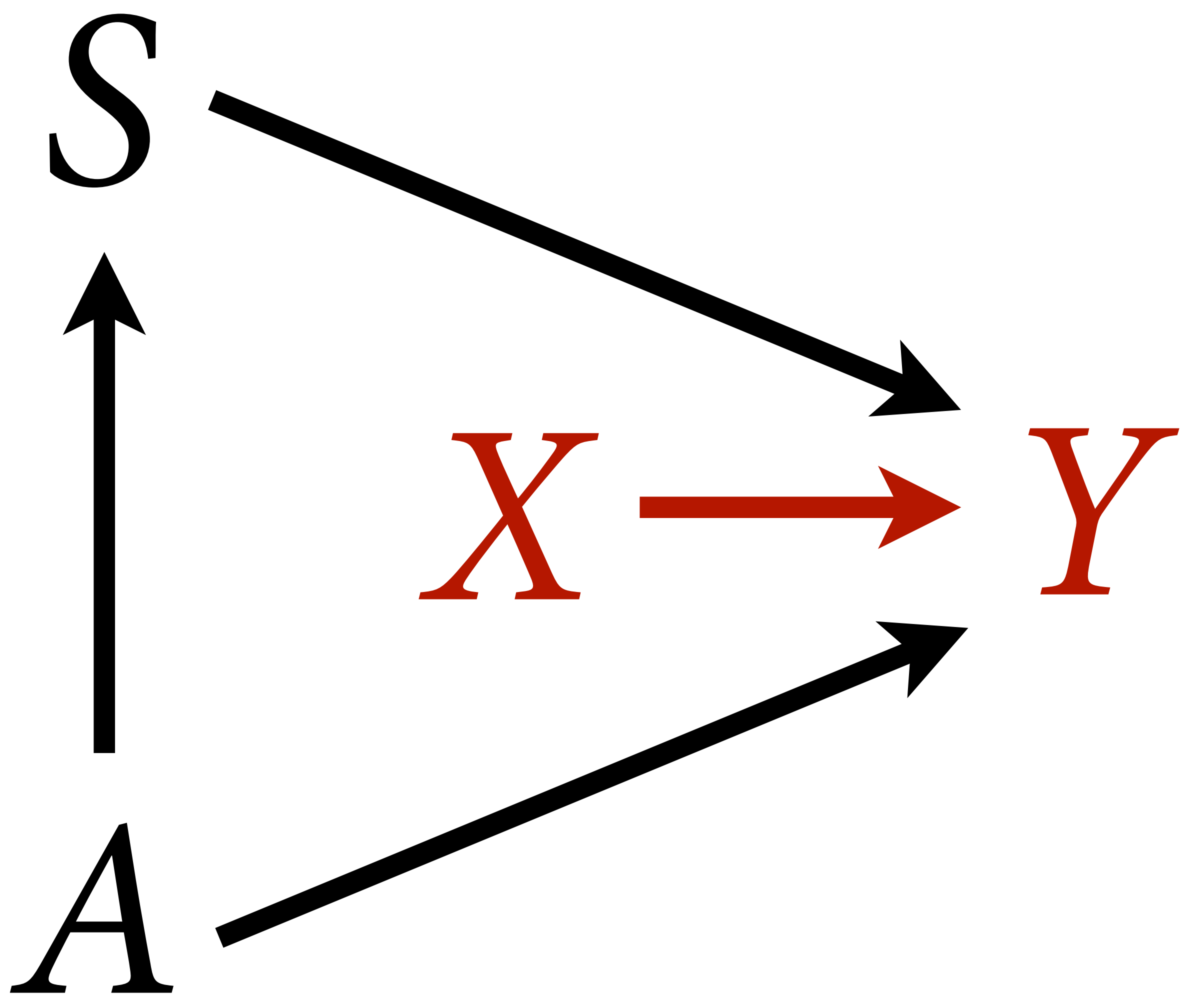
X

Unconditional



Confounded by A and S

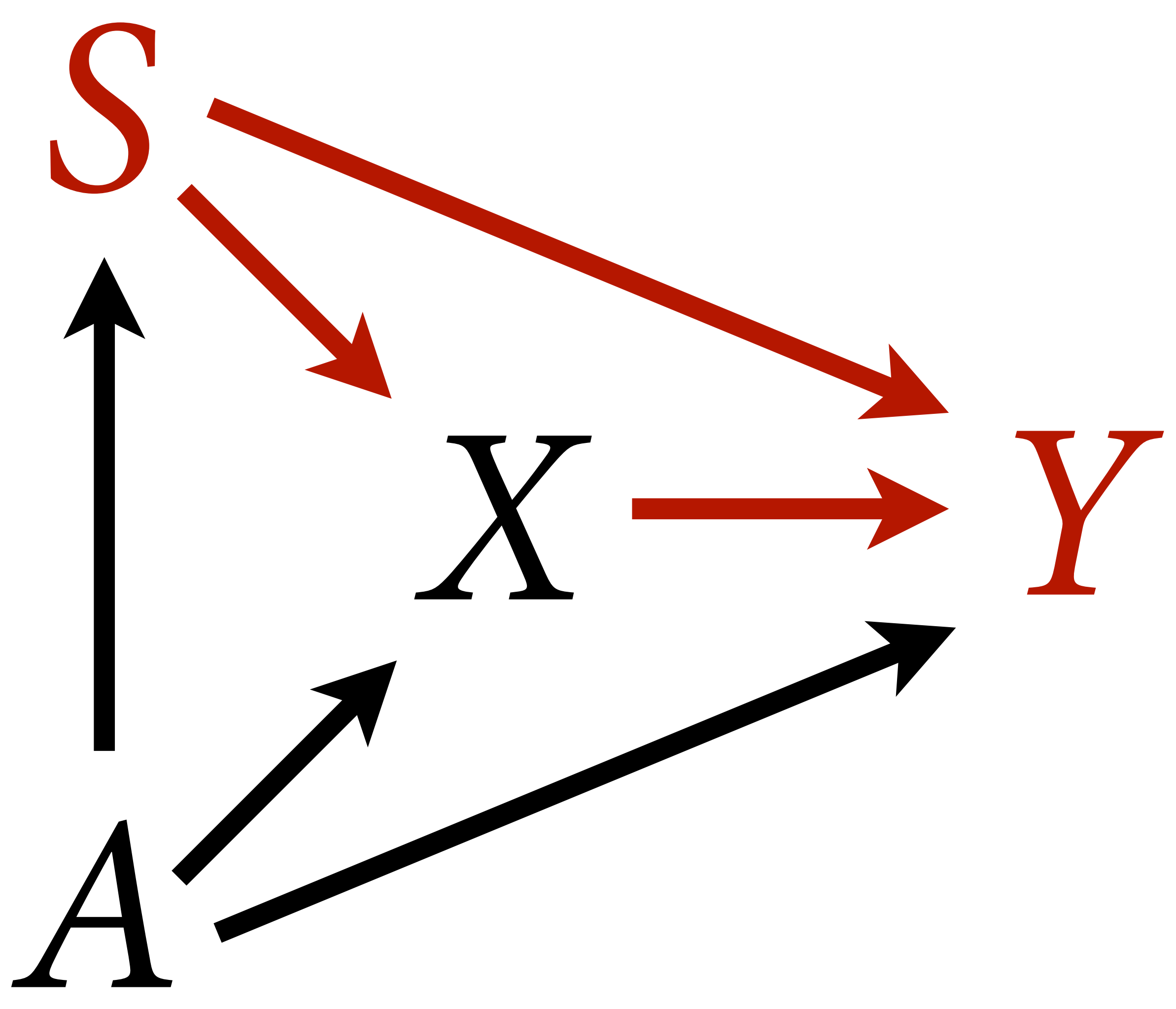
Conditional on A and S



Coefficient for ***X***:
Effect of X on Y
(still must marginalize!)

S

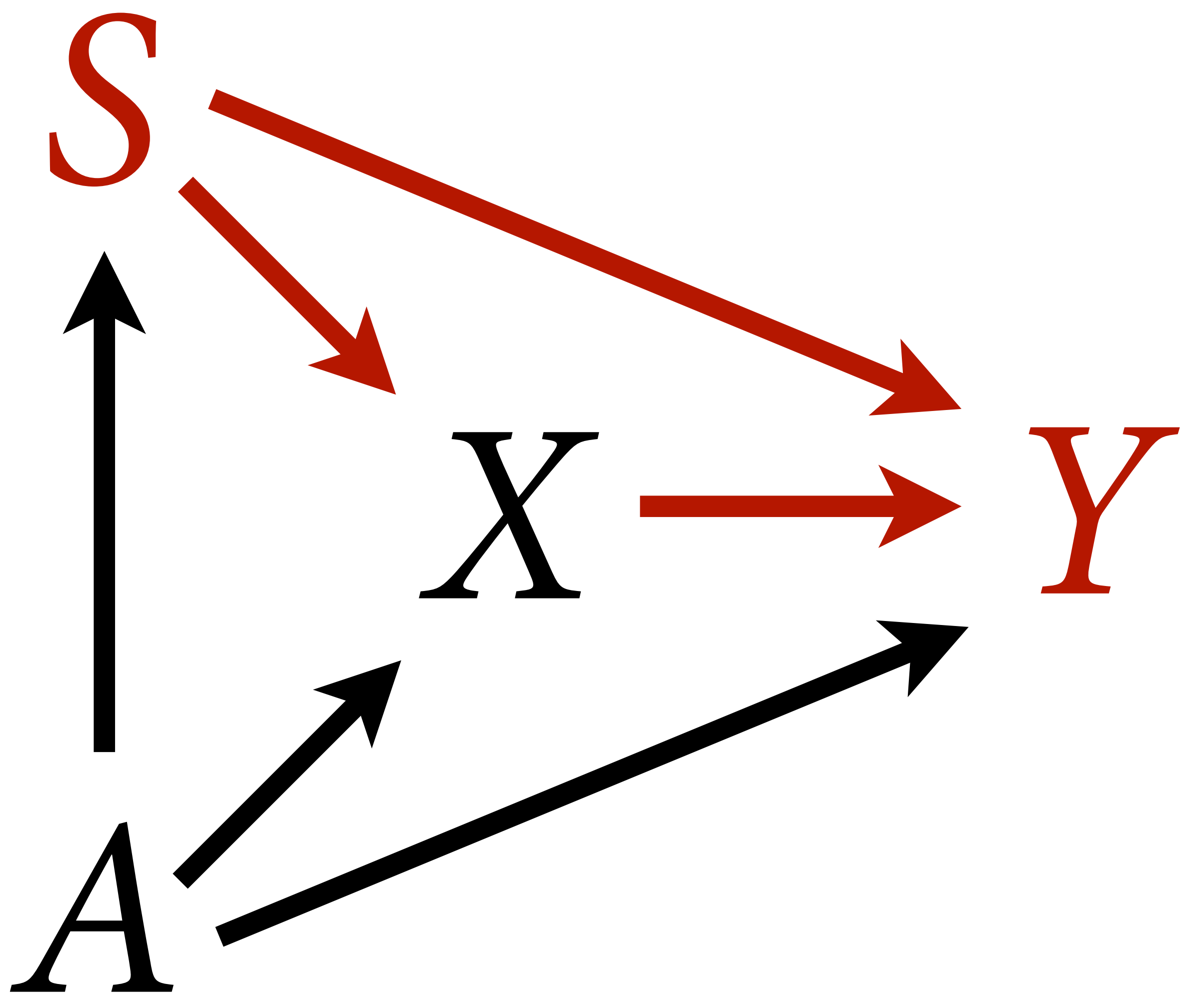
Unconditional



Effect of *S*
confounded by *A*

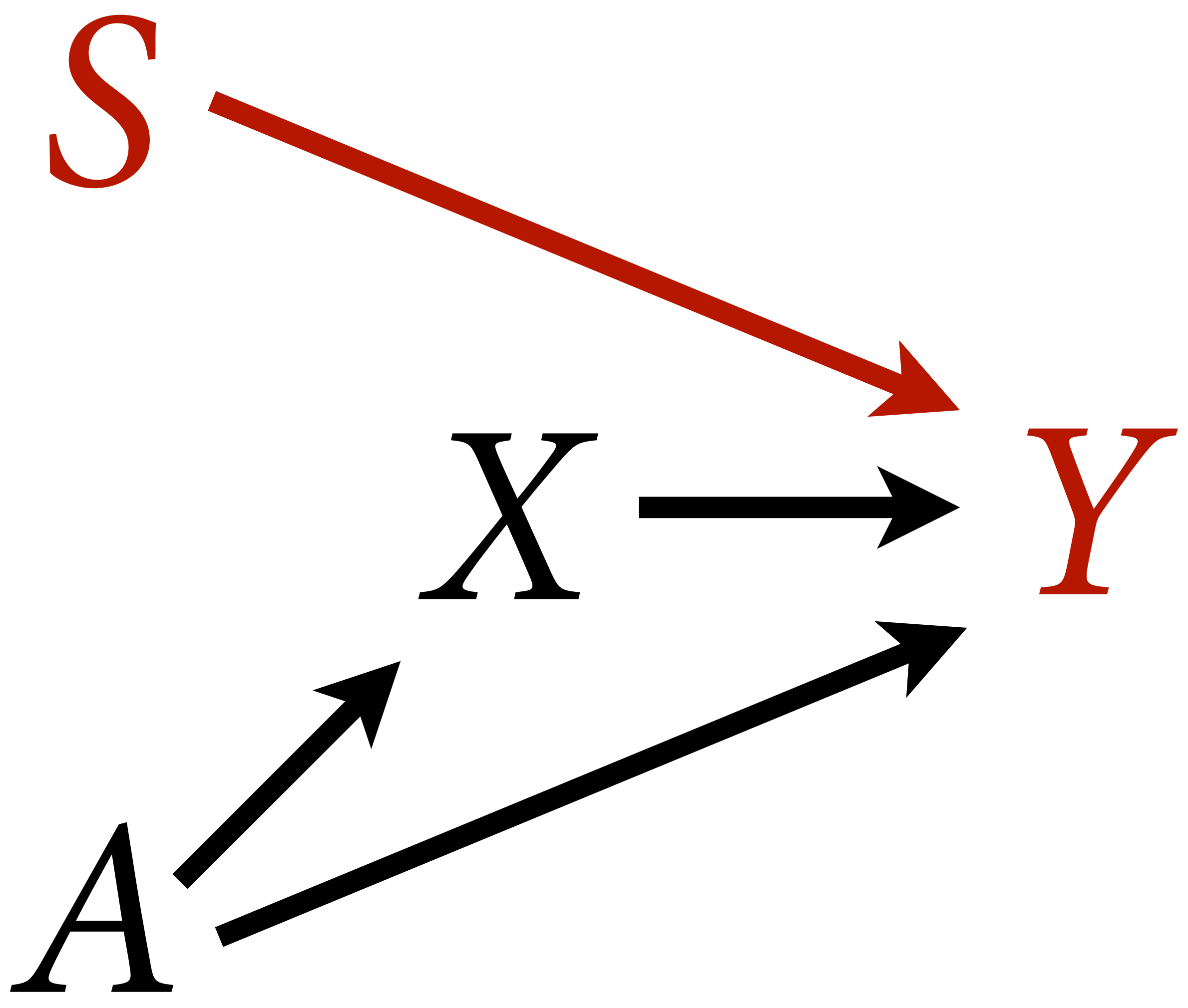
S

Unconditional



Effect of S
confounded by A

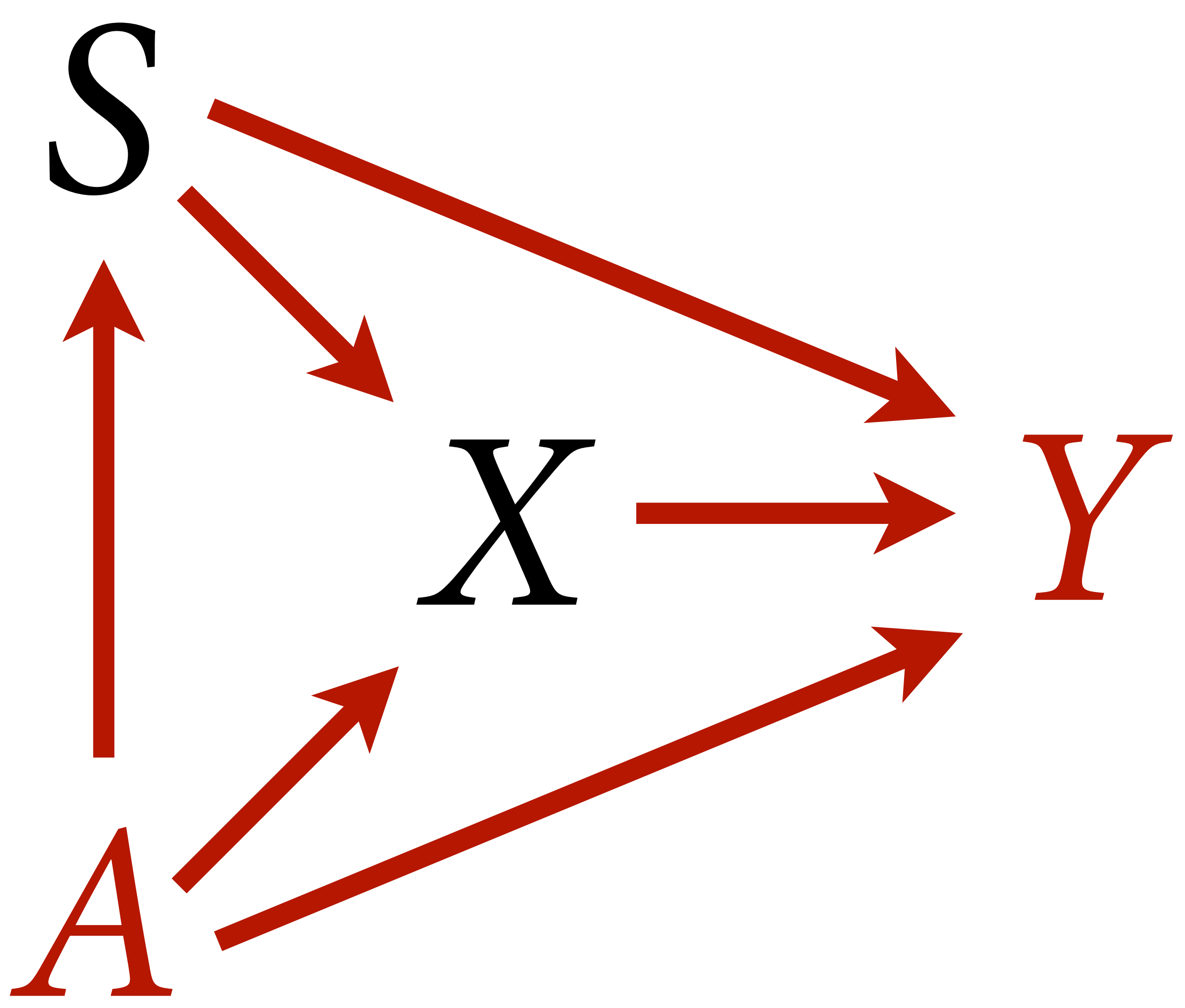
Conditional on A and X



Coefficient for **S**:
Direct effect of S on Y

A

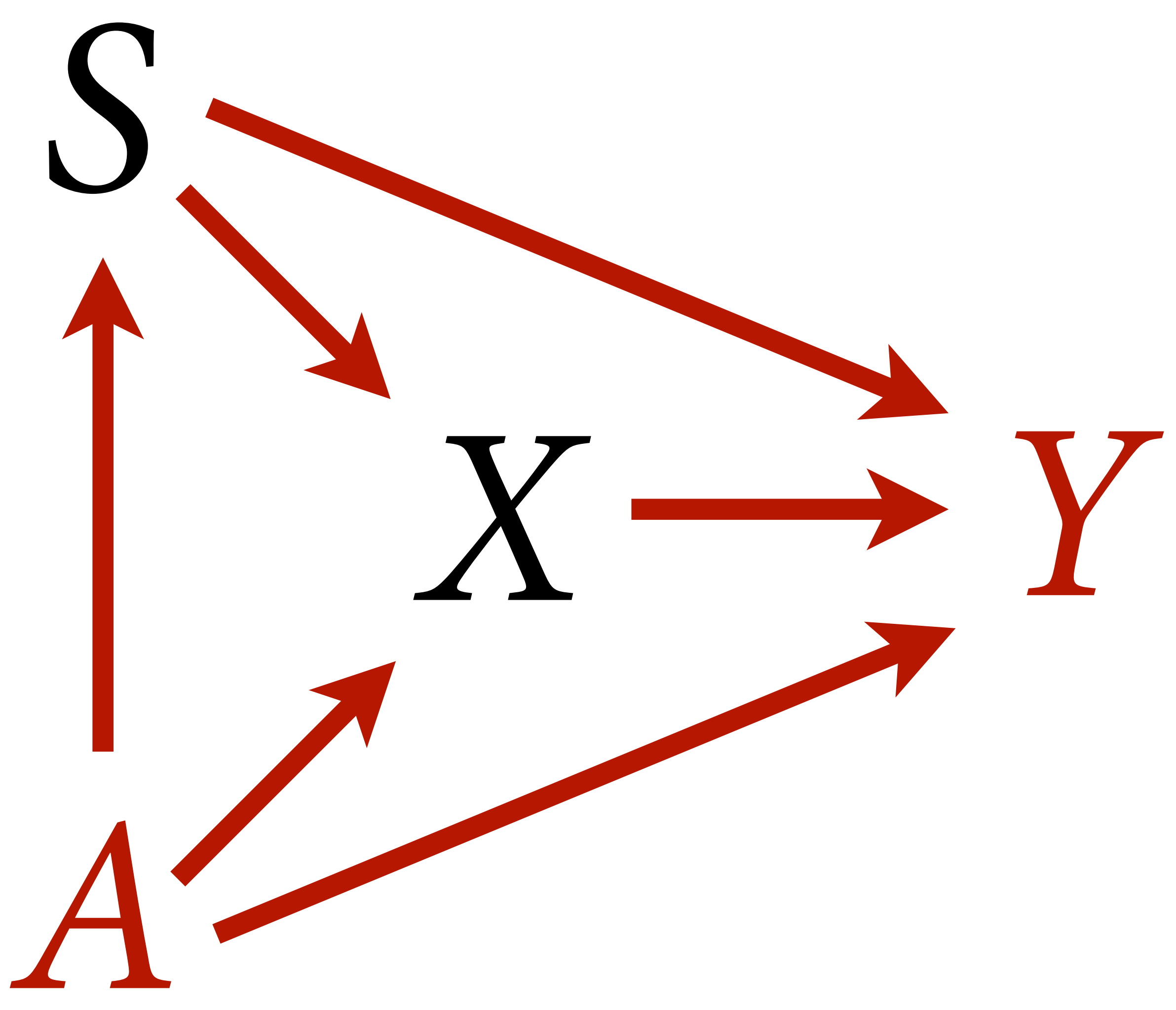
Unconditional



Total causal effect of A on Y flows through all paths

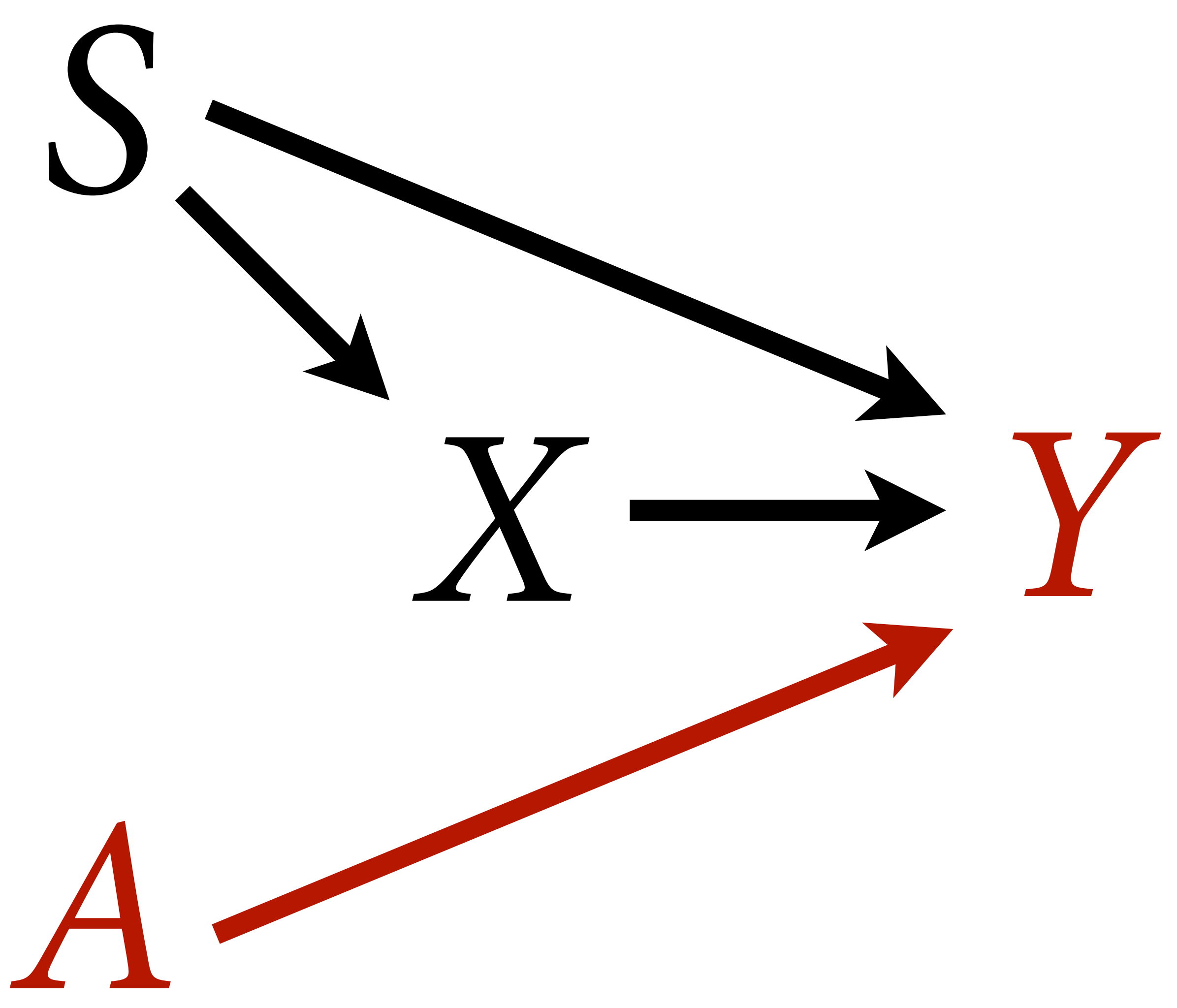
A

Unconditional



Total causal effect of *A* on *Y* flows through all paths

Conditional on X and S



Coefficient for ***A***:
Direct effect of *A* on *Y*

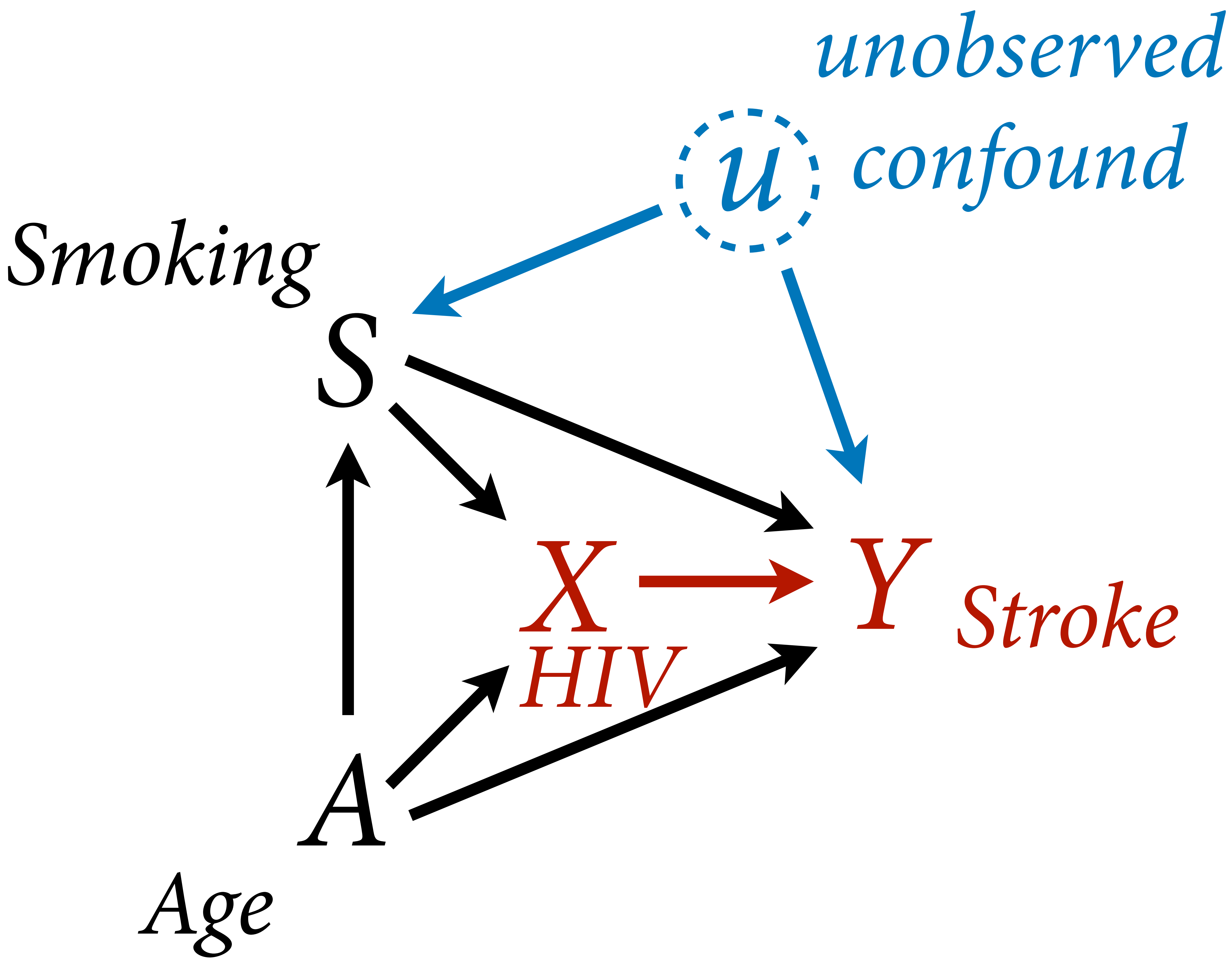

unobserved
confound
u
Smoking
S
X
HIV
Y
Stroke
A
Age

Table 2 Fallacy

Not all coefficients created equal

So do not present them as equal

Options:
Do not present control coefficients
Give explicit interpretation of each

No interpretation without causal representation