Does X cause Y?

Lord's Paradox

- a rationale for Longitudinal Data analysis

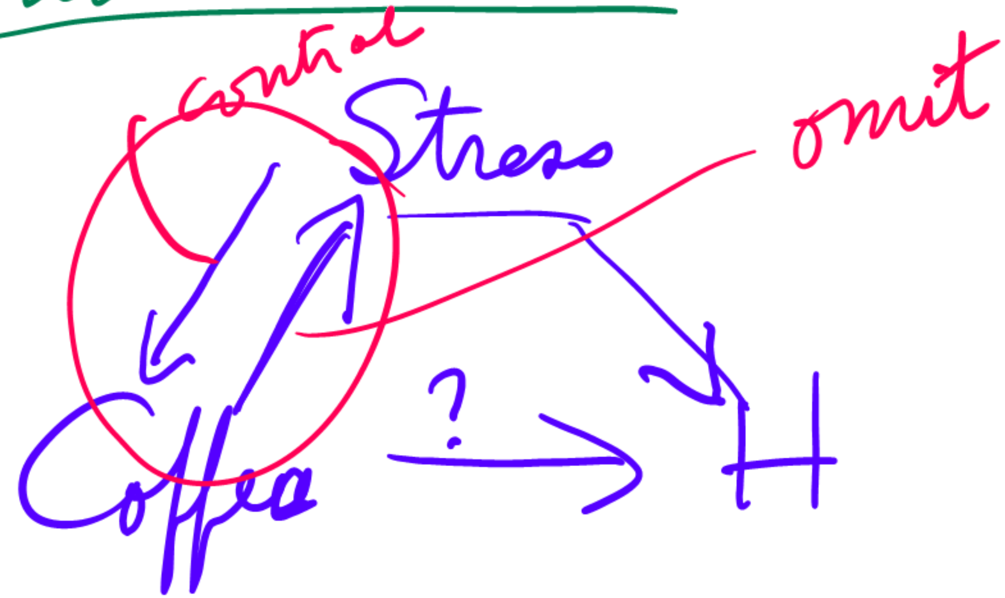
3 Qs?

1) Causal – predictive

random assignment

2) Experimental – Observational

3) Random selection.



Stres

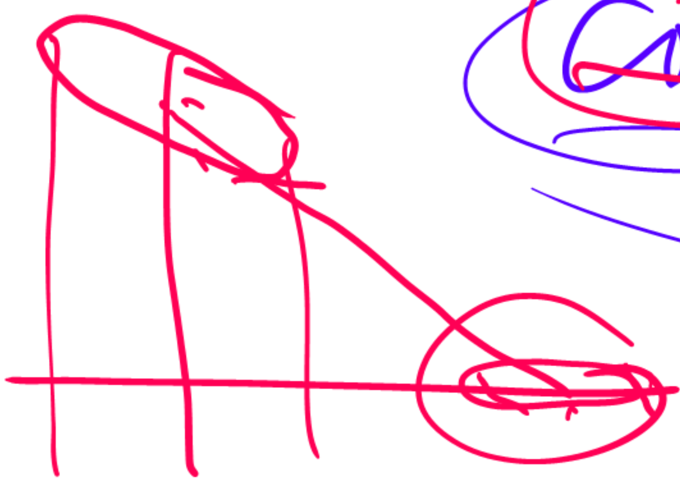
Confounding

Stress

Coffee

Coffee → H

Medi



Coffee → H

Coffee → Stress → H

mediating

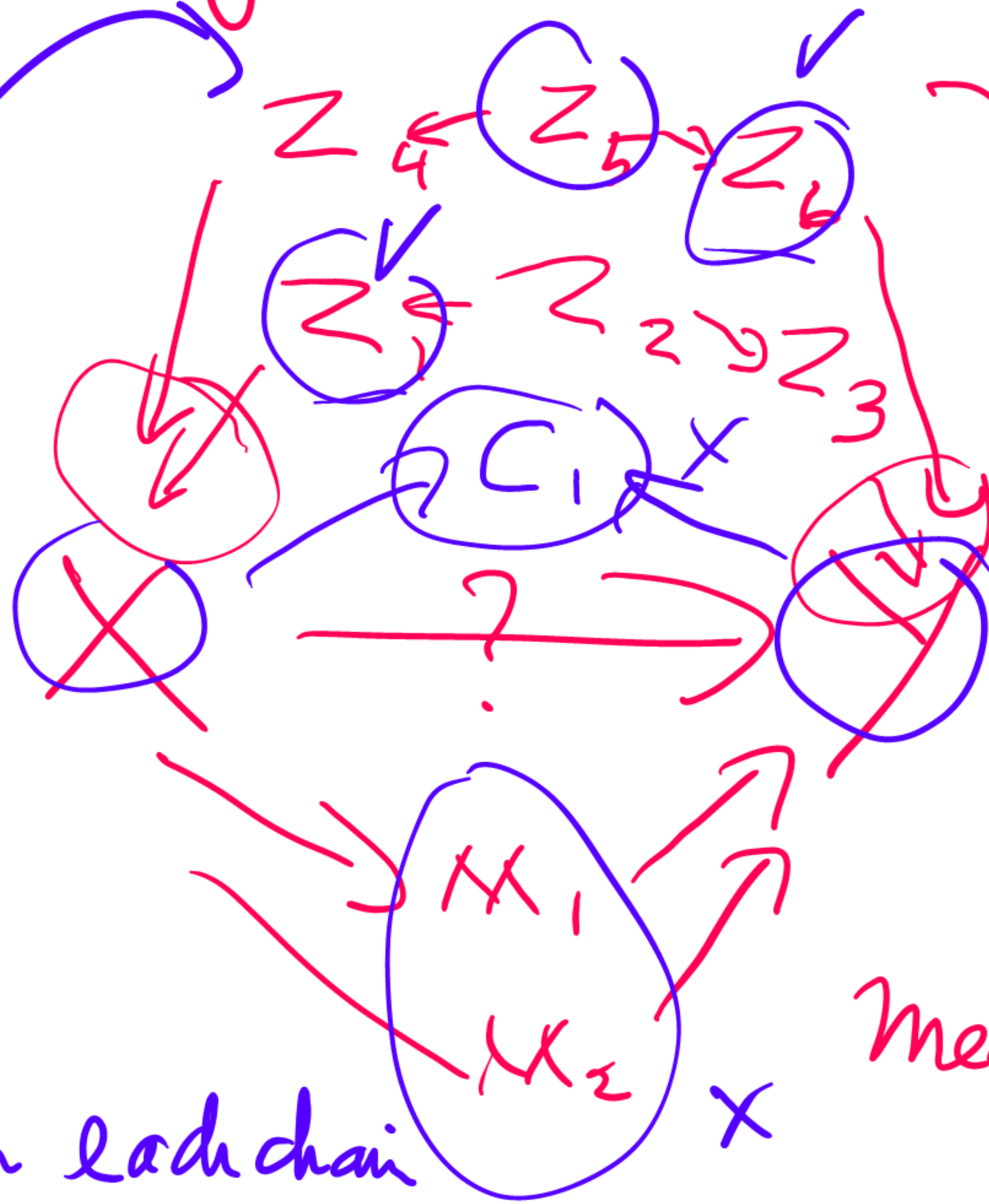
The Book of Why — Judea Pearl (2018)

Est causal effect of X on Y

"Backdoor criterion"

Block each path

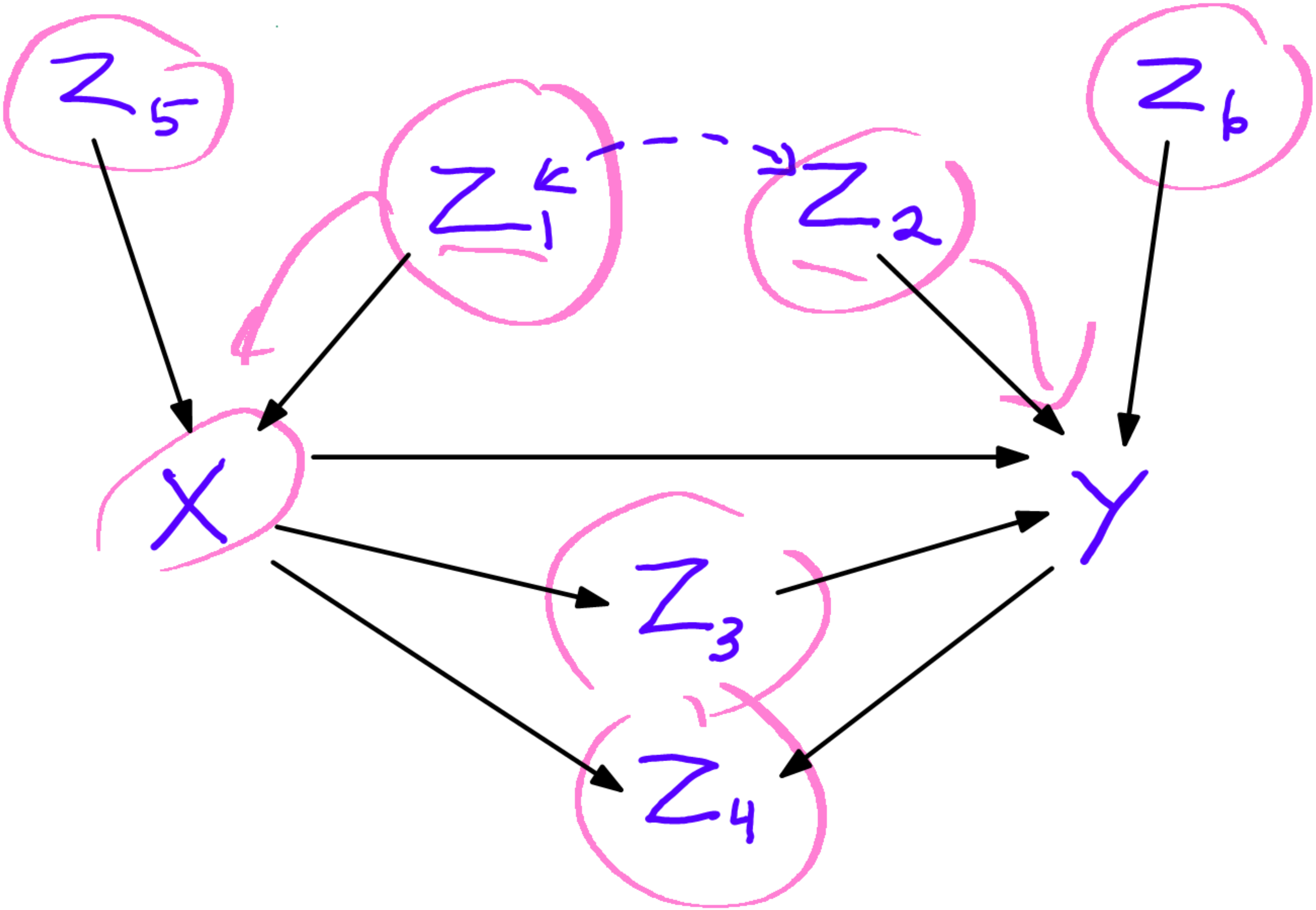
control at least one var. in each chain

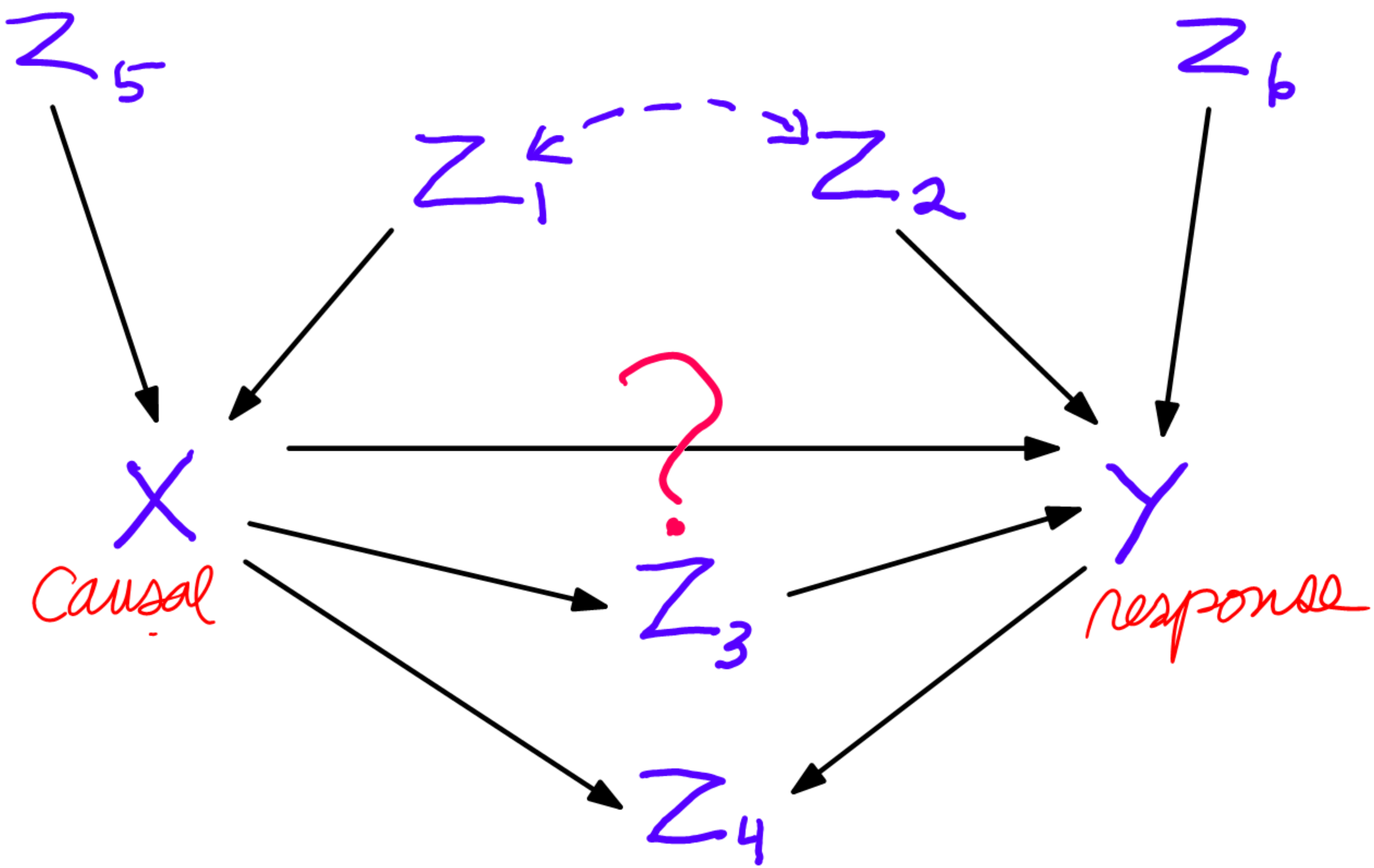


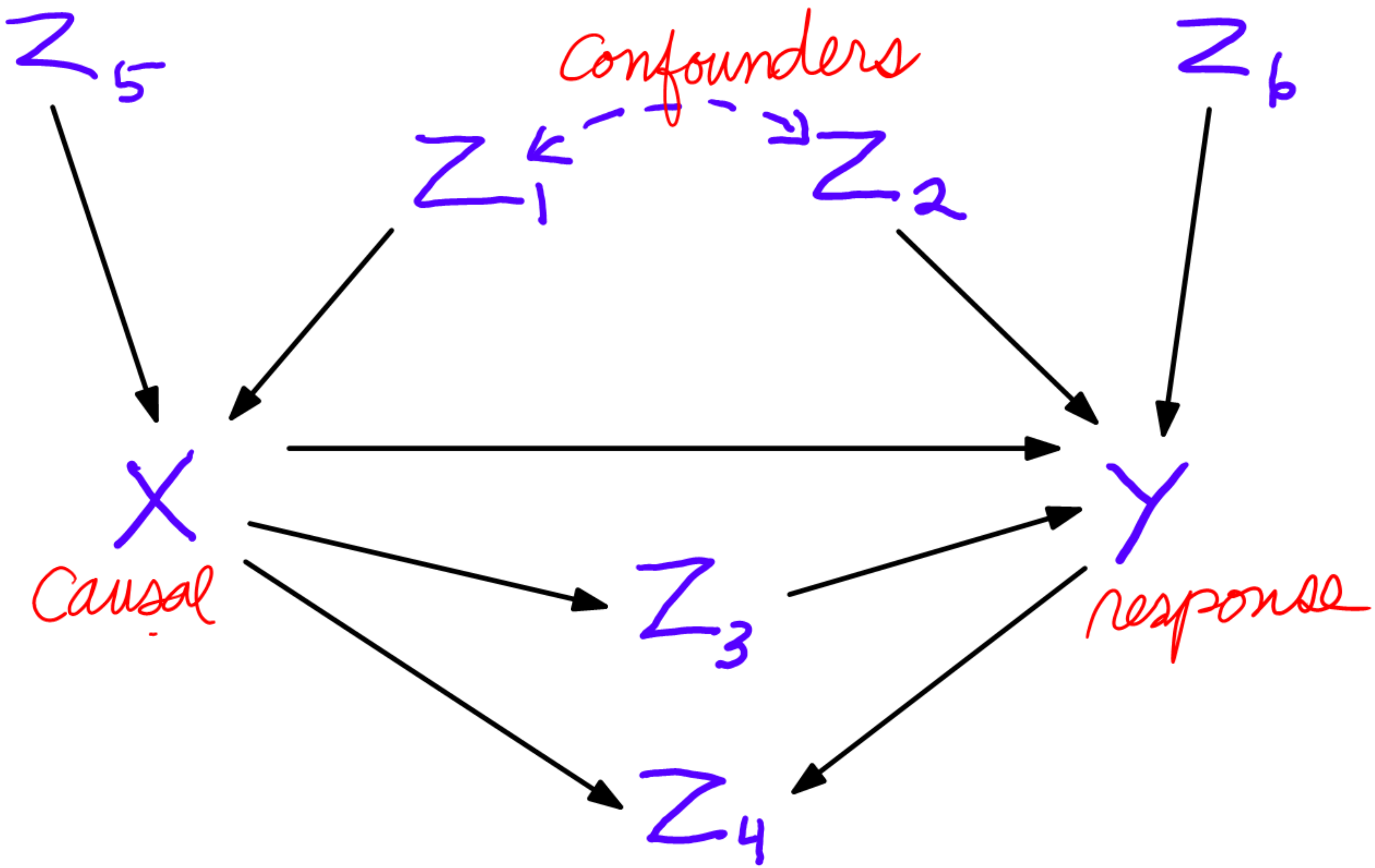
Confounder, chain

Mediators.

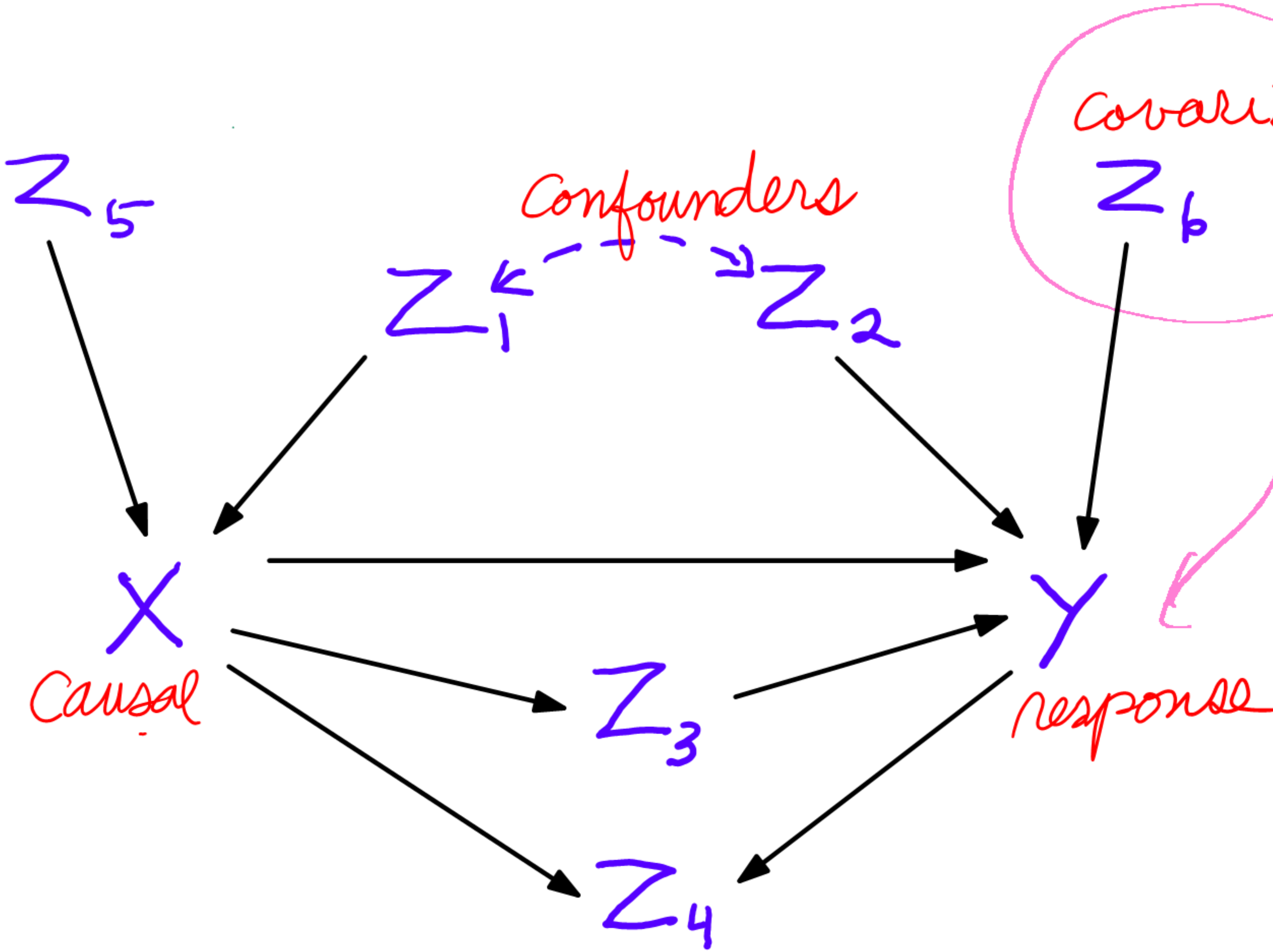
Causal graph



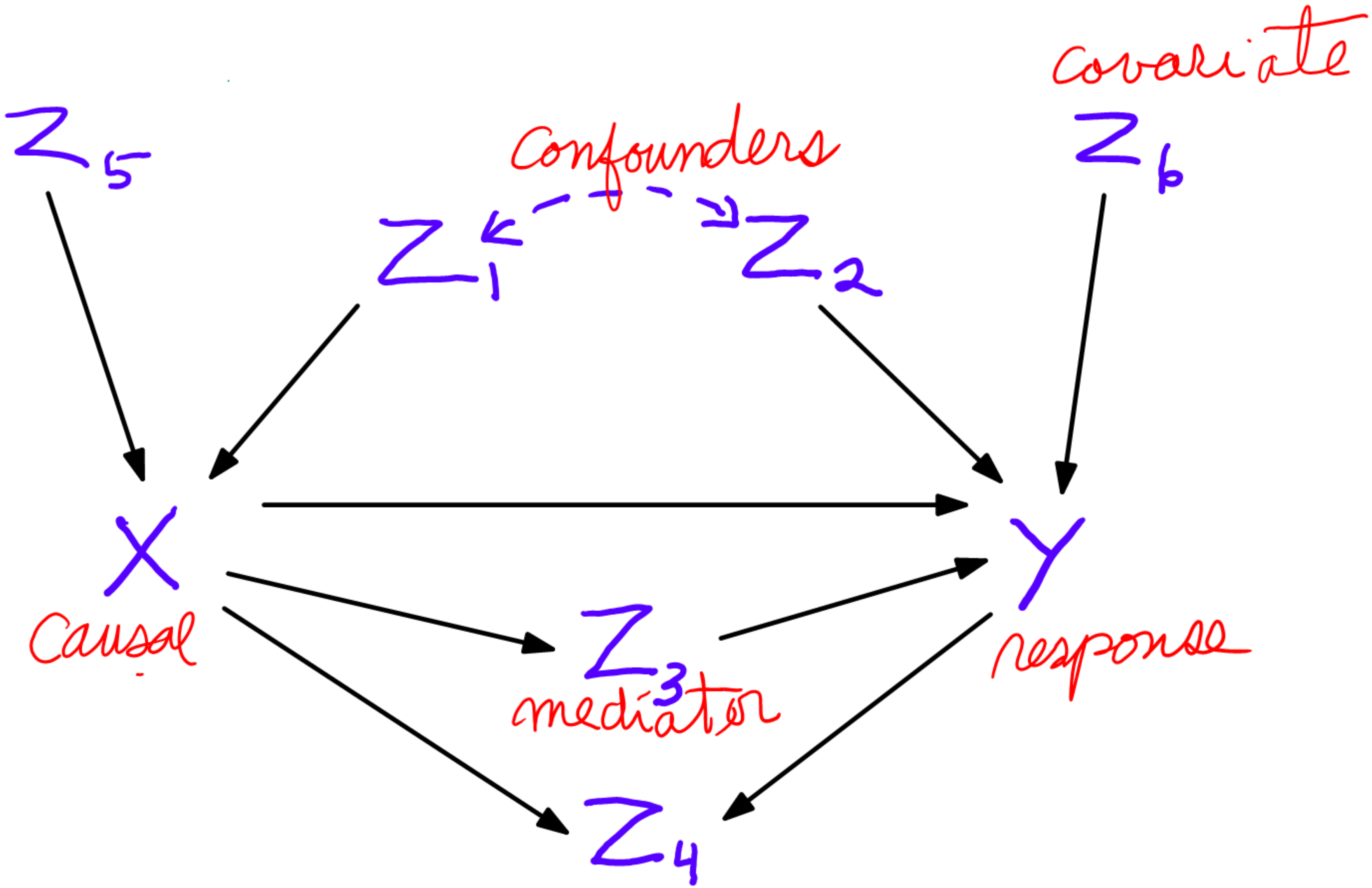

Z5
Z6
Z1
Z2
X
Y
Z3
Z4
?
causal
response

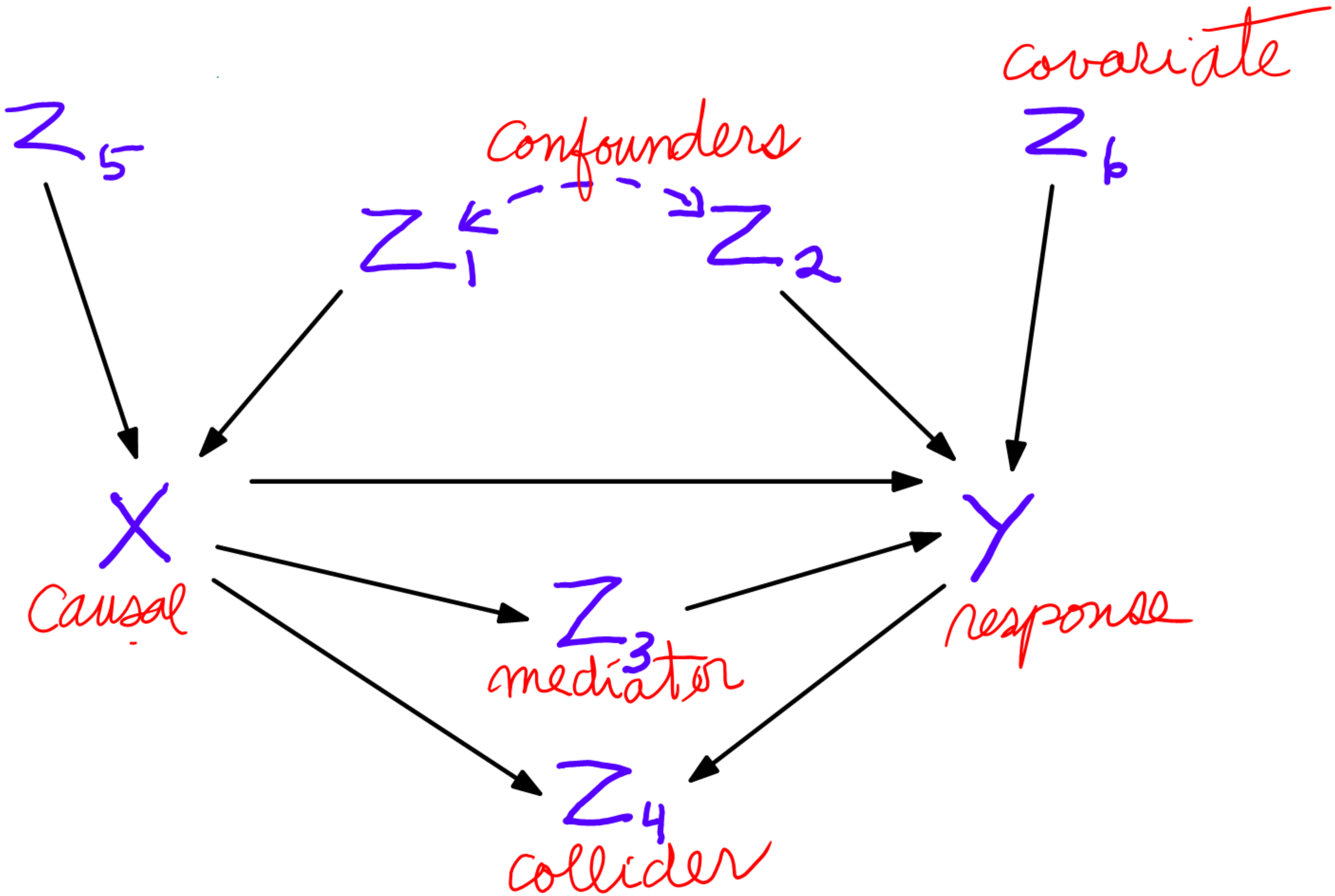


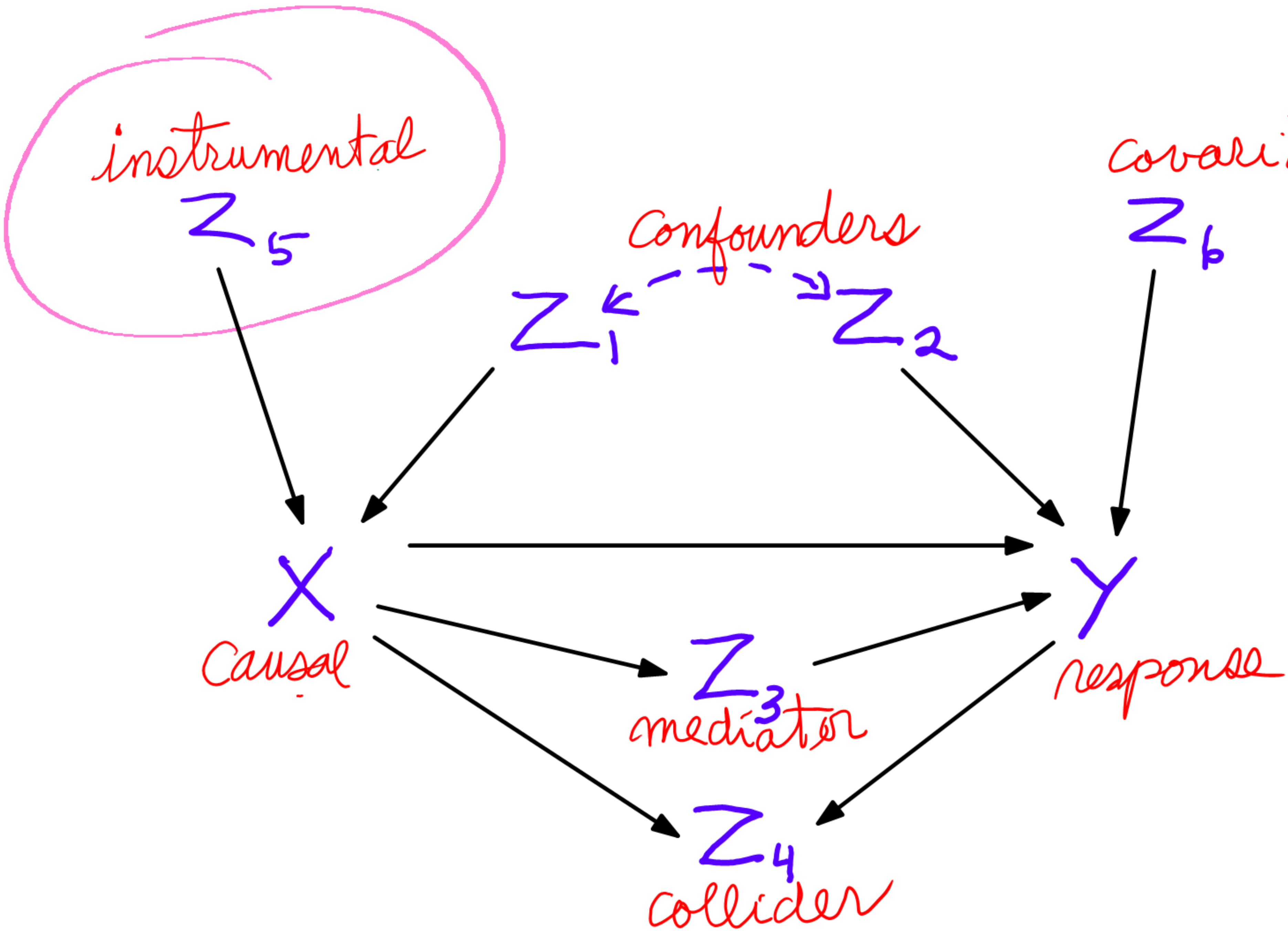
Confounders
Z_5
Z_1
Z_2
Z_6
X
Causal
Y
response
Z_3
Z_4

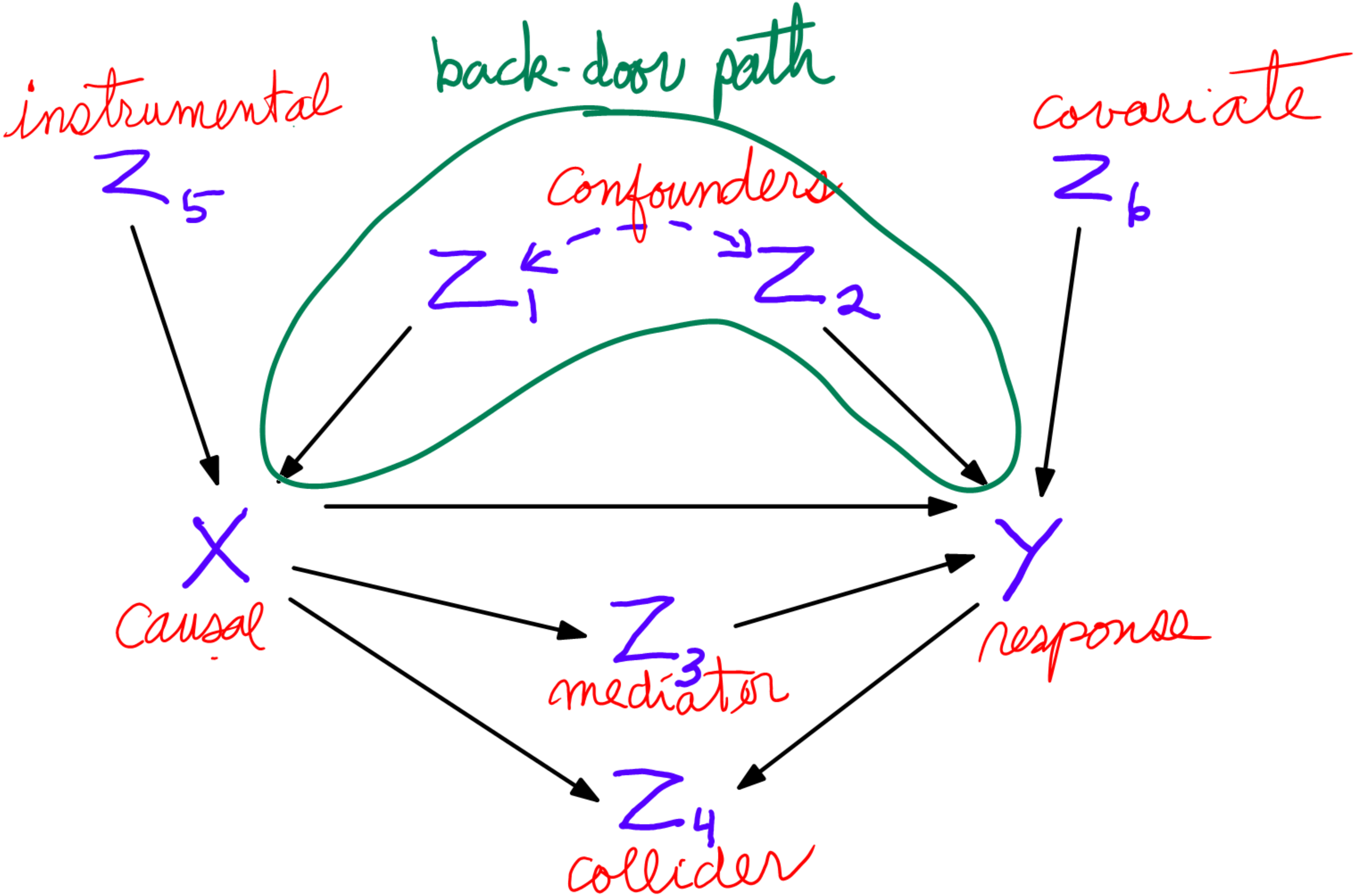


covariate
Confounders
Z_5
Z_1
Z_2
Z_6
X
Y
Causal
response
Z_3
Z_4

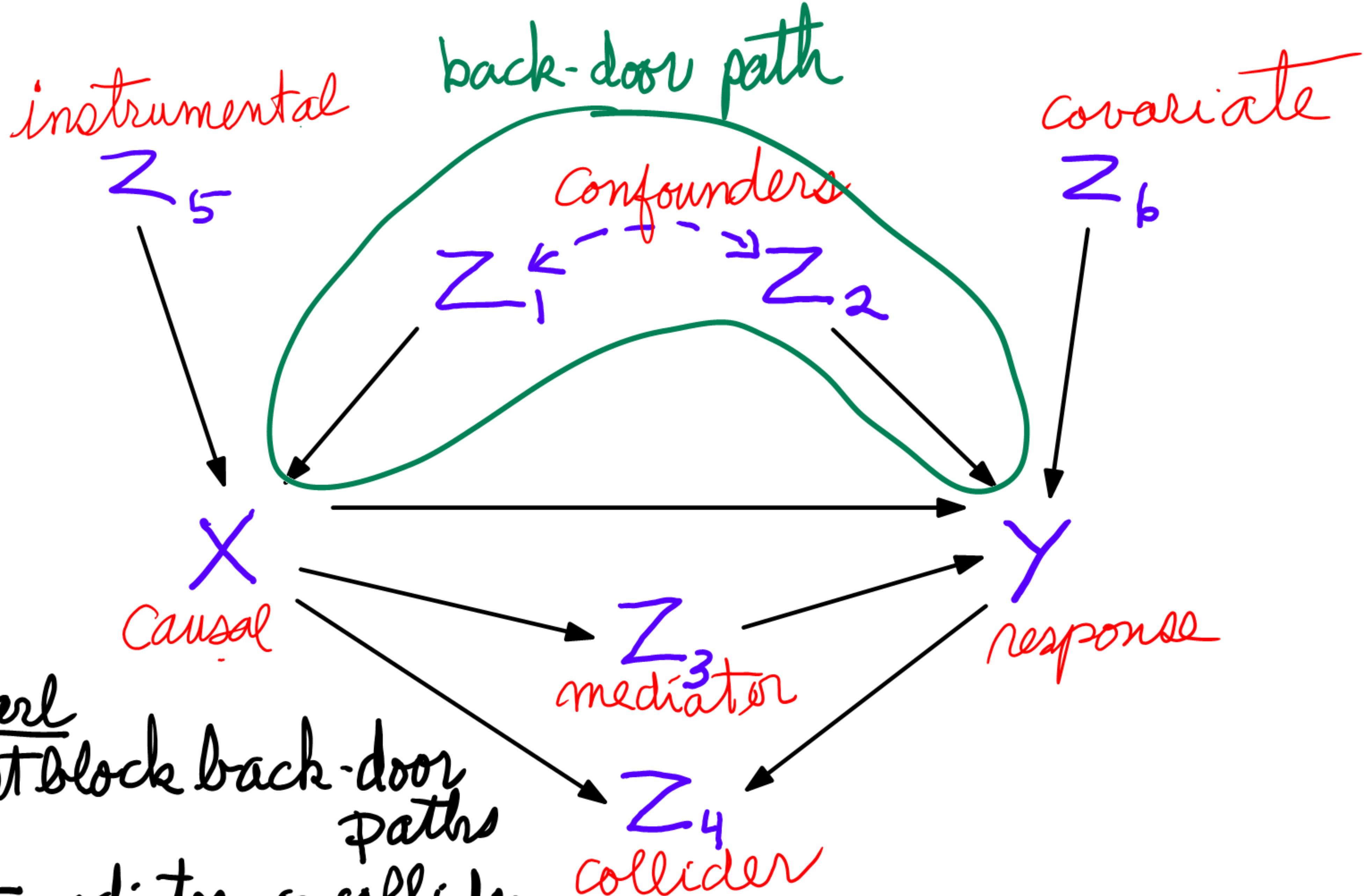

covariate
confounders
Z_5
Z_1
Z_2
Z_6
X
causal
Y
response
Z_3
mediator
Z_4


covariate
Z_5
confounders
Z_1
Z_2
Z_6
X
causal
Y
response
Z_3
mediator
Z_4
collider


instrumental
Z_5
confounders
Z_1
Z_2
covariate
Z_6
X
causal
Y
response
Z_3
mediator
Z_4
collider

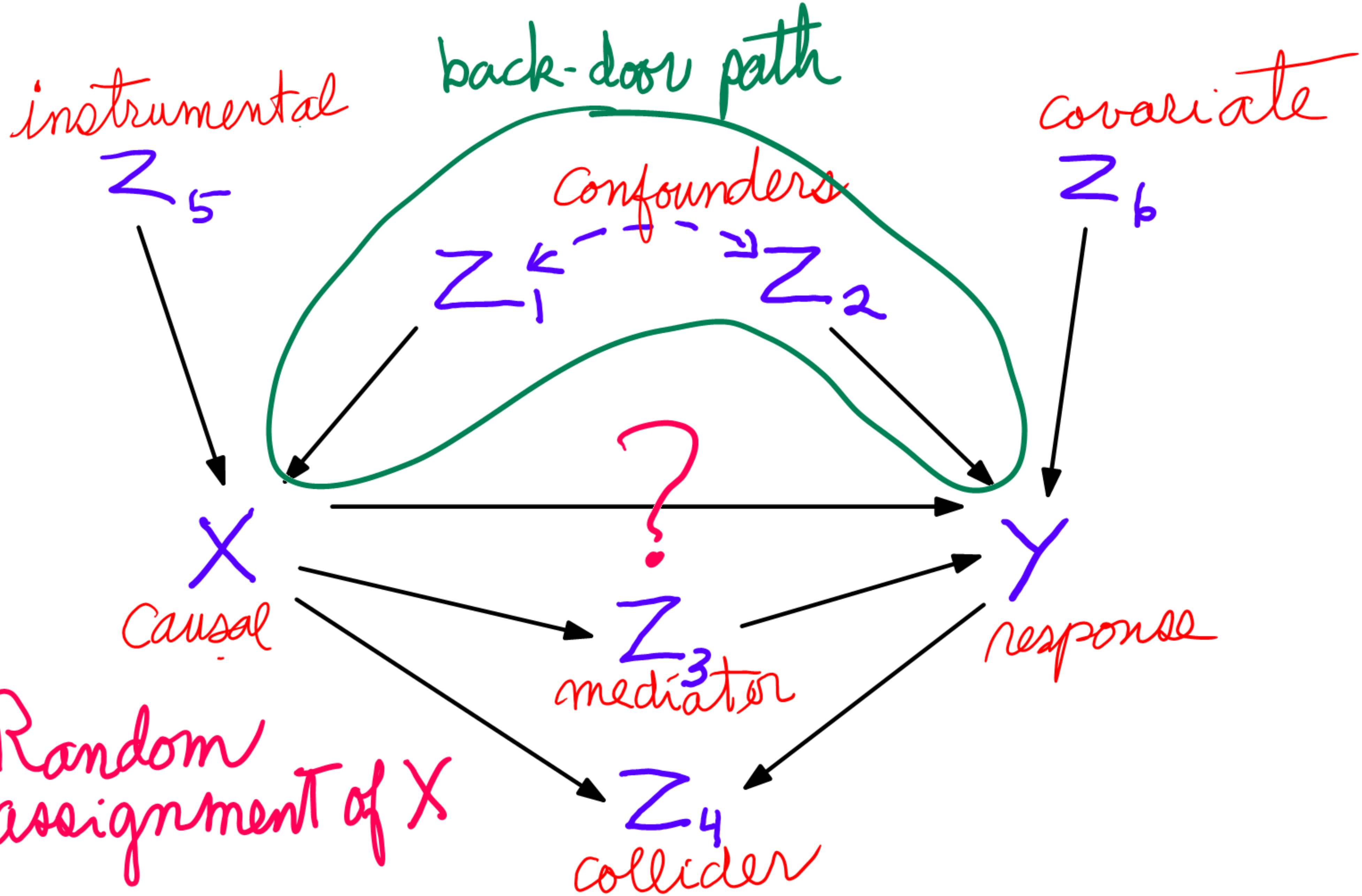


back-door path
instrumental
Z_5
confounders
Z_1
Z_2
covariate
Z_6
X
causal
Y
response
Z_3
mediator
Z_4
collider

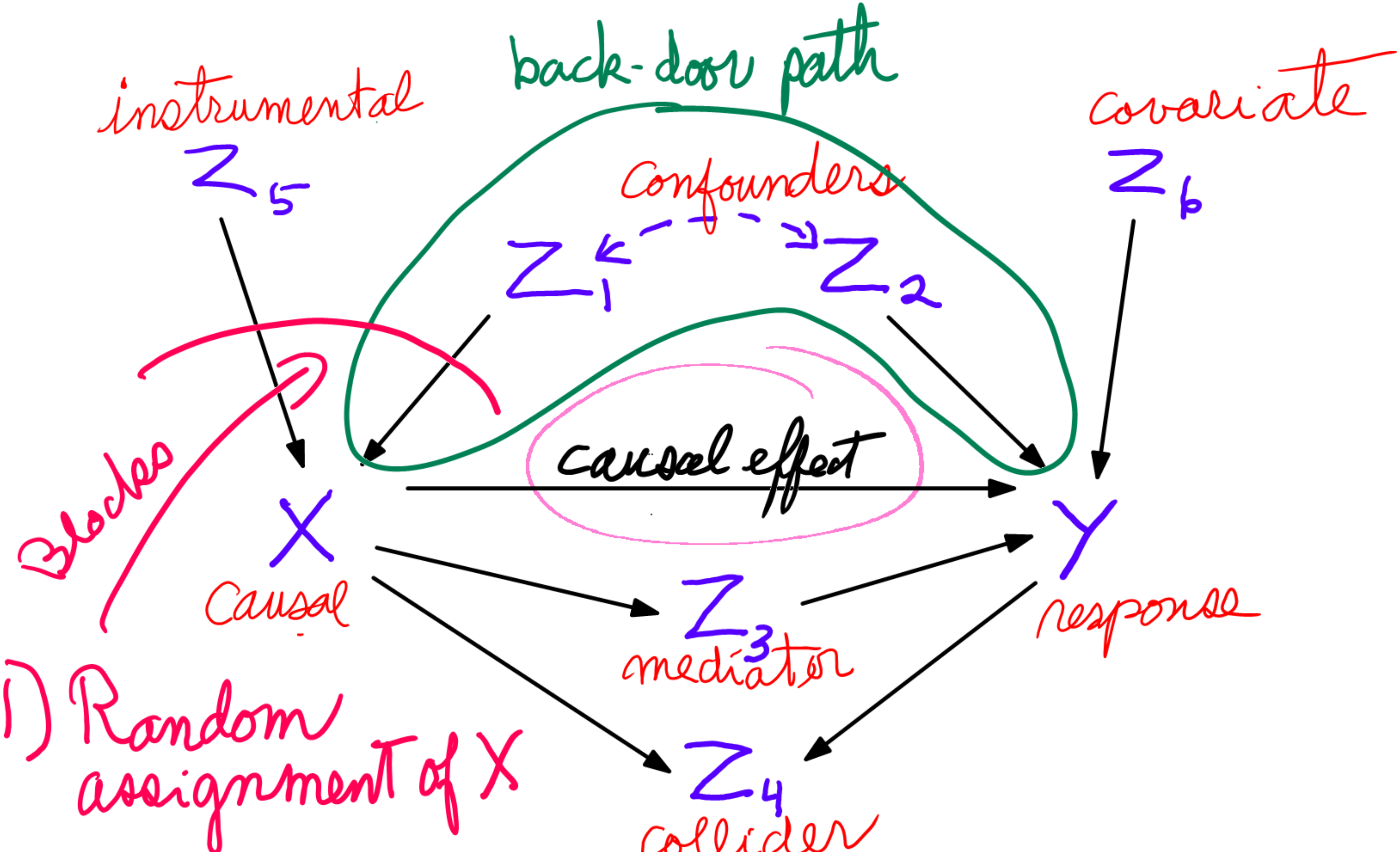


Pearl

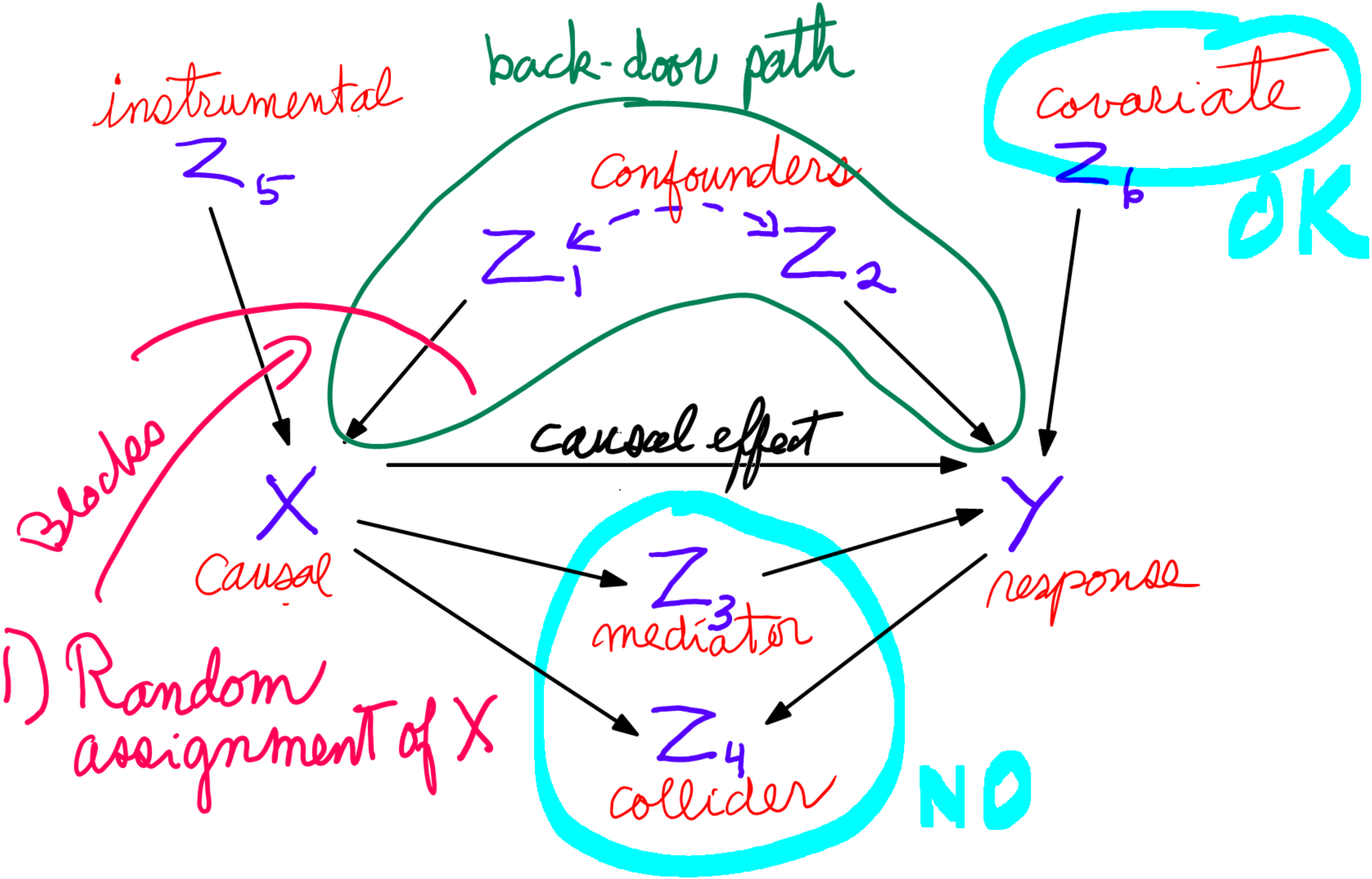
- Must block back-door paths
- NOT mediator or collider

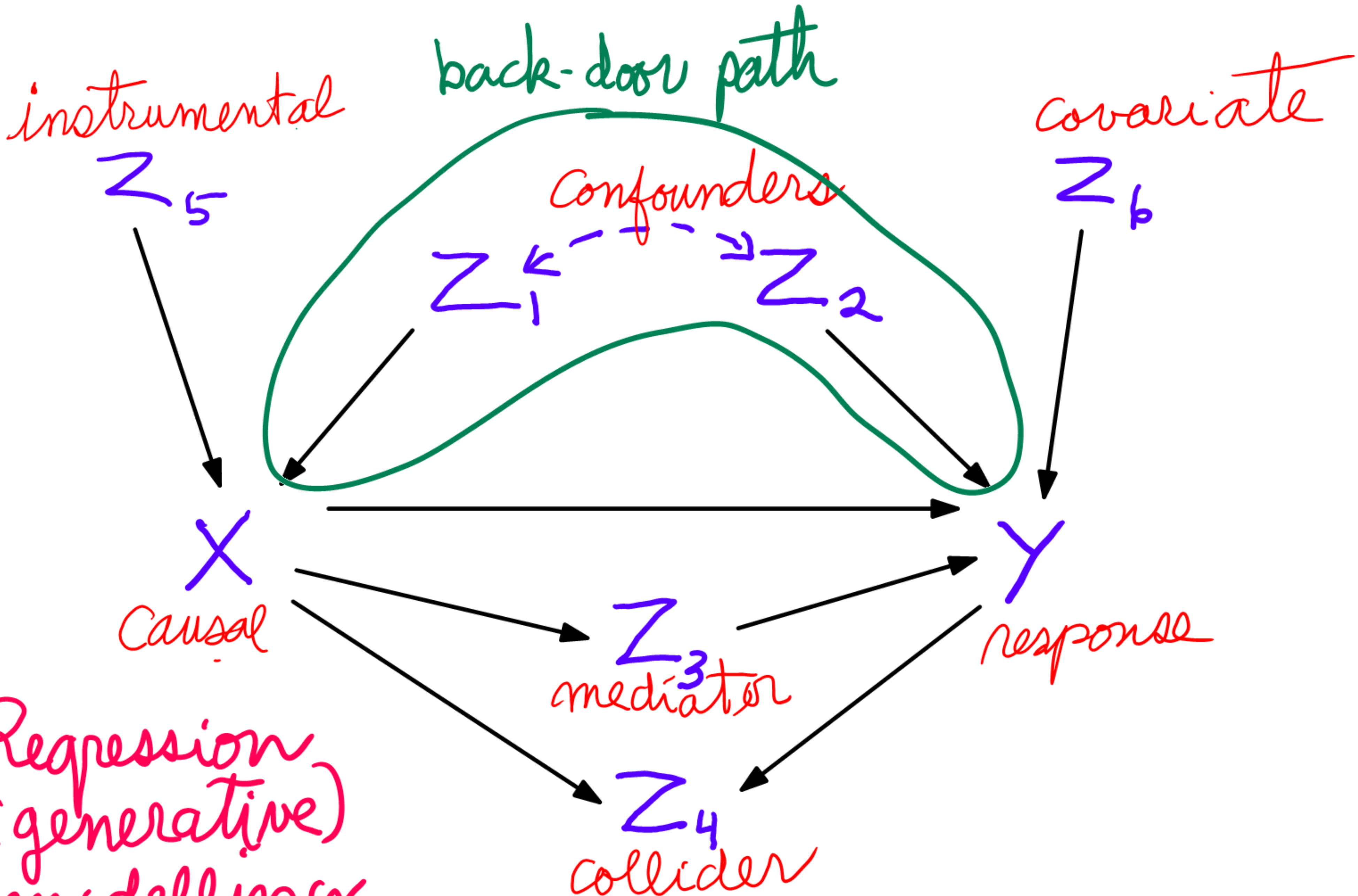


1) Random assignment of X

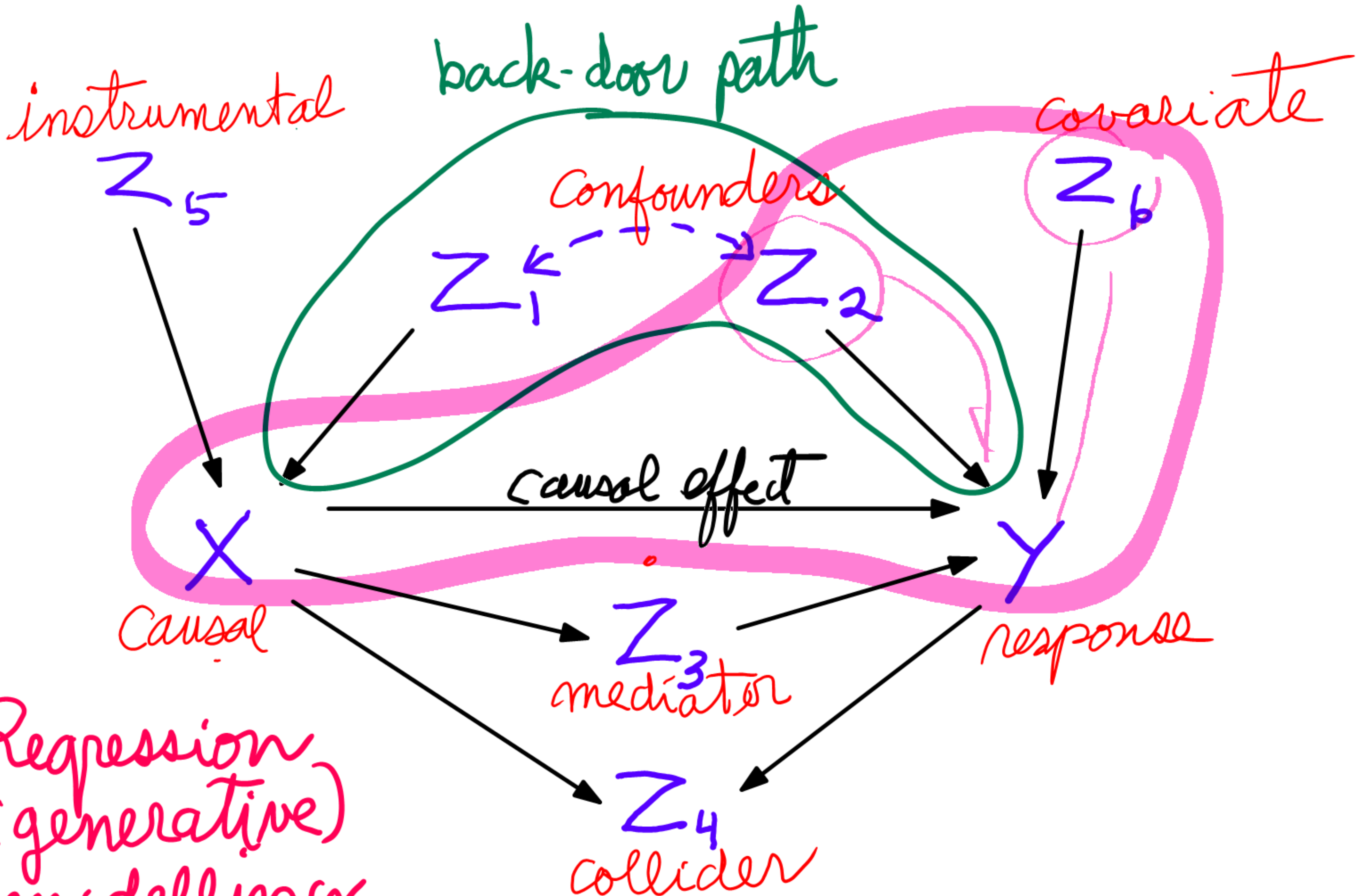


back-door path
instrumental
covariate
Confounders
Z_5
Z_6
Z_1
Z_2
causal effect
Blocks
X
causal
Y
response
Z_3
mediator
Z_4
collider
1) Random assignment of X

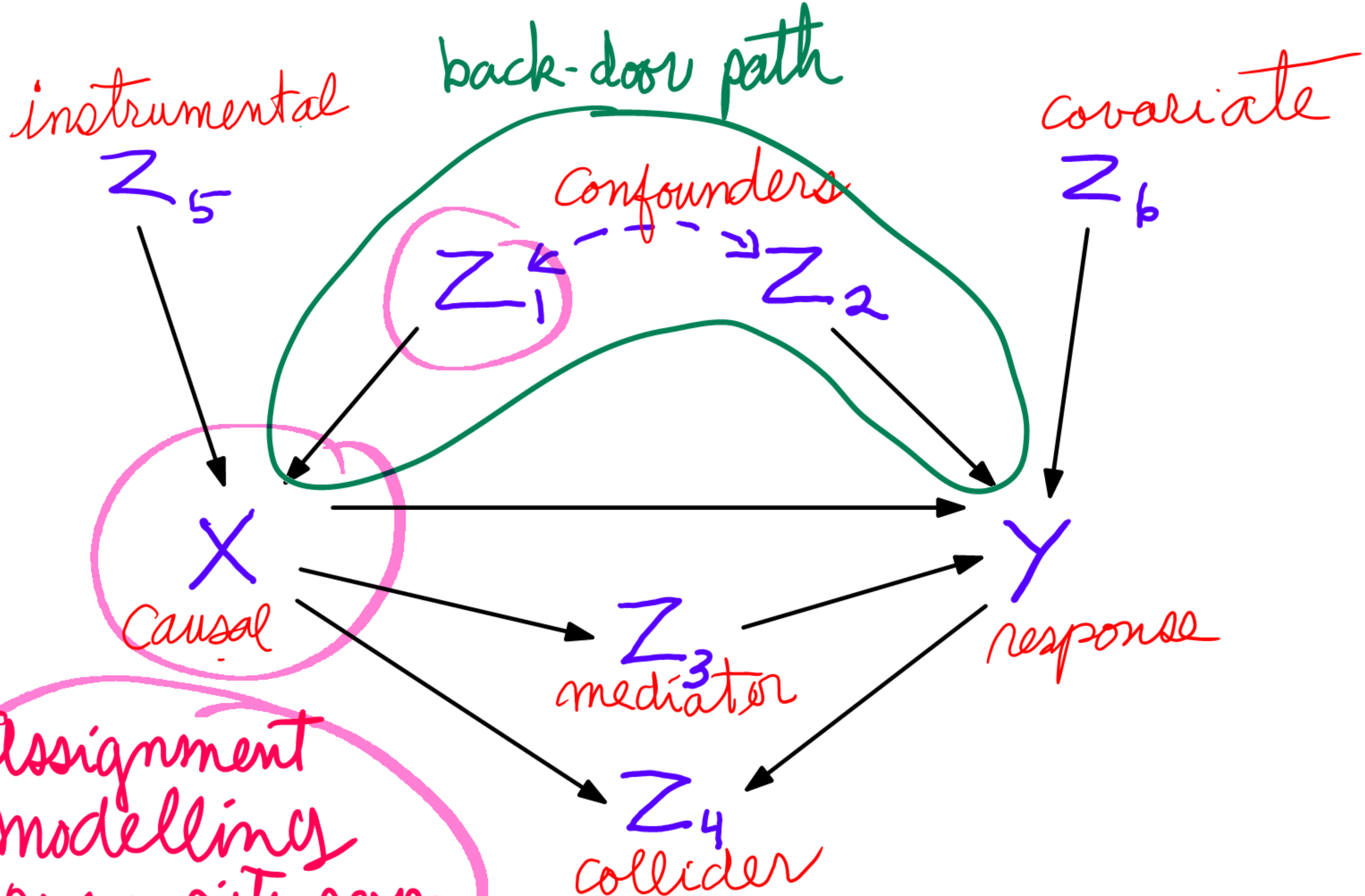

back-door path
instrumental
Z_5
covariate
Z_6
OK
Confounders
Z_1
Z_2
Blocks
causal effect
X
Causal
Y
response
Z_3
mediator
Z_4
collider
NO
1) Random assignment of X



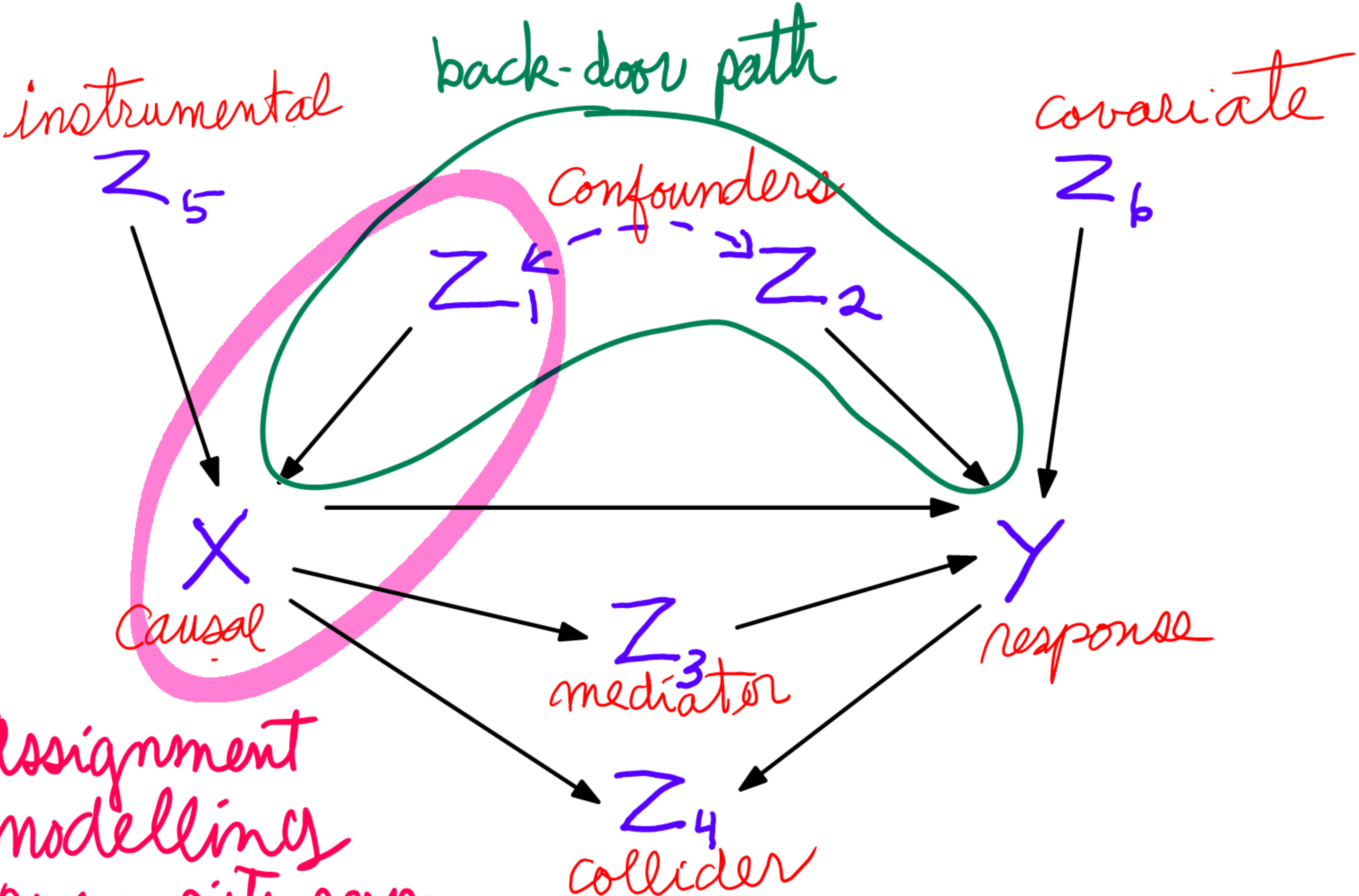
2) Regression (generative) modelling



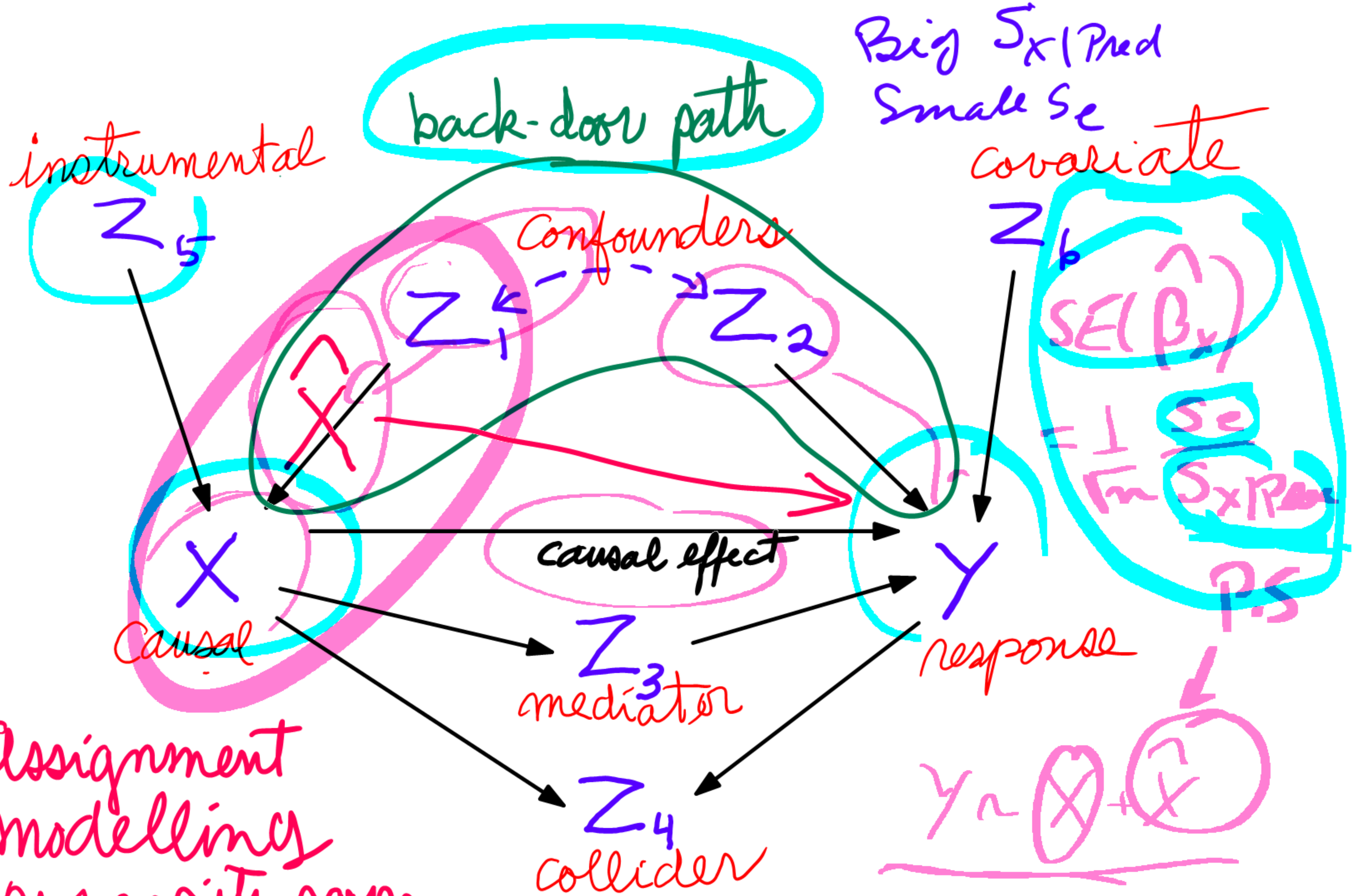
2) Regression (generative modelling)



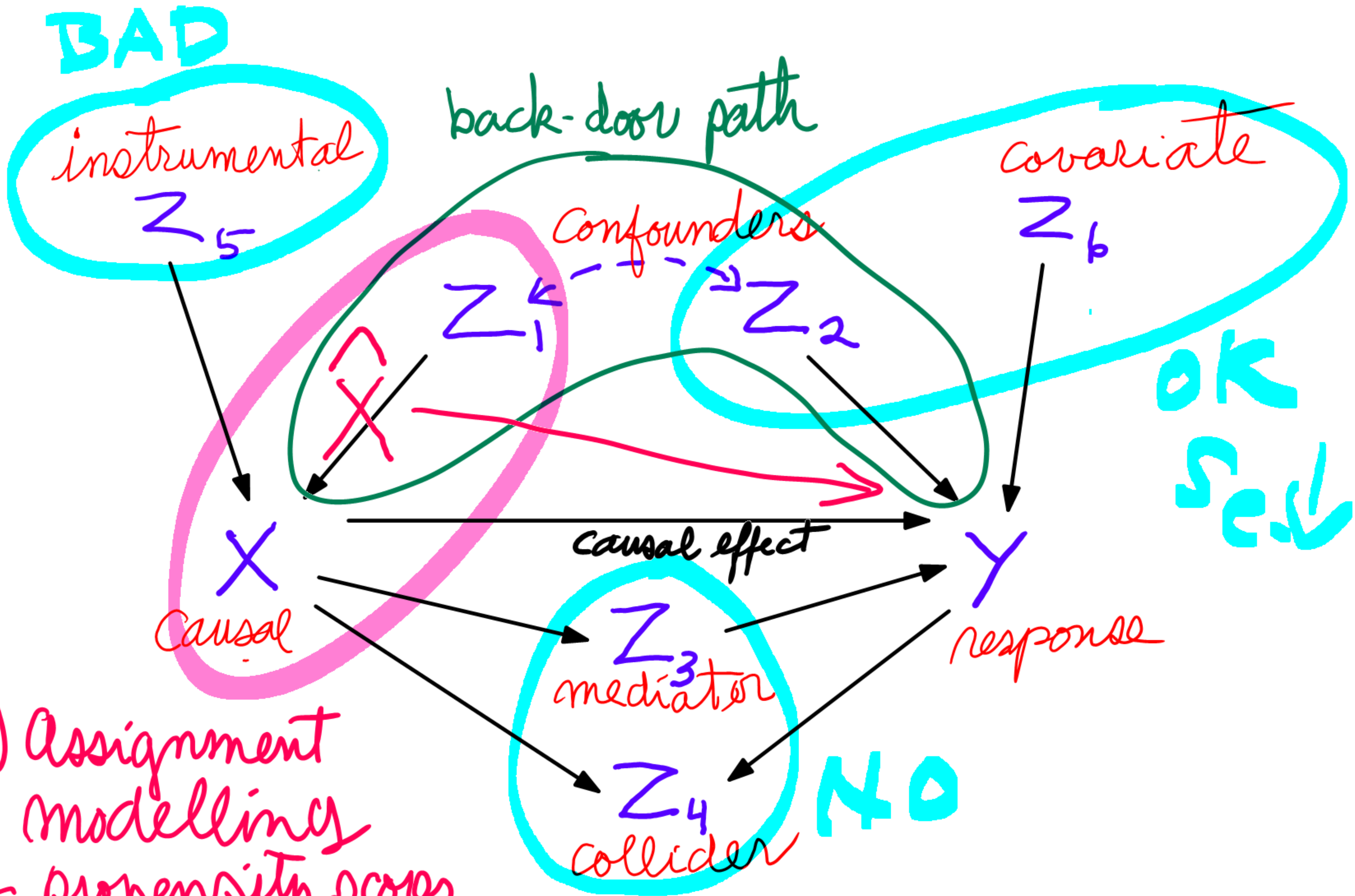
3) Assignment modelling
– propensity scores



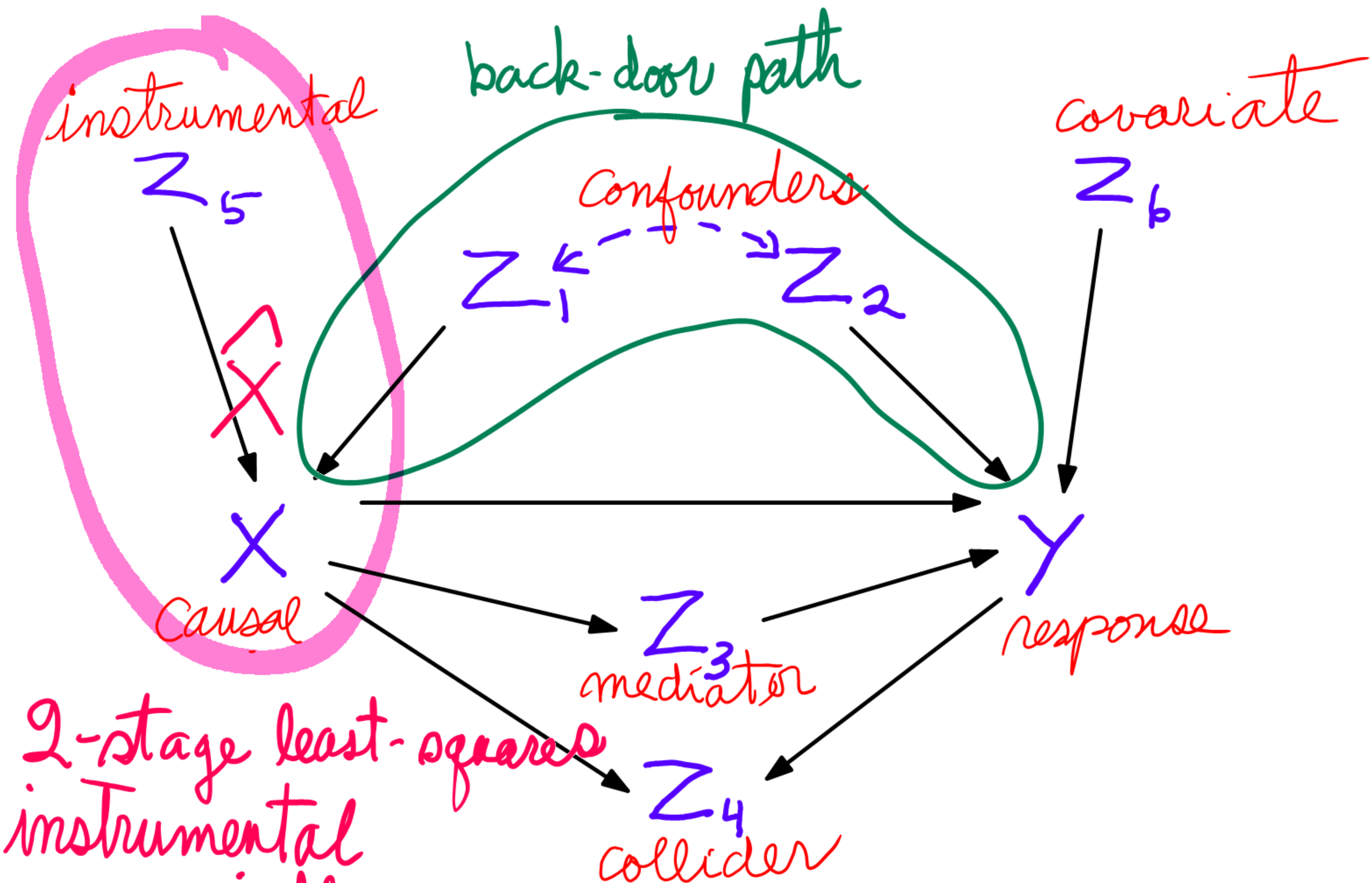
3) Assignment modelling
- propensity scores



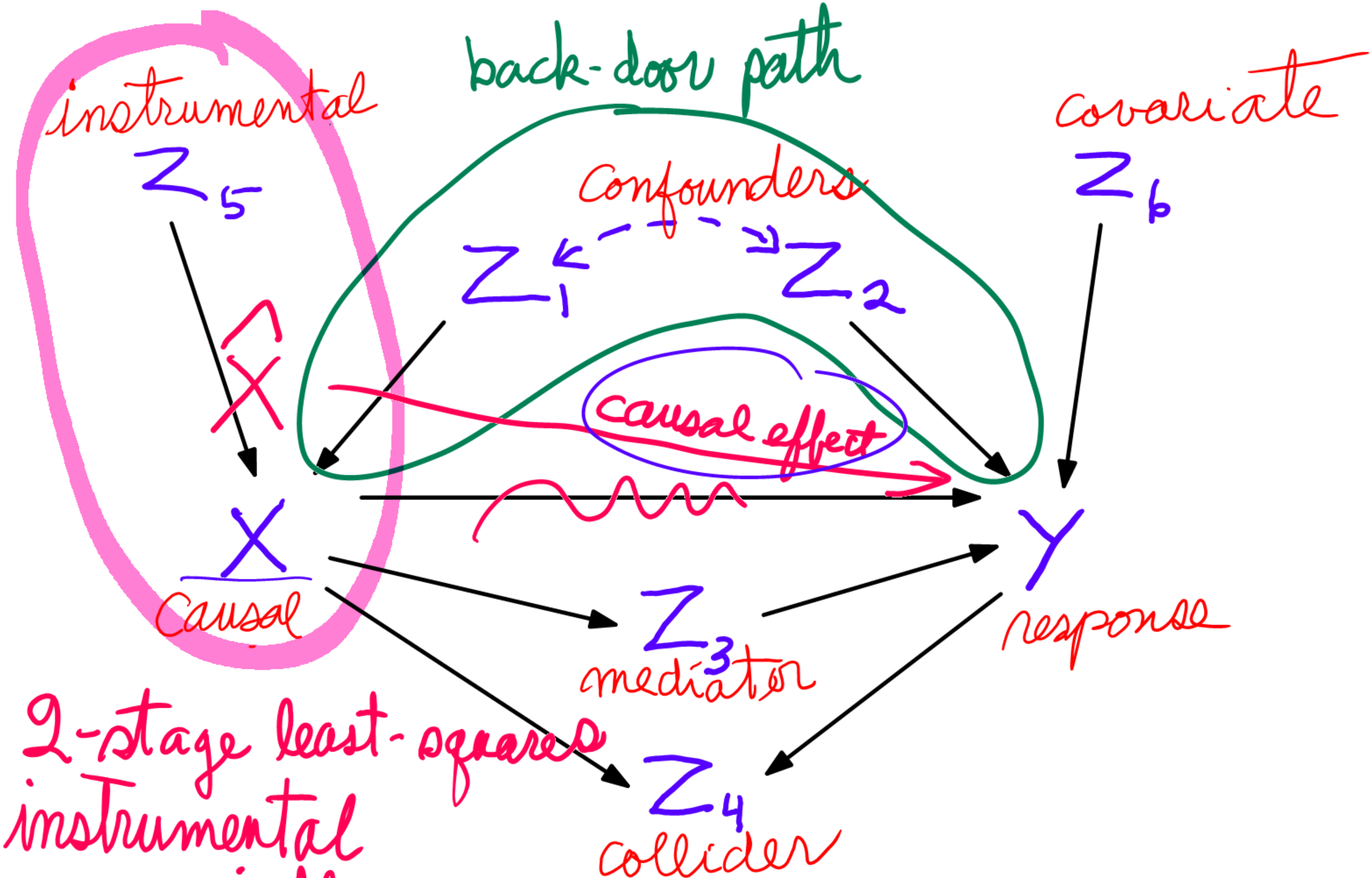
3) Assignment modelling
– propensity scores



3) Assignment modelling
- propensity scores

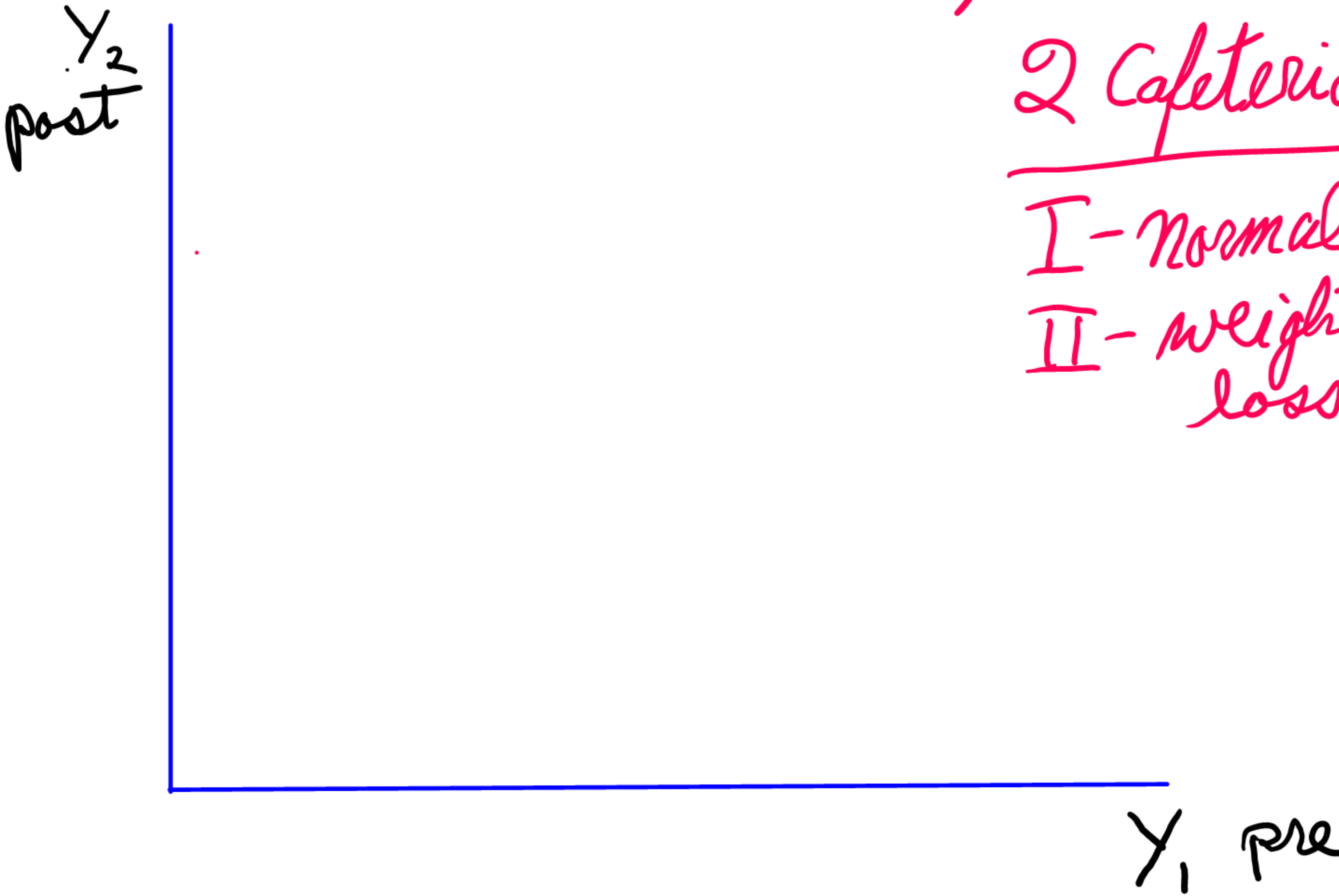


4) 2-stage least-squares
instrumental
variables



4) 2-stage least-squares
instrumental
variables

Lord's Paradox (Wainer version)



2 Cafeterias

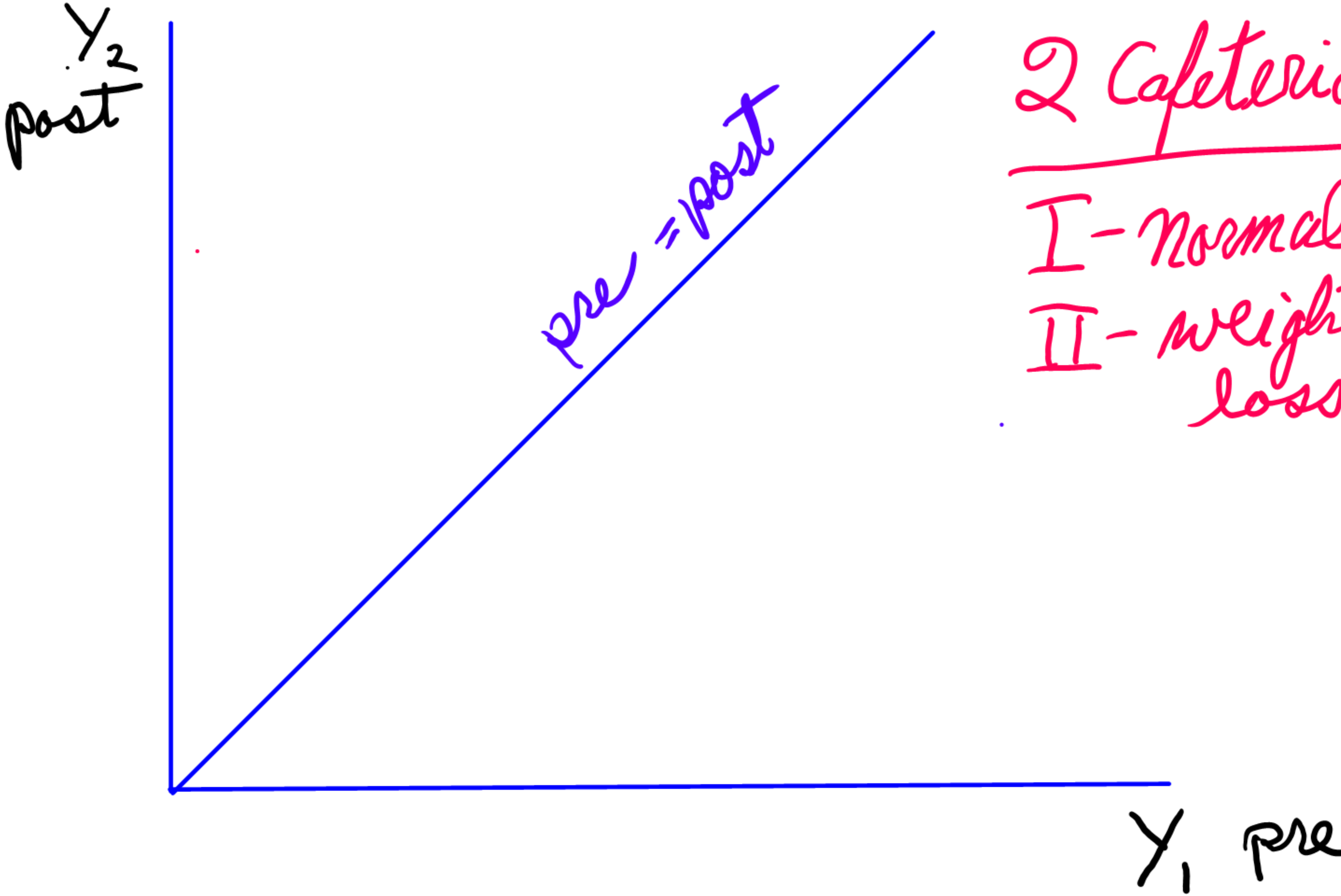
I - normal

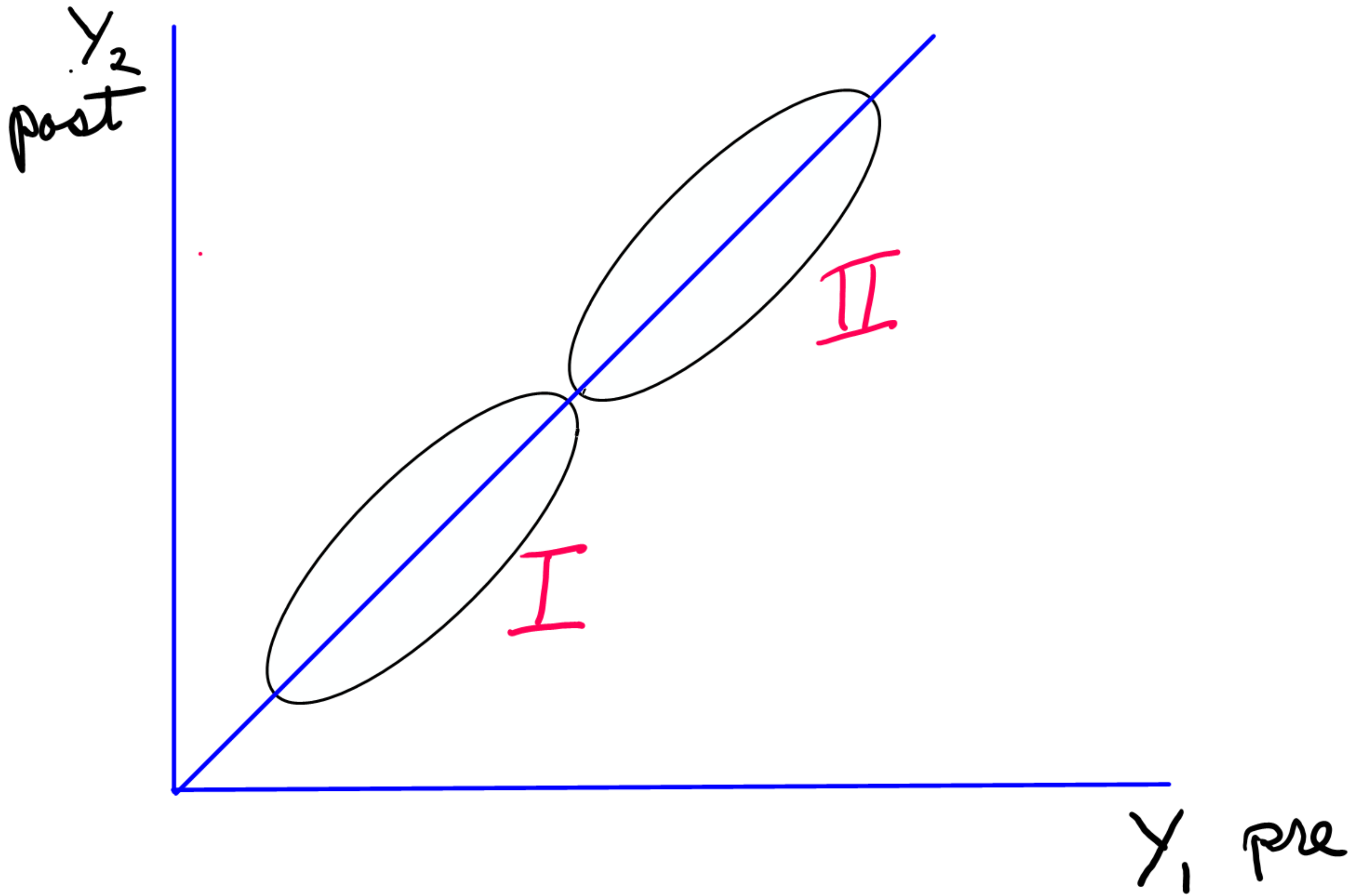
II - weight loss

2 Cafeterias

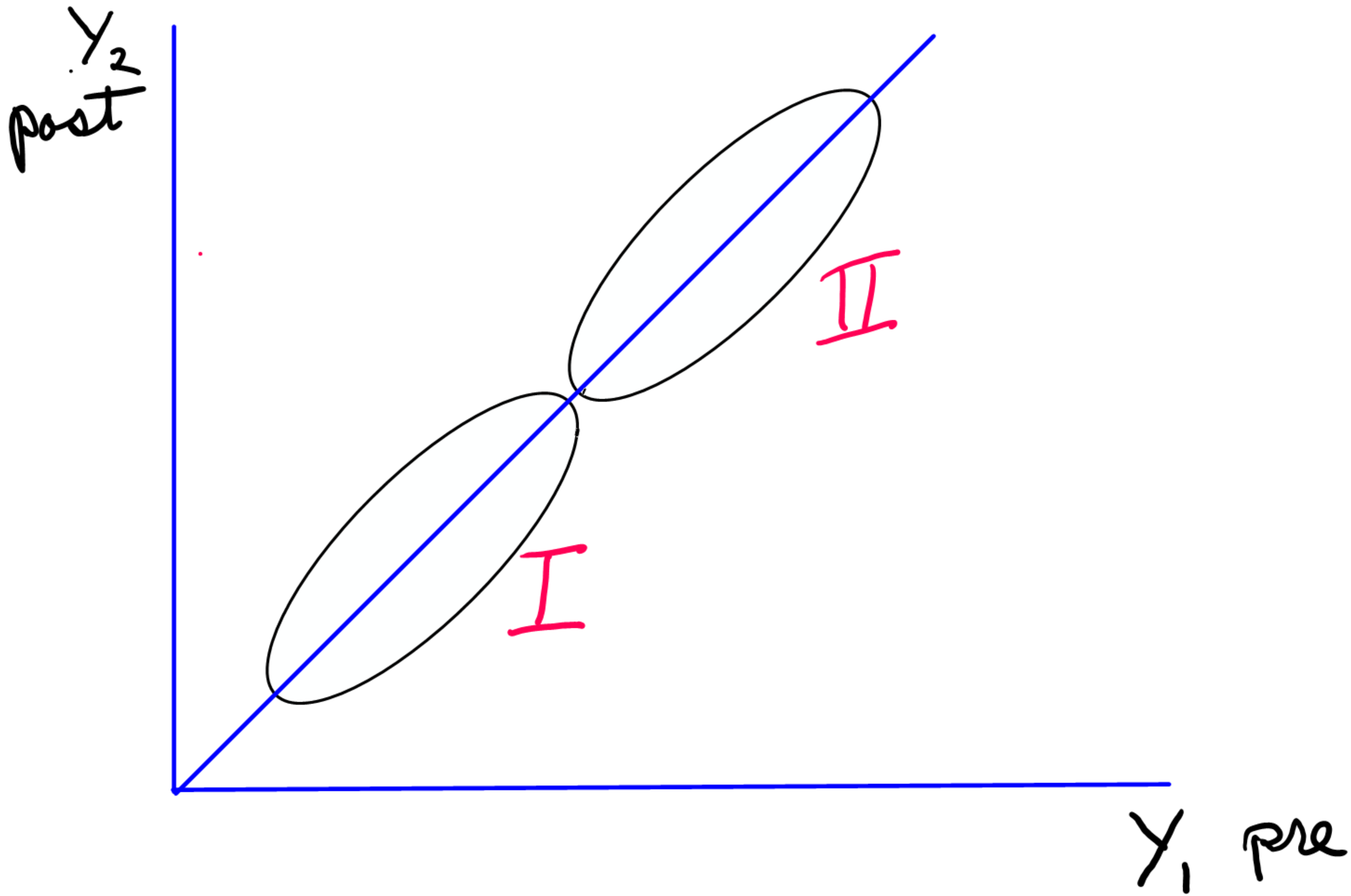
I - normal

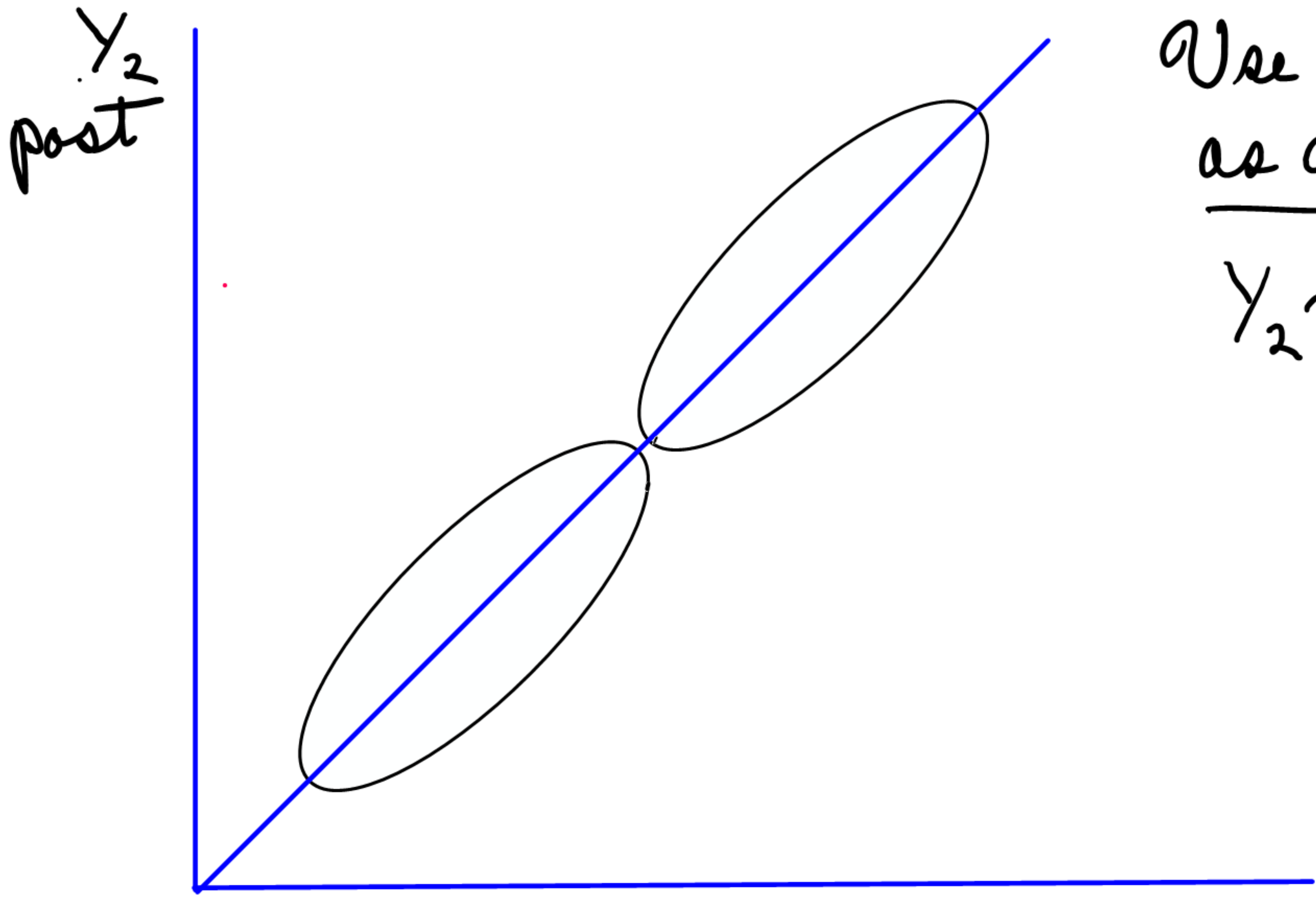
II - weight loss





Y_2
post
I
II
Y_1 pre


Y_2
post
I
II
Y_1 pre

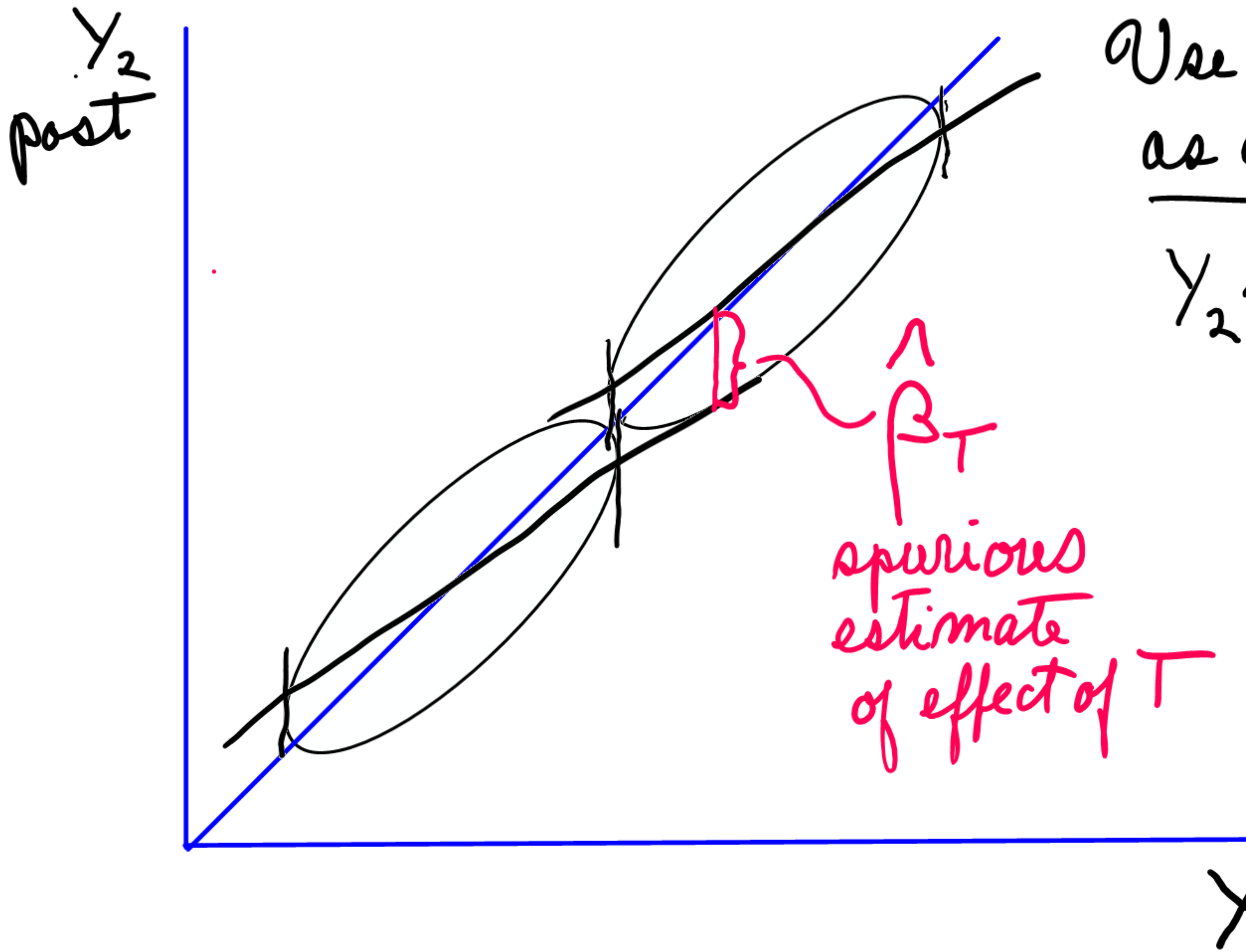


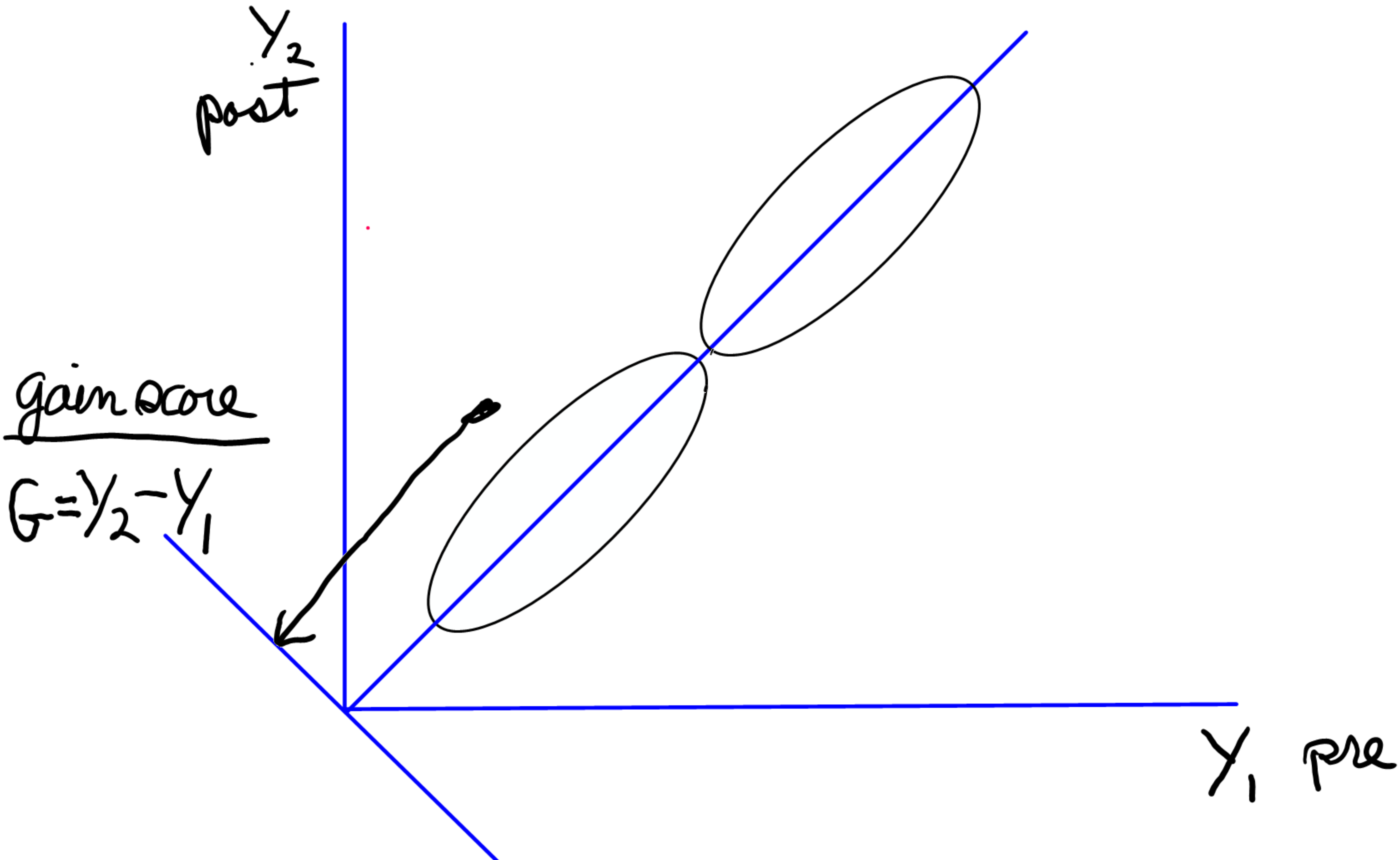
Use pretest as covariate

$Y_2 \sim T + Y_1$

Use pretest as covariate

$Y_2 \sim T + Y_1$

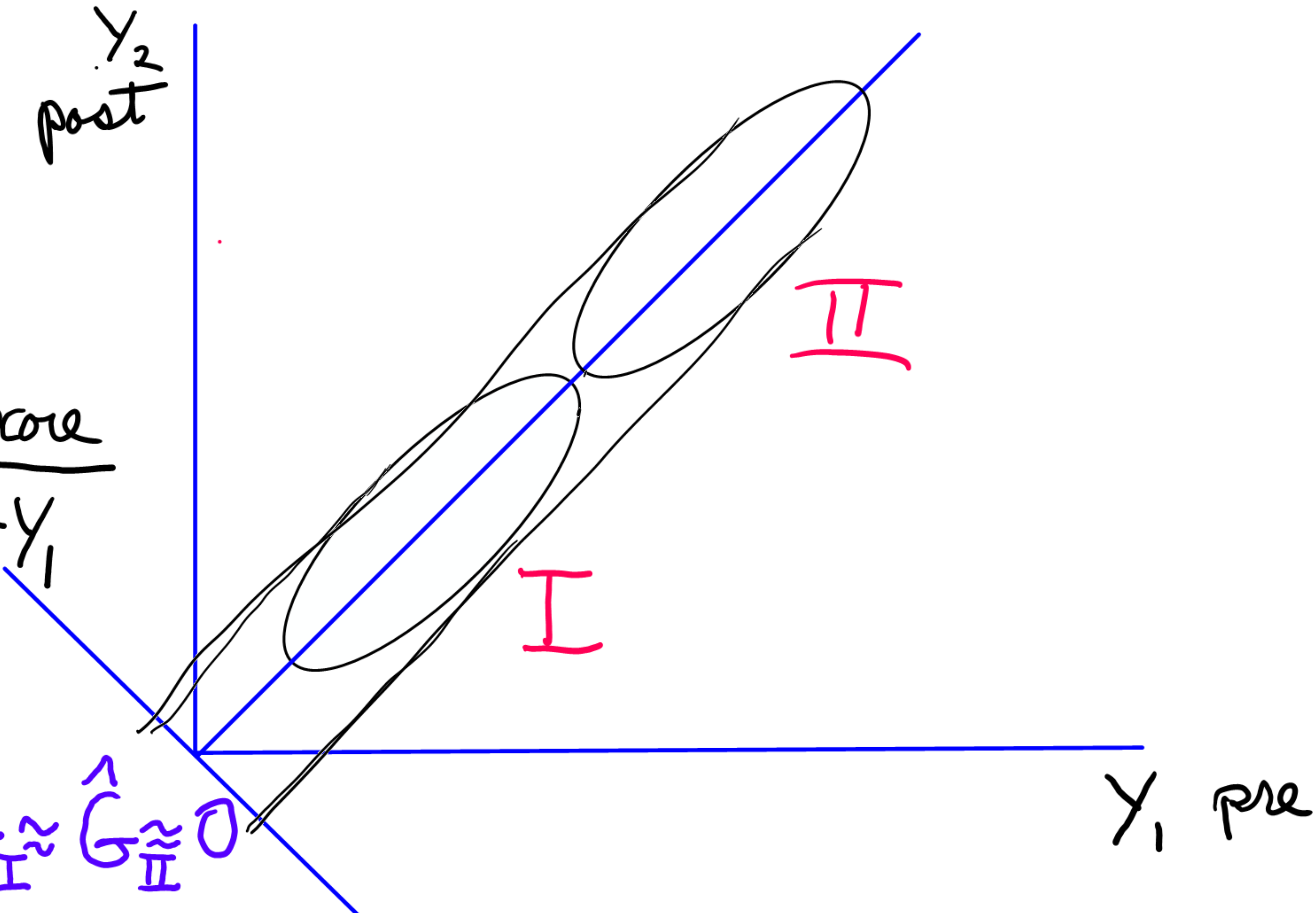




Y_2
post
gain score
$G = Y_2 - Y_1$
Y_1 pre

gain score

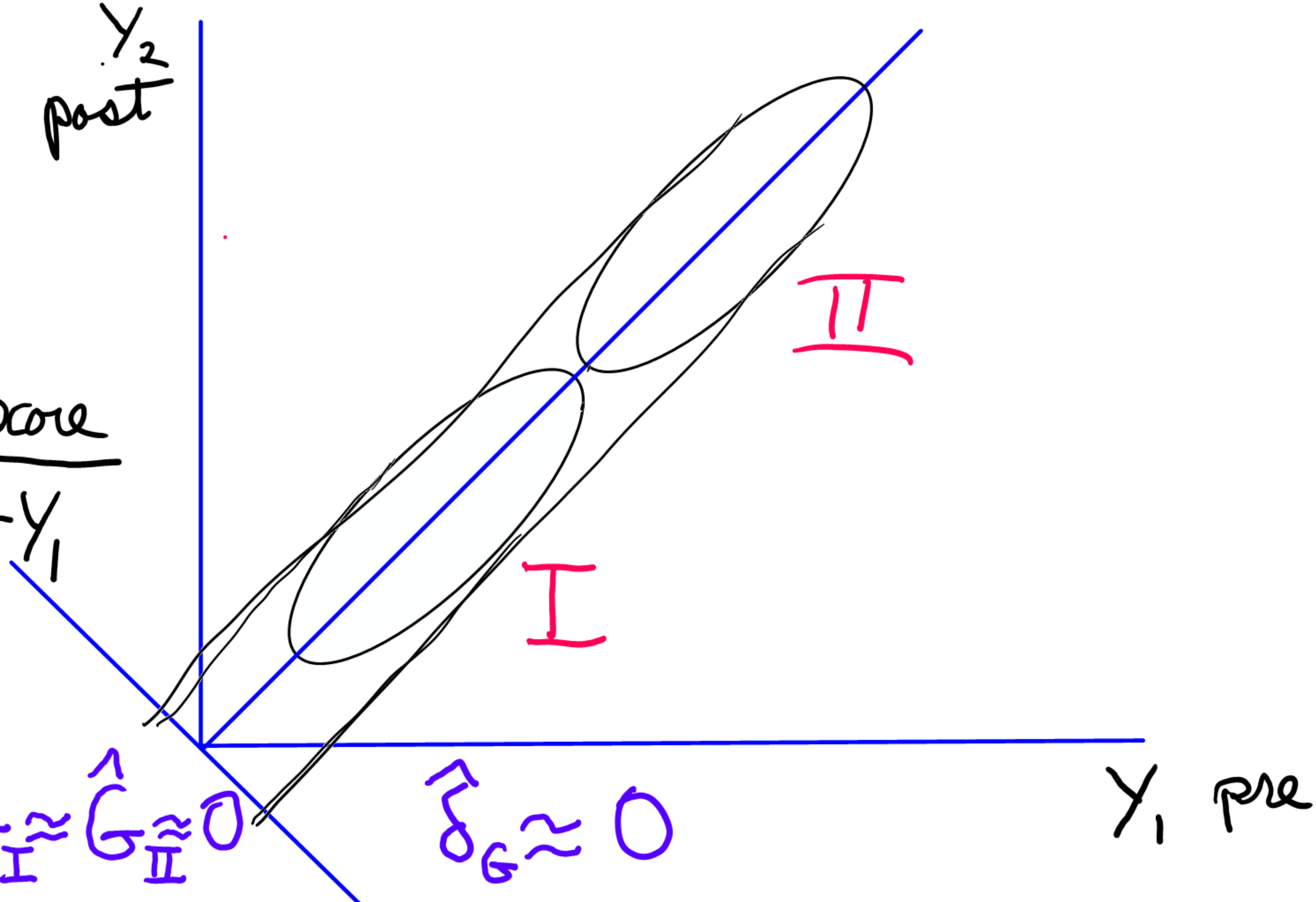
$G = Y_2 - Y_1$



$\hat{G}_{I} \approx \hat{G}_{II} \approx 0$

gain score

$G = Y_2 - Y_1$

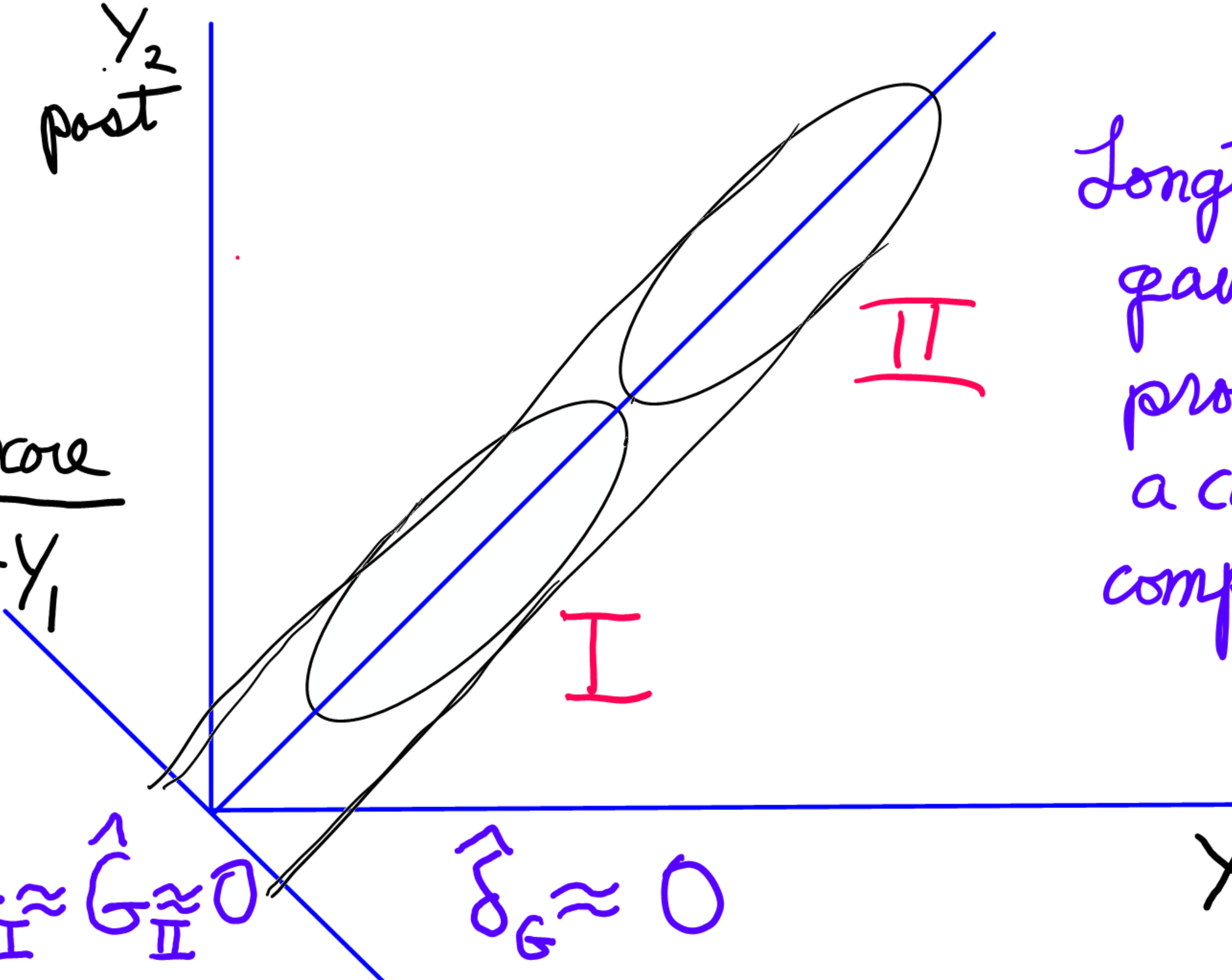


$\hat{G}_{I} \approx \hat{G}_{II} \approx 0$

$\hat{\delta}_G \approx 0$

gain score

$G = Y_2 - Y_1$



Y_1 pre

Longitudinal gain score provides a correct comparison

$\hat{G}_{I} \approx \hat{G}_{II} \approx 0$

$\hat{\delta}_G \approx 0$

Conditions

- Same scale for Y_{pre} & Y_{post}
- No time-varying confounders

Within-subject effect adjusts for between-subject confounders whether measured or not.

Good model?

$Y \sim X + Z_i + Z_j$

Want:

1) Unbiased : consistent

Block backdoors – NOT mediators & colliders

2) Low SE $= SD(Y_{res}) / SD(X_{res})$

Small $SD(Y_{res})$, Large $SD(X_{res})$

3) Honest SE

4) Robust Propensity scores – focus on X

Use the AVP to compare models.

Using confounders close to Y

$$\left.\begin{array}{l}\downarrow SD(Y_{res}) \\ \uparrow SD(X_{res})\end{array}\right\} \downarrow SE(\hat{\beta}_T)$$

But may not have knowledge about structure of model.

Using confounders close to X

$$\left.\begin{array}{l}\uparrow SD(Y_{res}) \\ \downarrow SD(X_{res})\end{array}\right\} \uparrow SE(\hat{\beta}_T)$$

But may have external knowledge about model.