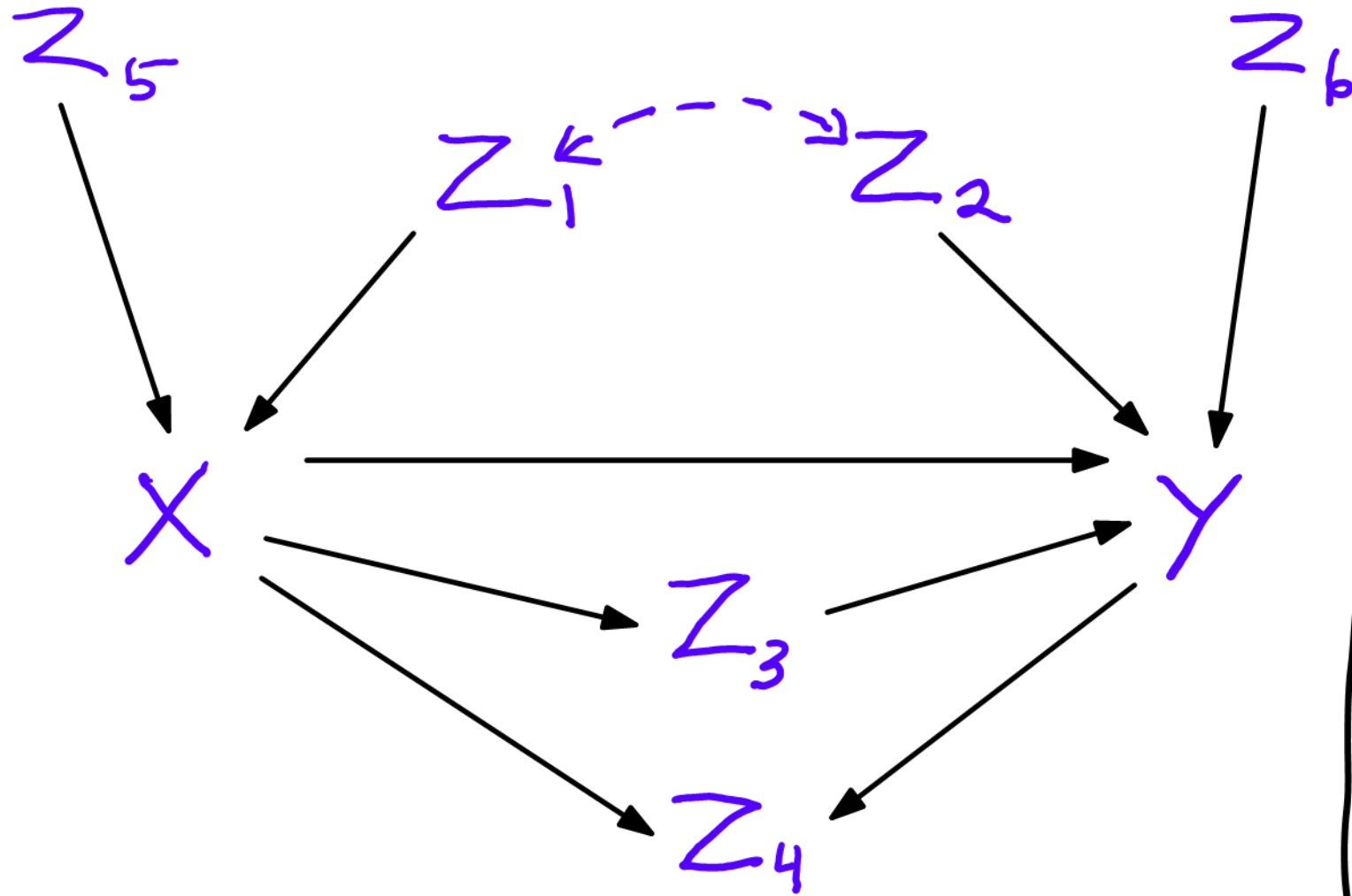


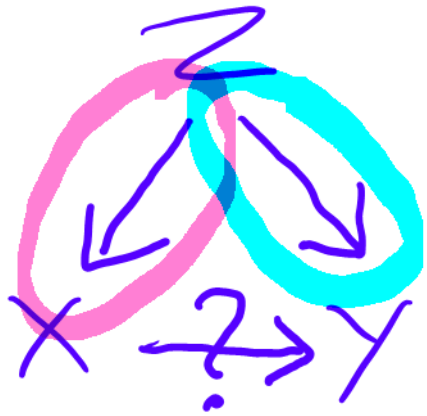
# Does X cause Y?

- A bird's eye view of methods with observational data
- Lord's Paradox and the role of longitudinal data

# Causal Graph Pearl & Mackenzie (2019)

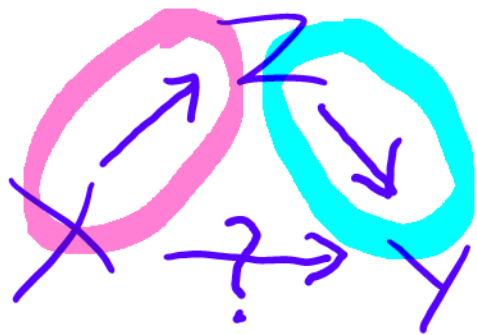


DAG =  
Directed  
Acyclic  
graph



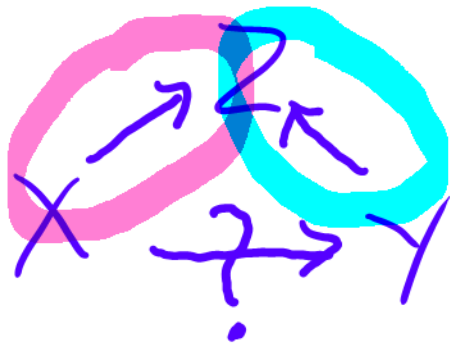
confounder

must include to see  
see causal effect  
of X on Y



mediator

must exclude  
including may  
wipe out a true effect



collider

e.g.  
selection

must exclude  
including may  
create the impression  
of an effect although there is  
none

# Moderators?

- Can have - Confounder - moderators
- mediator - moderators
- collider - moderators

Also mediator-colliders, etc

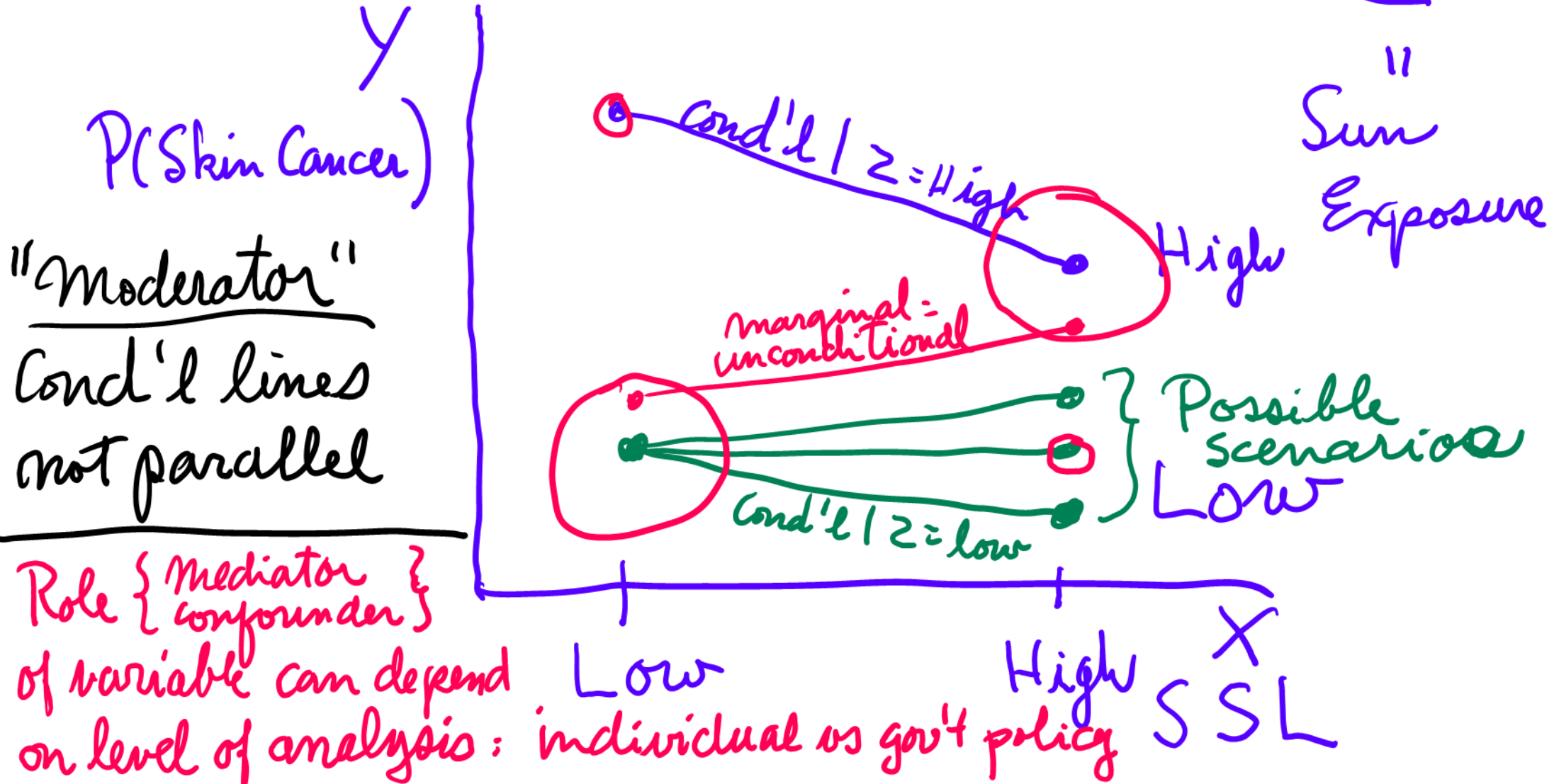
BUT not represented by DAGs

— So convenient visualizations of DAG are useful abstraction but limited in practice

— Avoiding inclusion of mediators more critical than avoiding colliders since inclusion of other confounders can correct for inclusion of a collider.

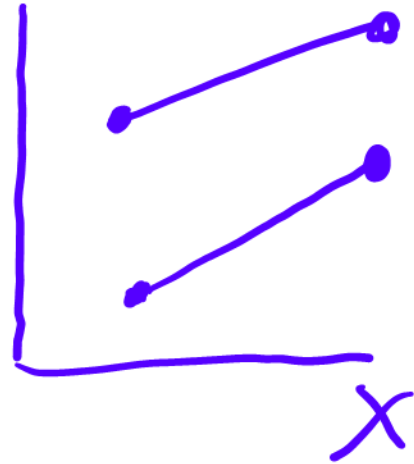
# Moderator (= interaction) in data space

SSL example

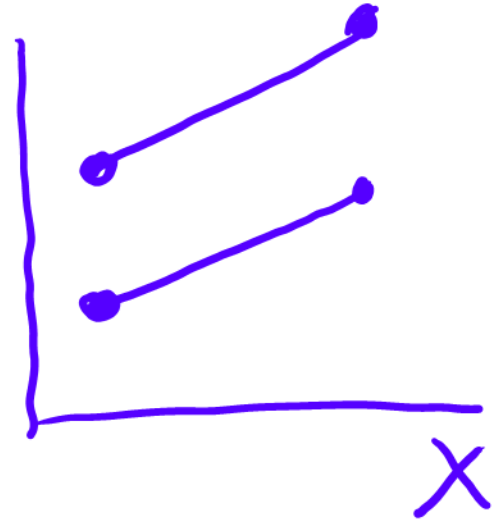


Moderation may be removable by <sup>monotonically</sup> transforming  $Y$   
 if 1) Cond'l effects in same direction  
 2) No crossings of lines

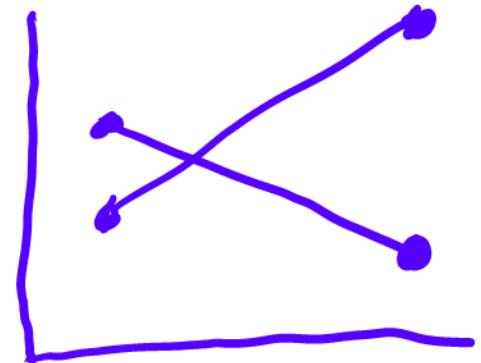
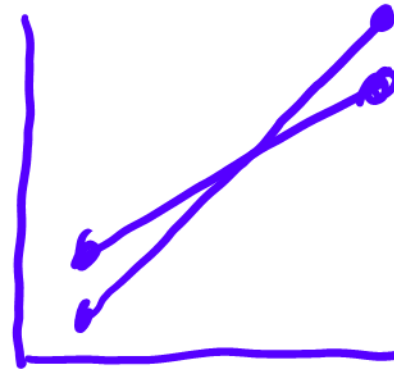
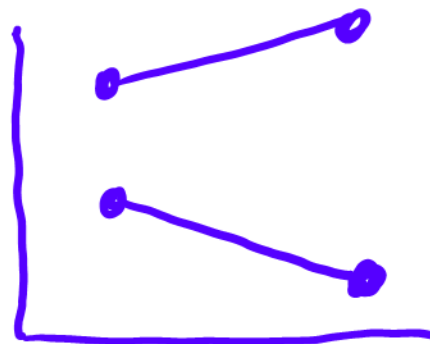
Removable  $P(Y)$



log odds  
 " "  
 $\log\left(\frac{P(Y)}{1-P(Y)}\right)$

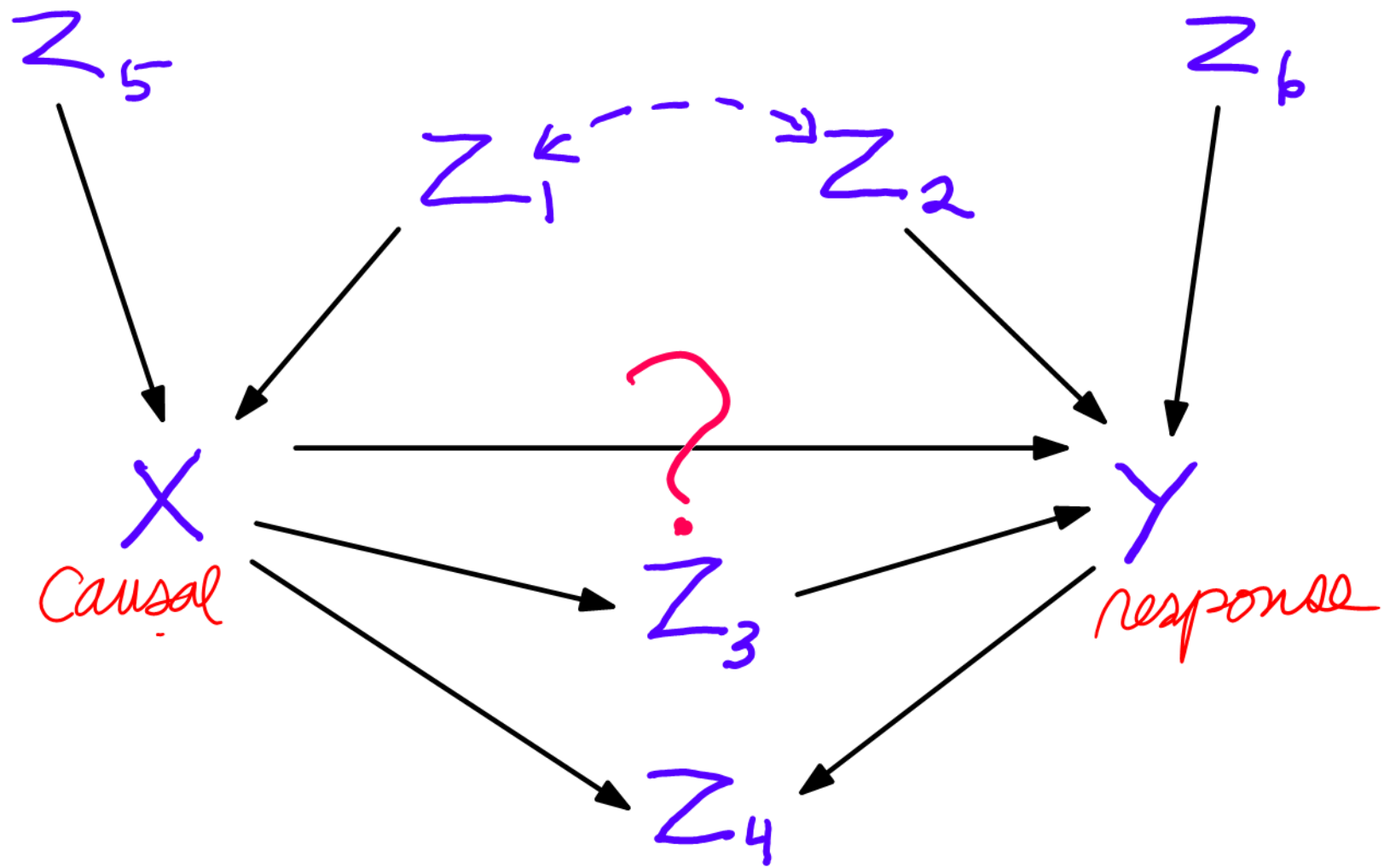


Nonremovable

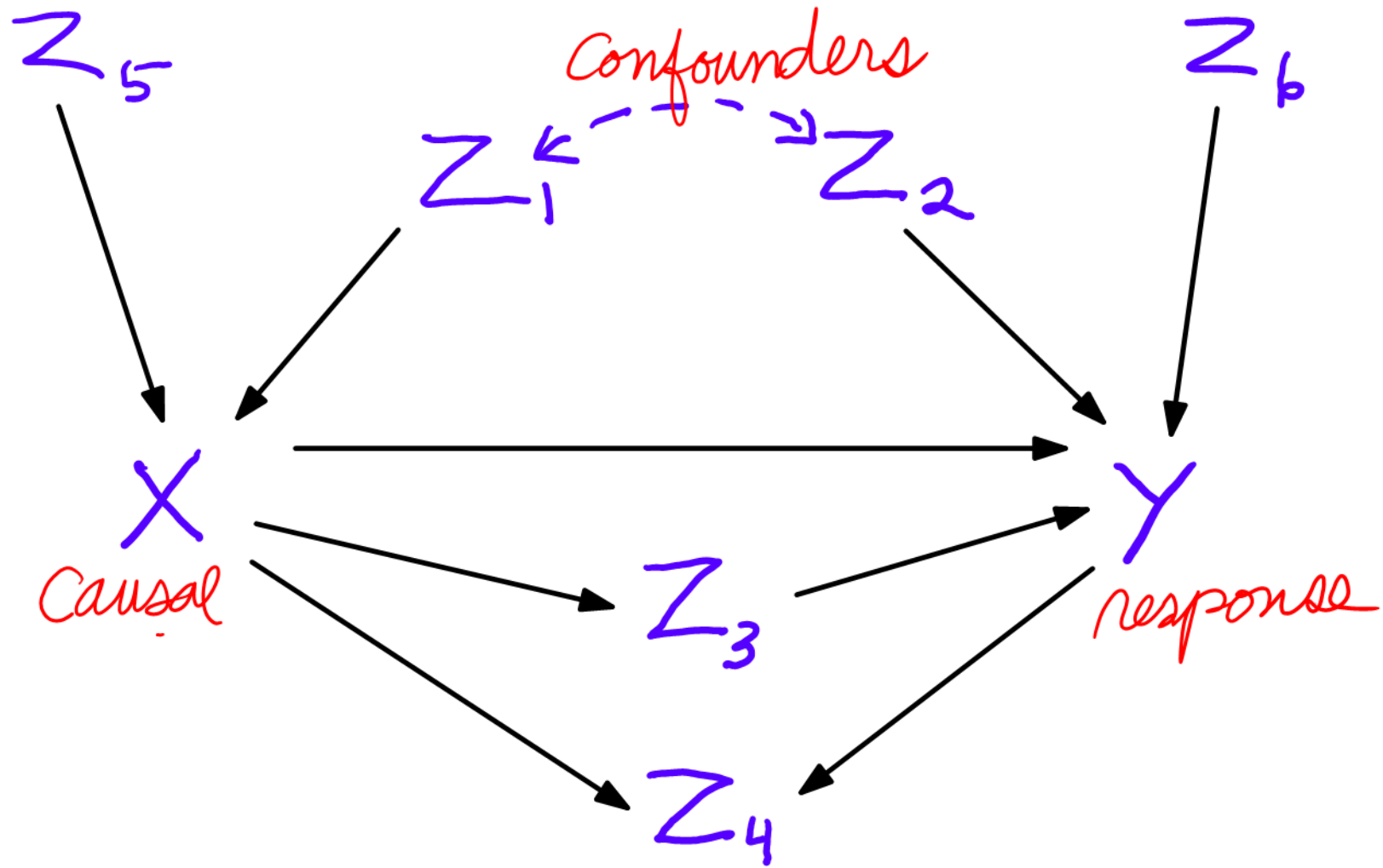


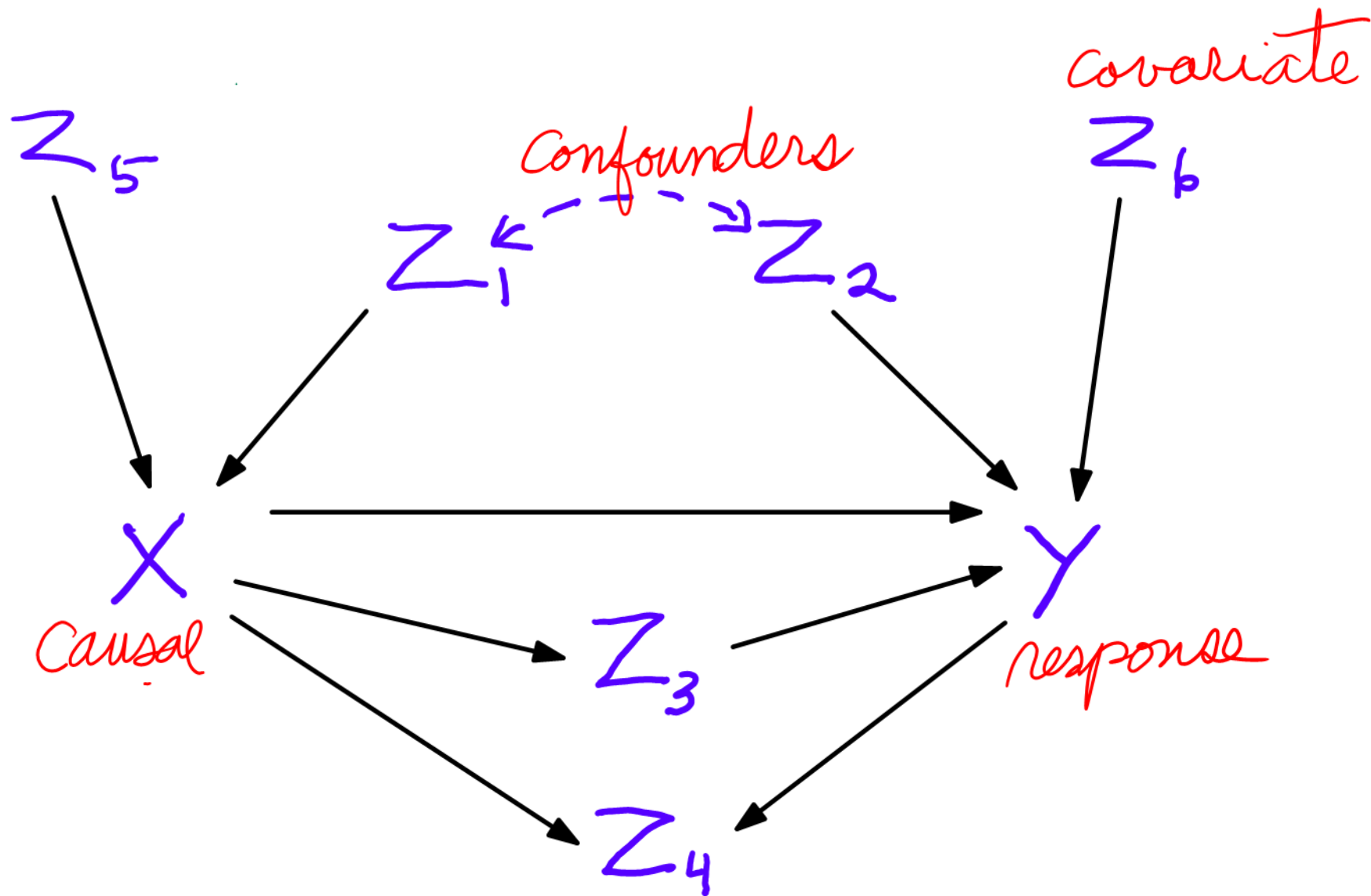
# Note:

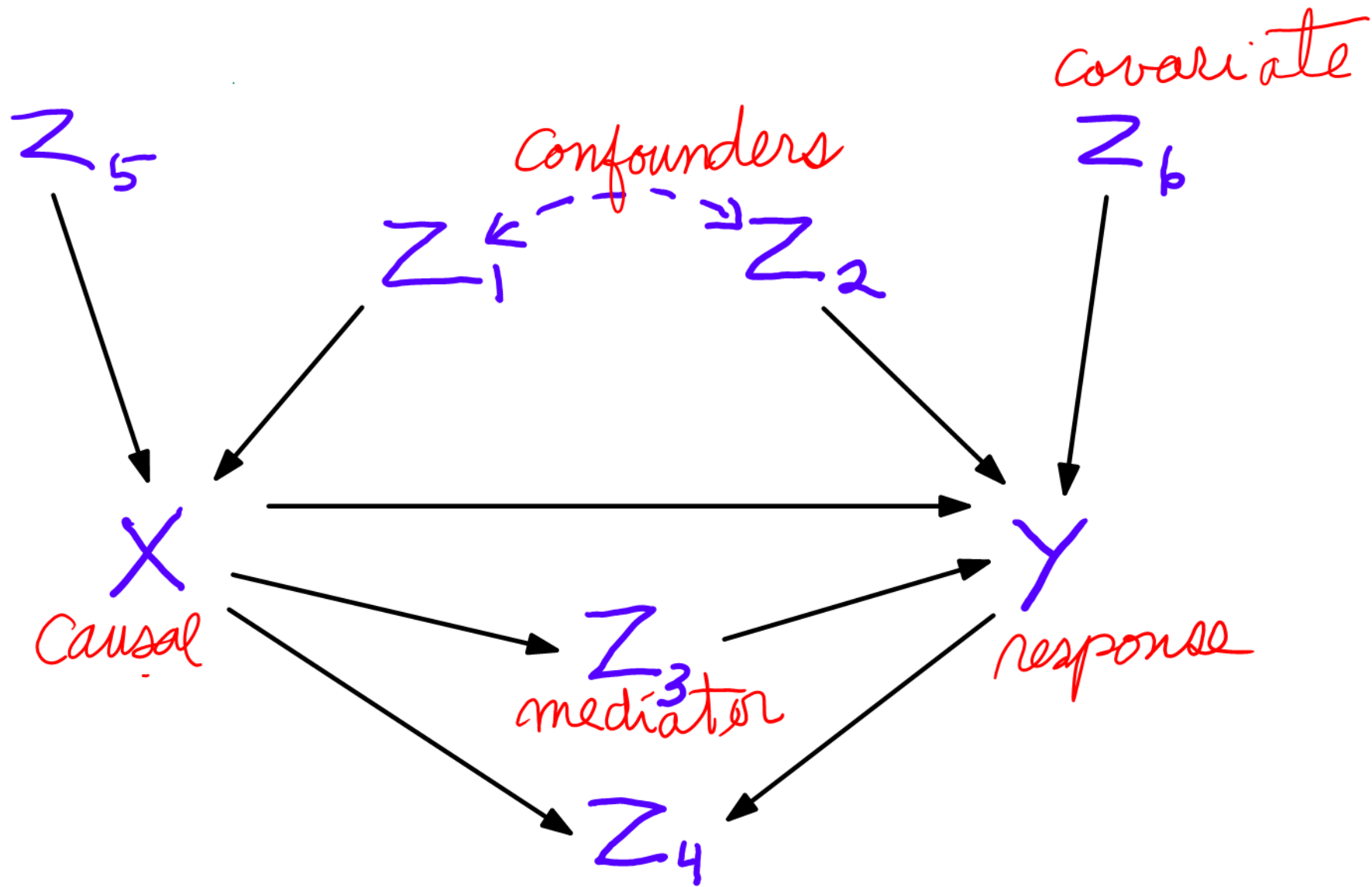
- Simpson effect: reversal of cond'l vs. marginal effects
  - Moderation (= interaction in relation of X and Z with Y)
  - Association (predictive)
  - Causality
- are distinct but not unrelated concepts

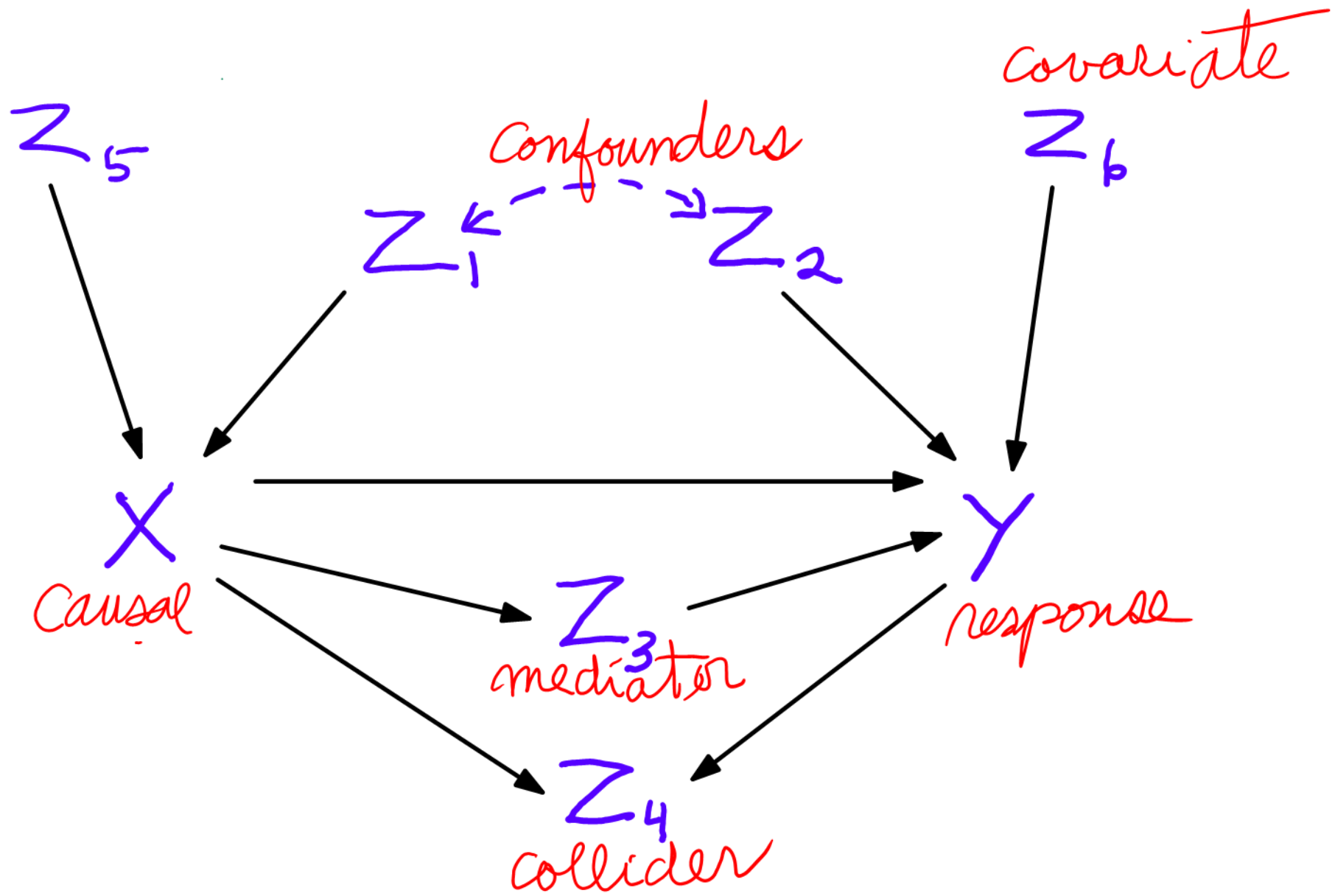


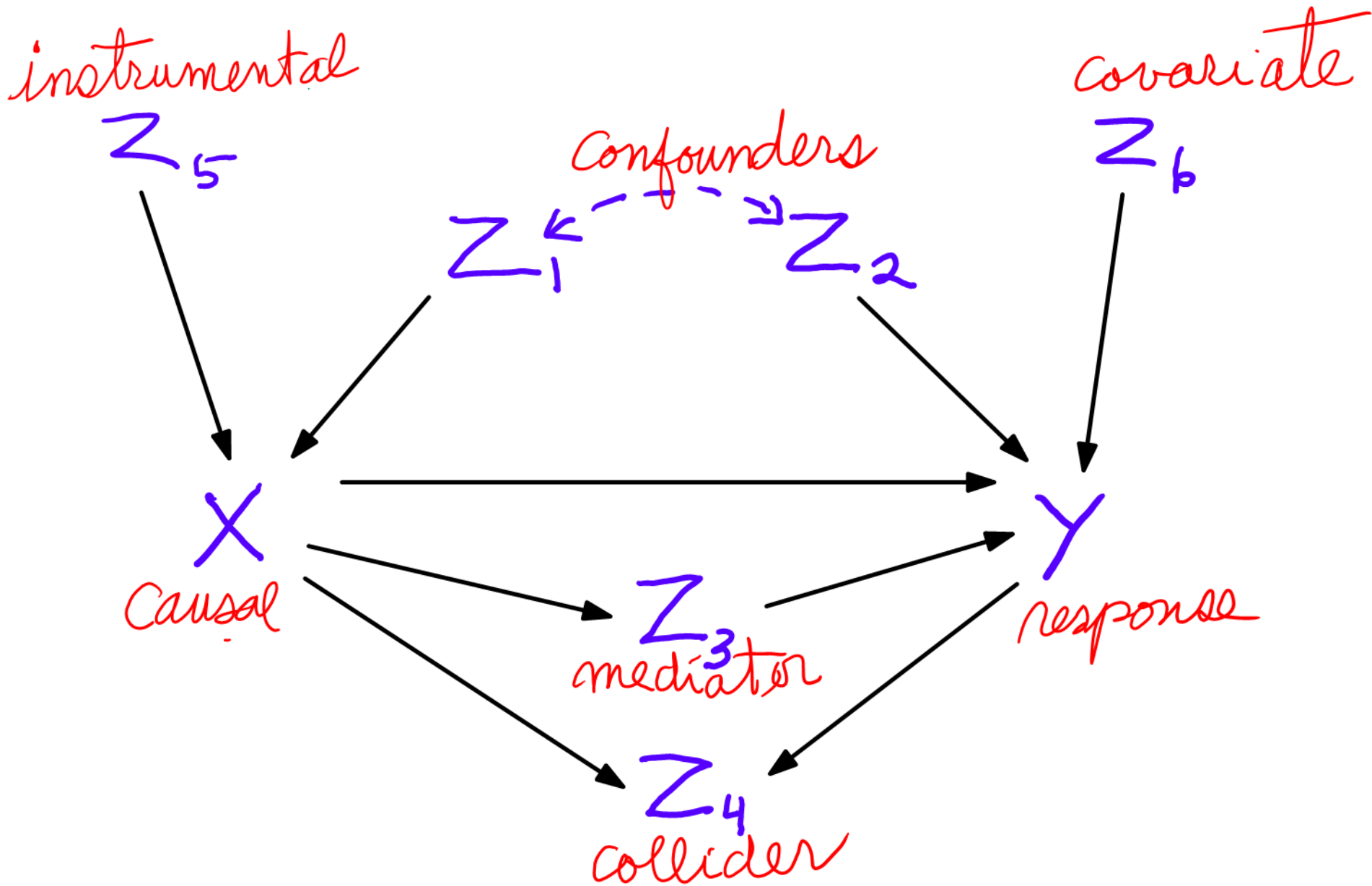


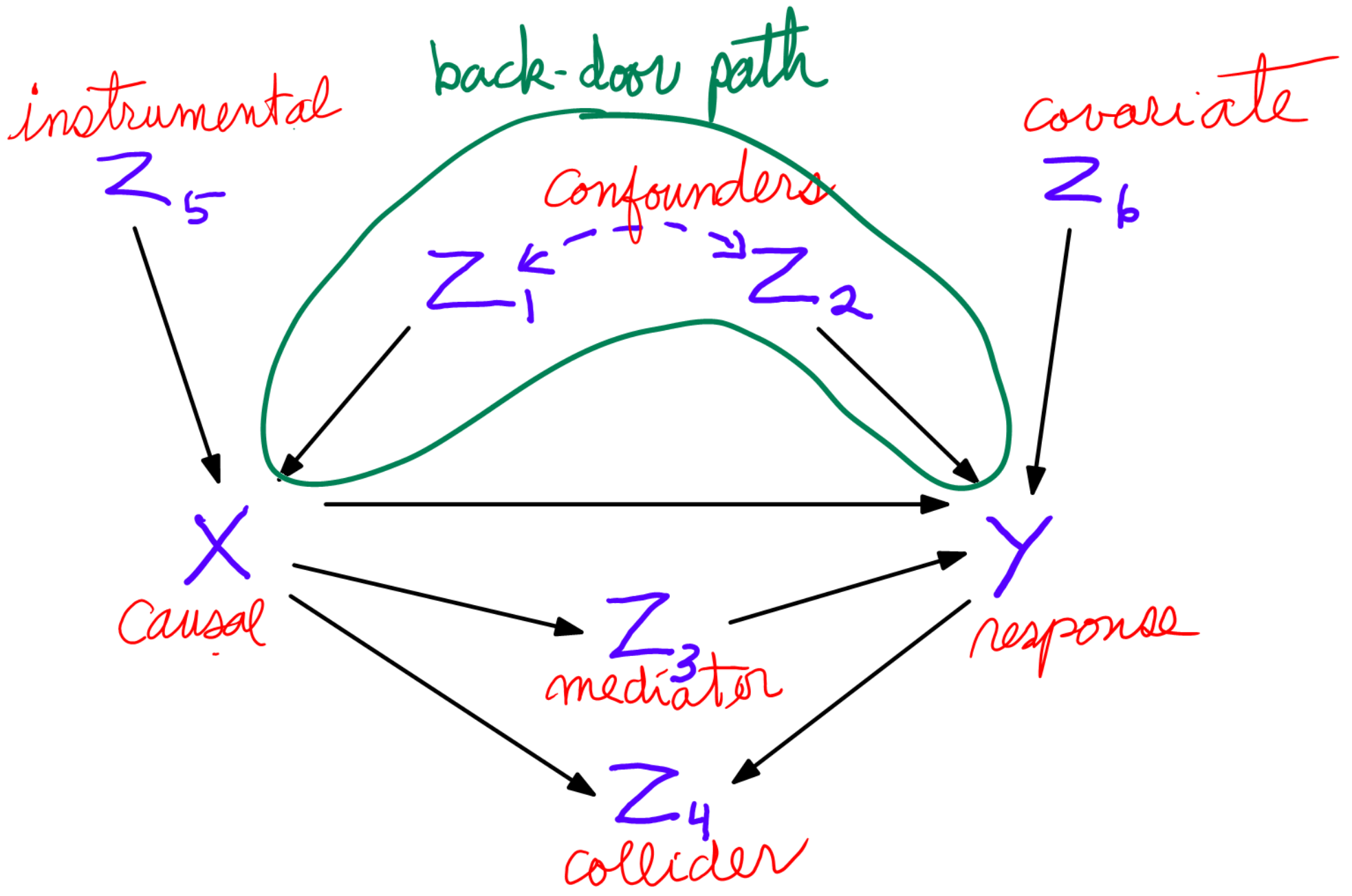


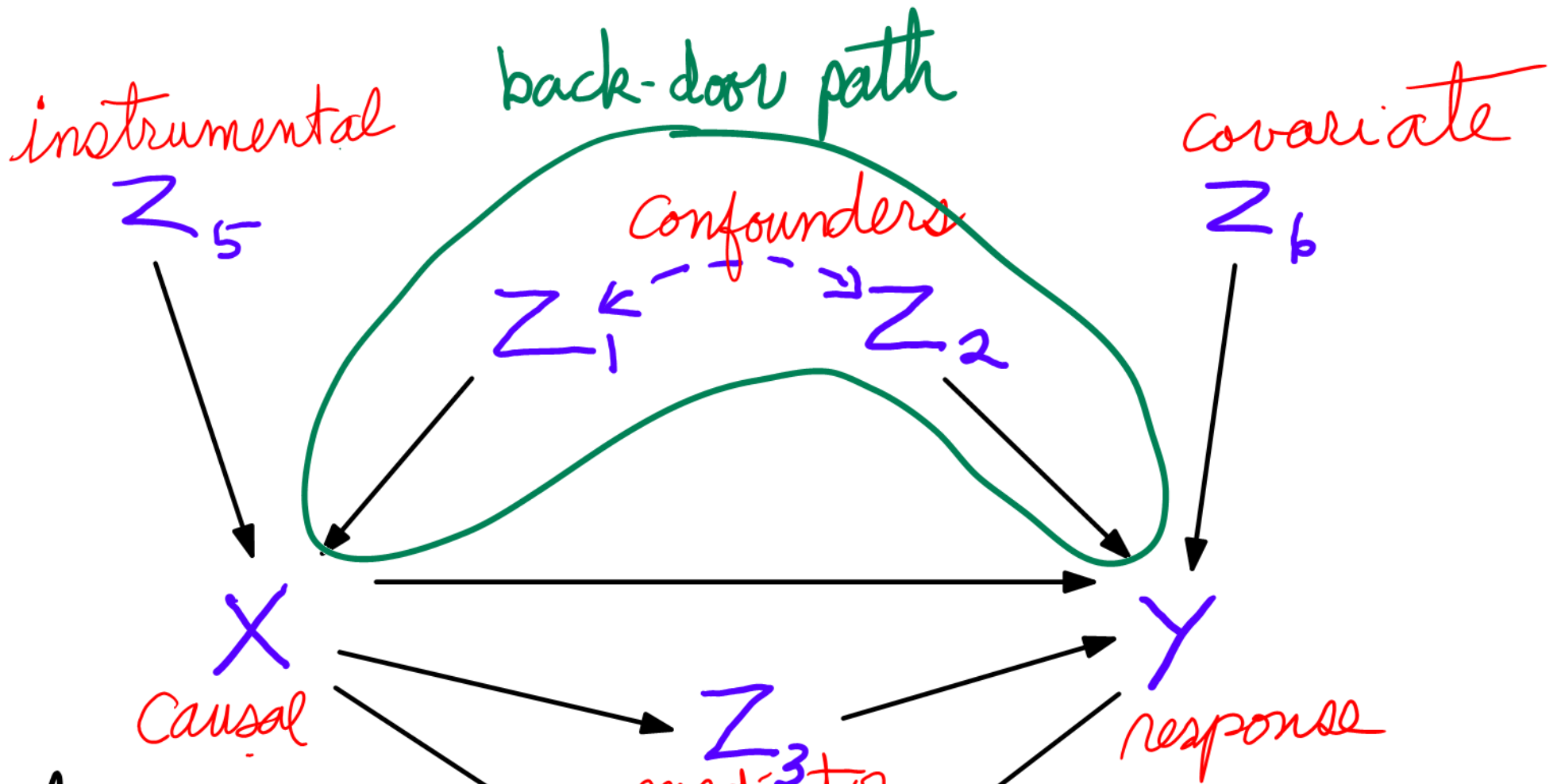




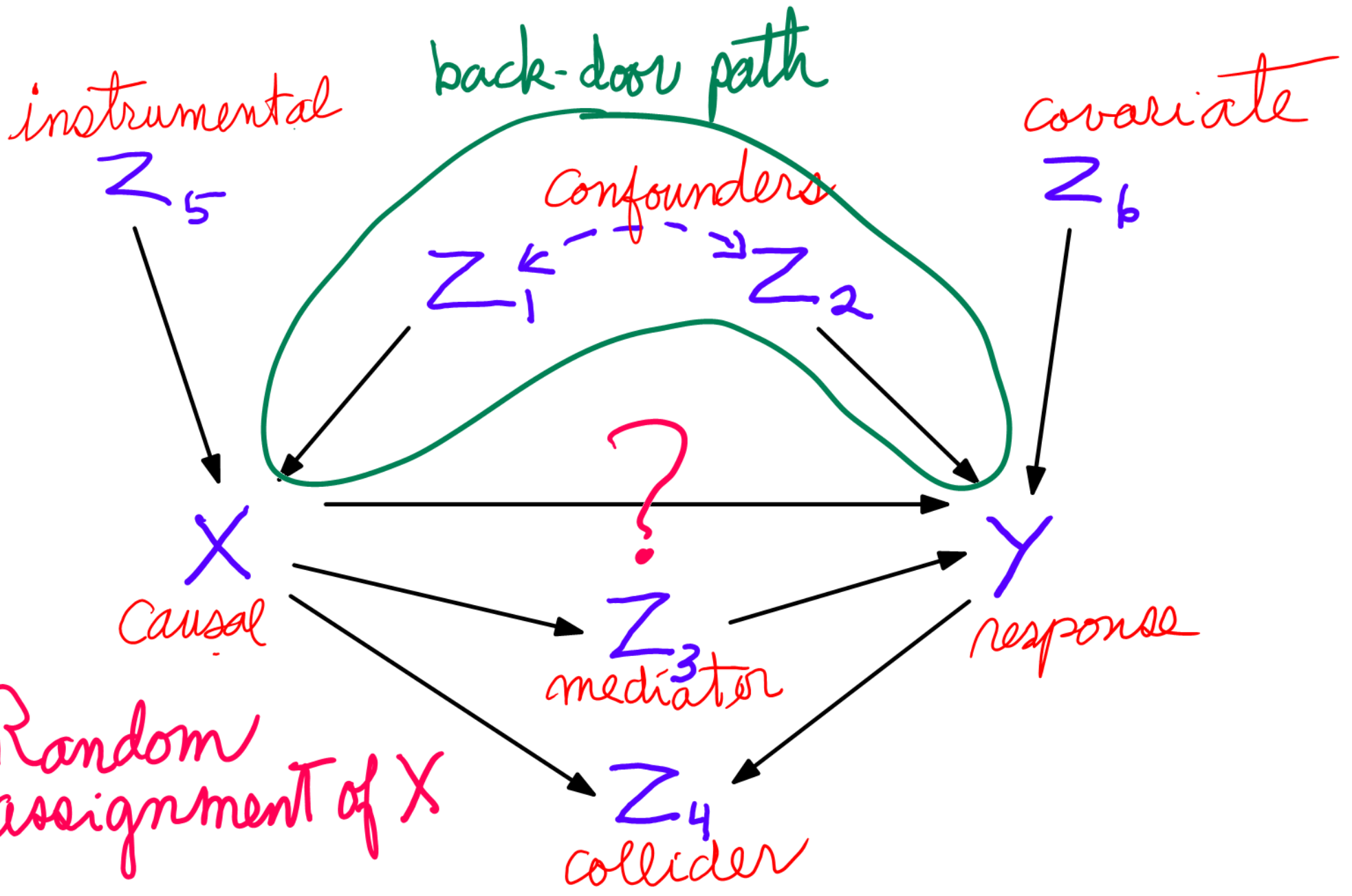






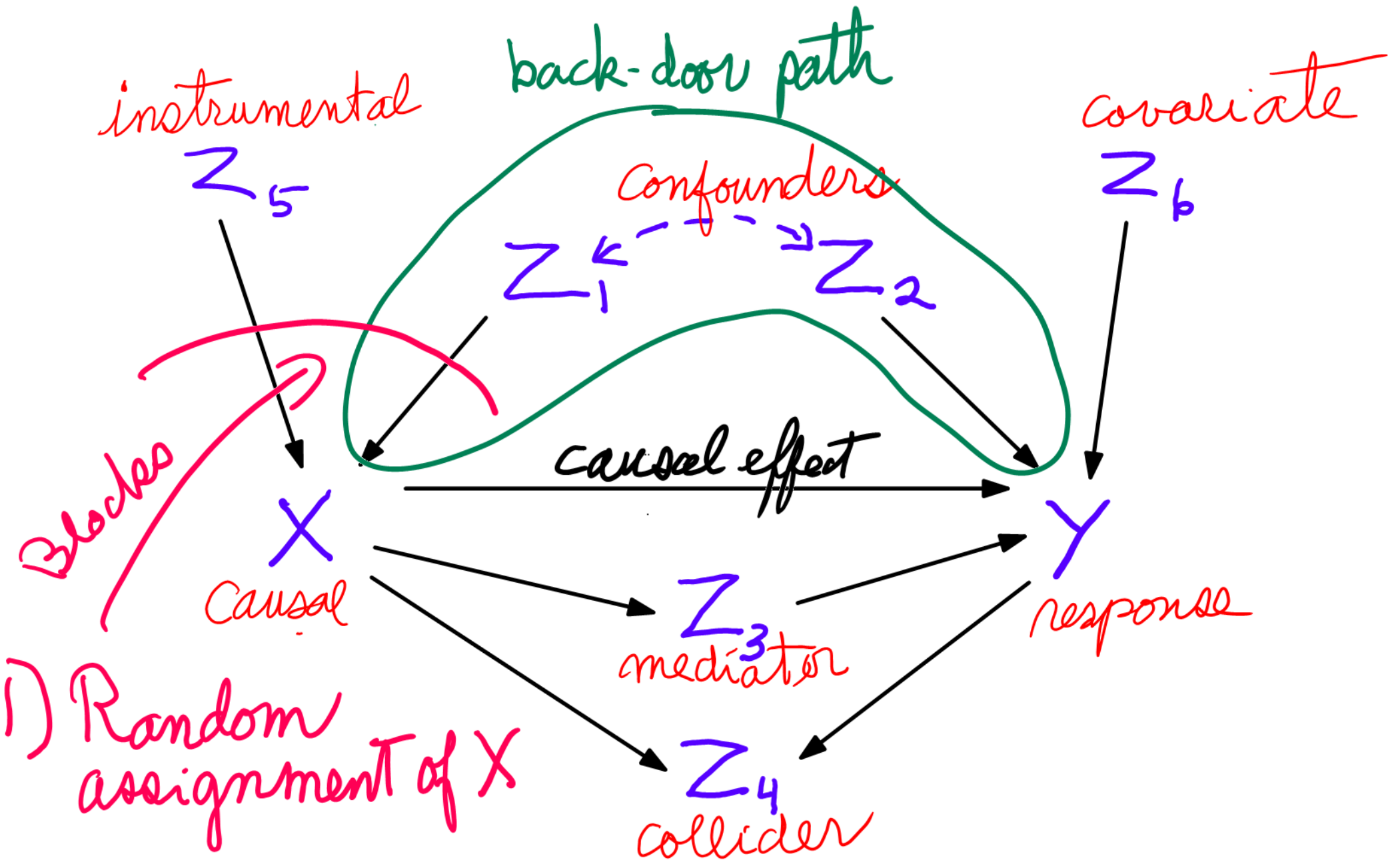


Pearl  
 Must block back-door paths  
 - NOT mediator or collider

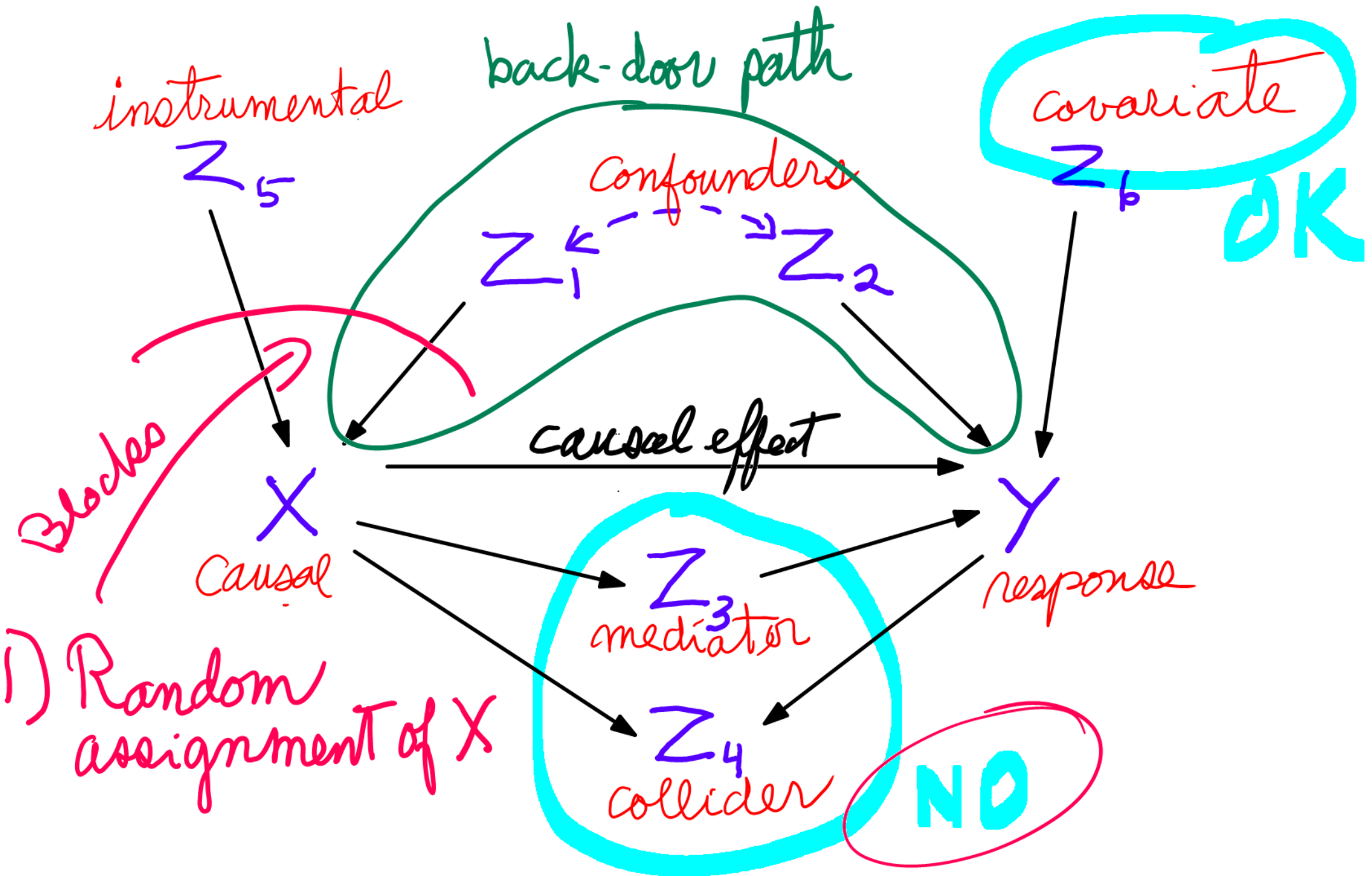


1) Random assignment of  $X$

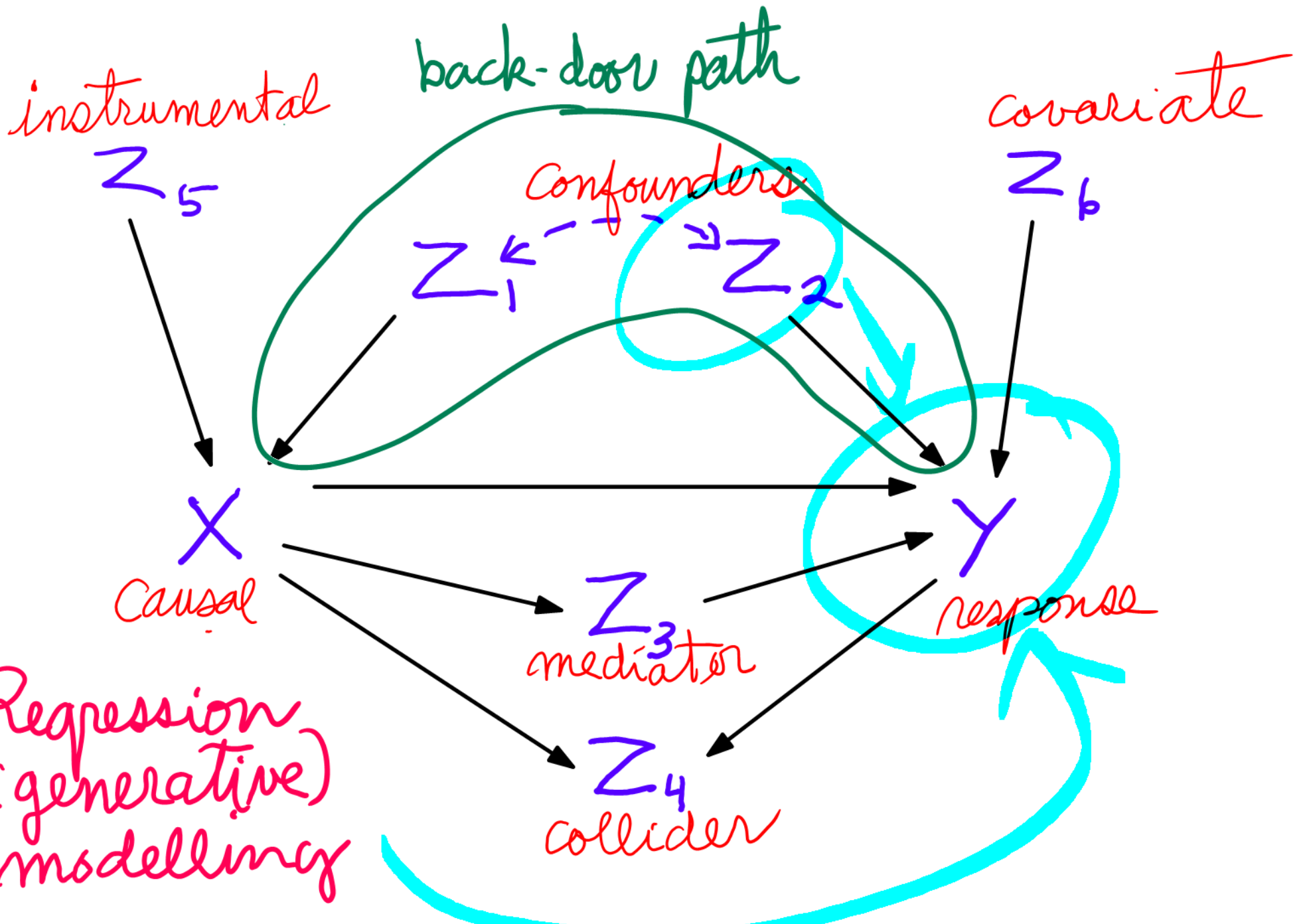




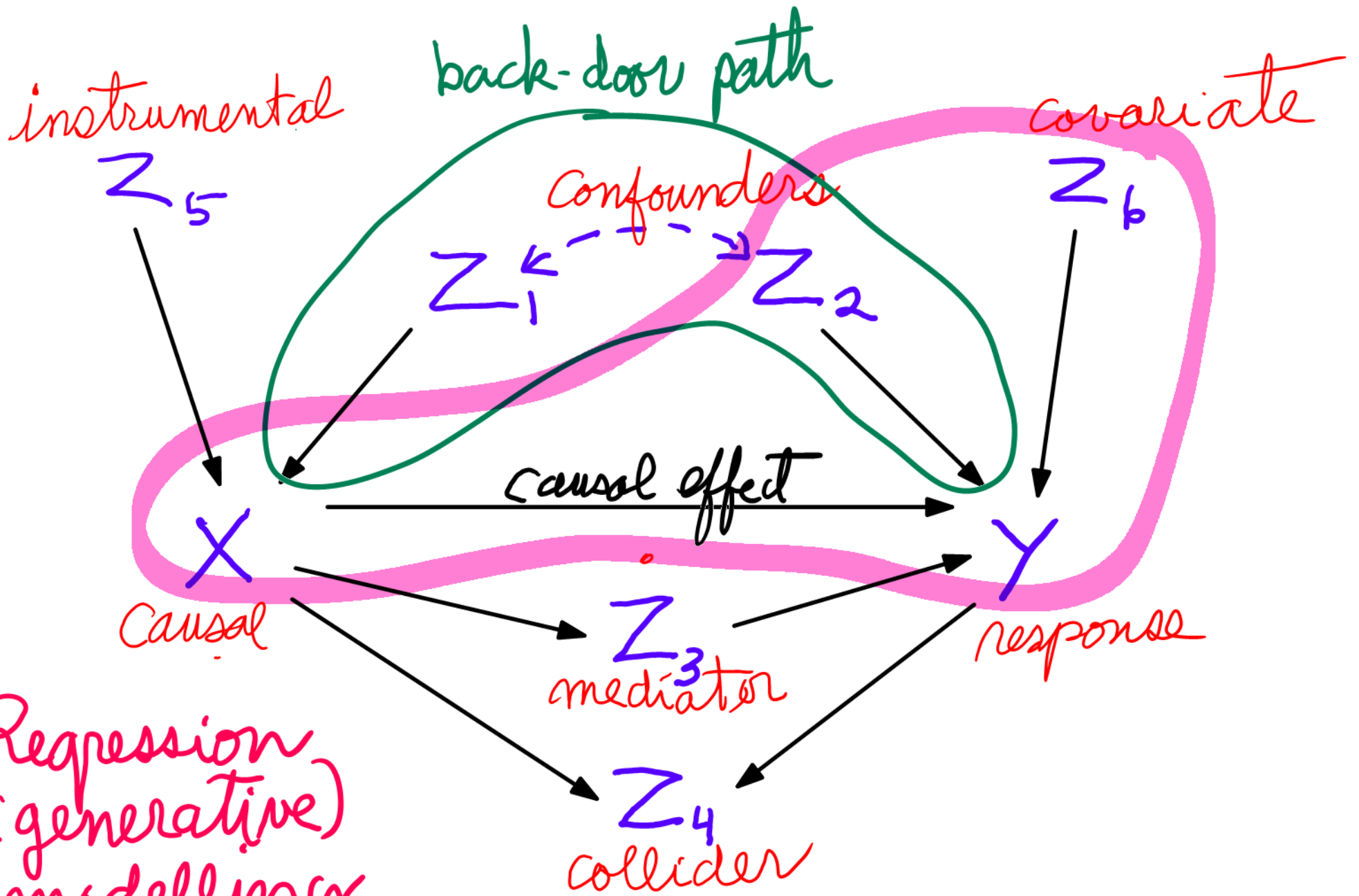
1) Random assignment of X



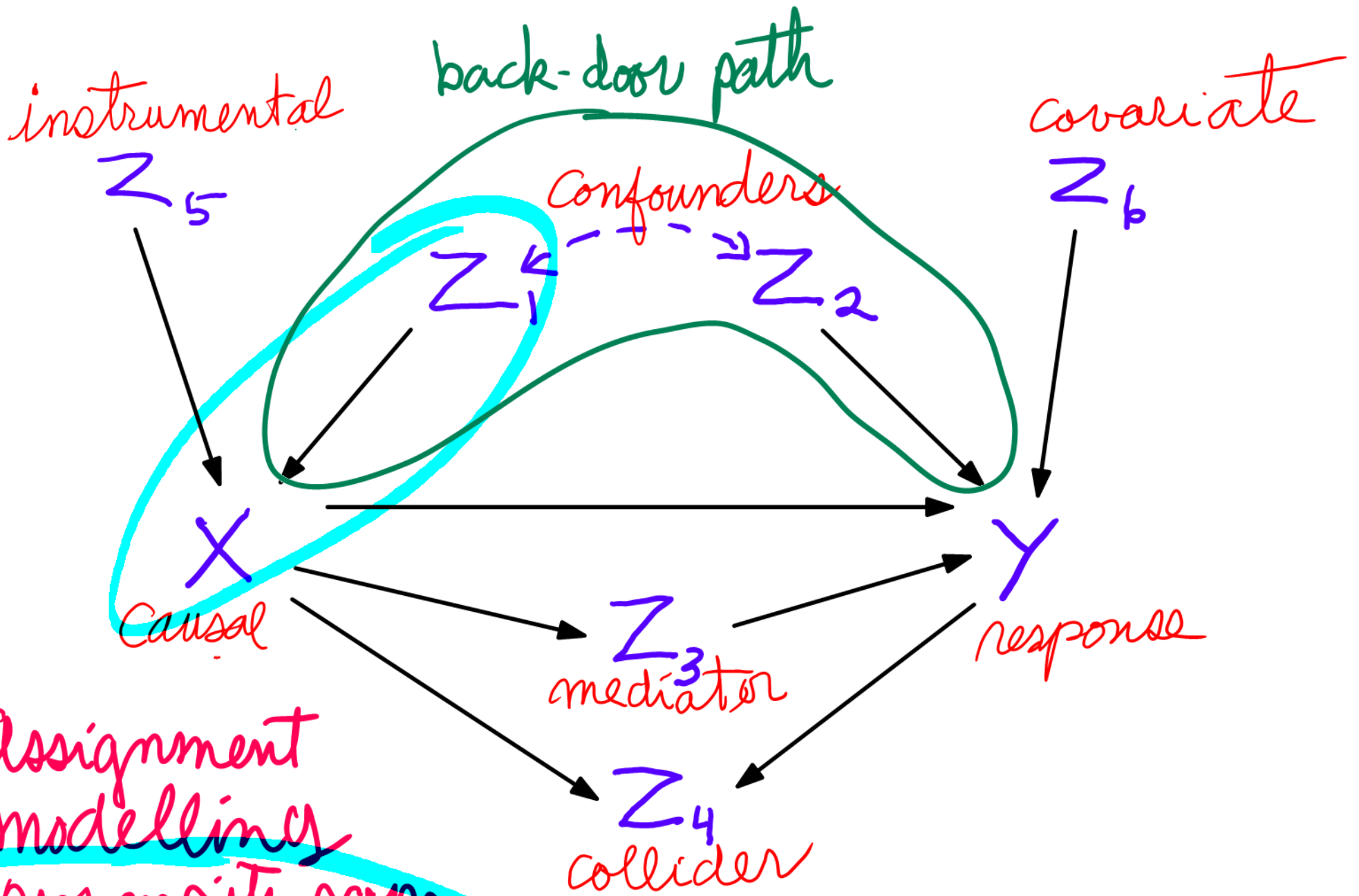
1) Random assignment of X



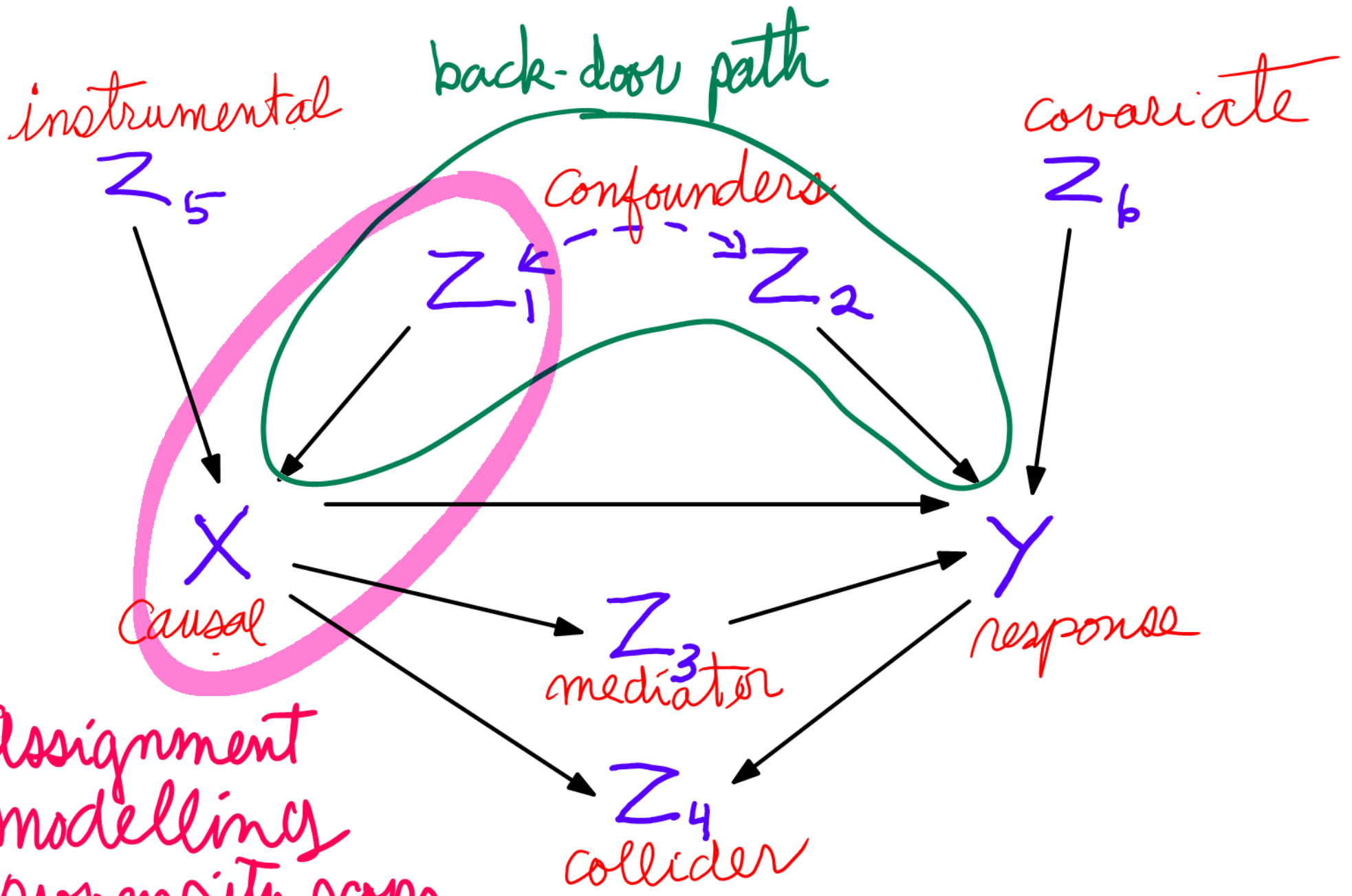
2) Regression (generative) modelling



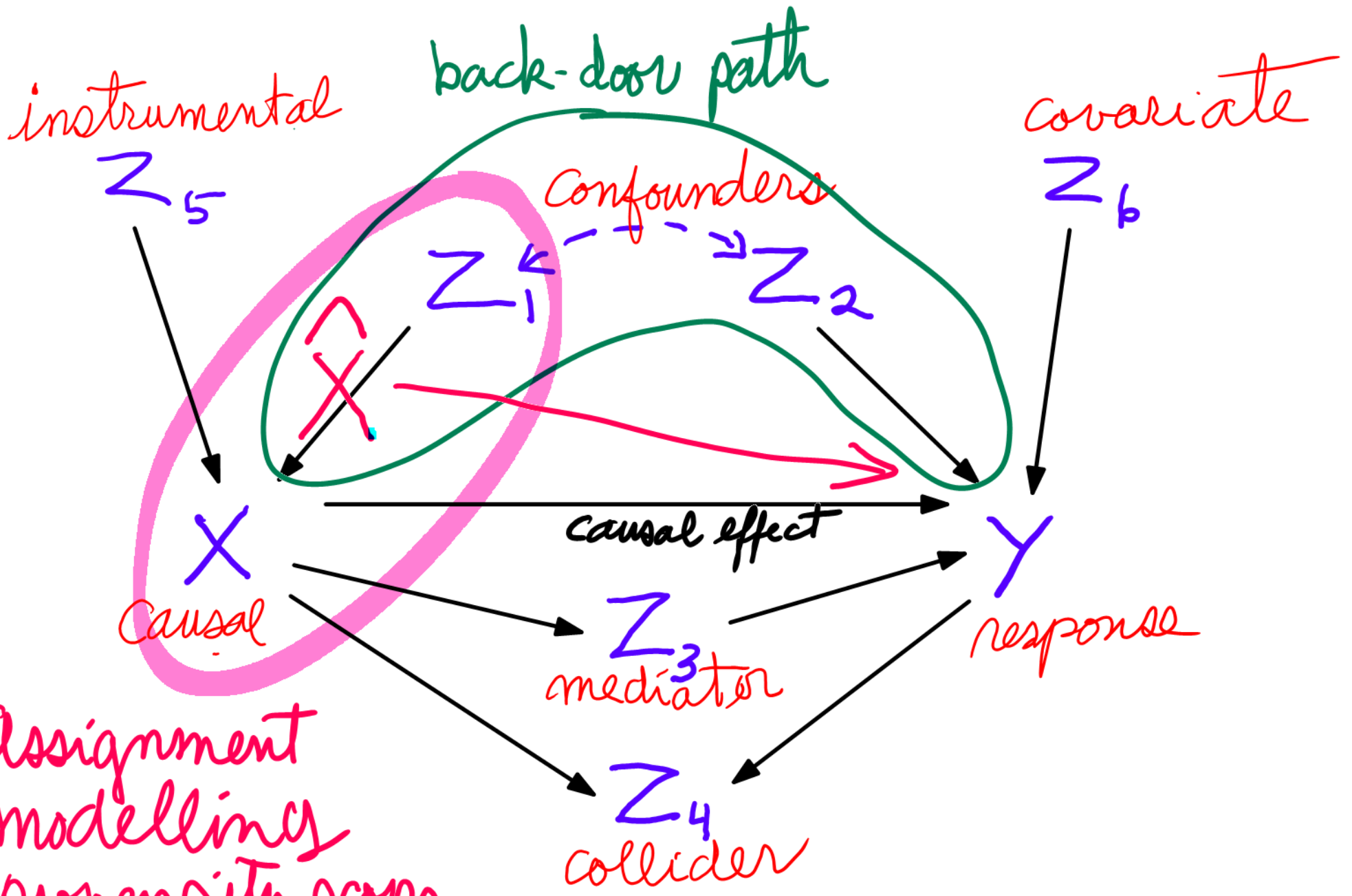
2) Regression (generative) modelling



3) Assignment modelling  
 - propensity scores

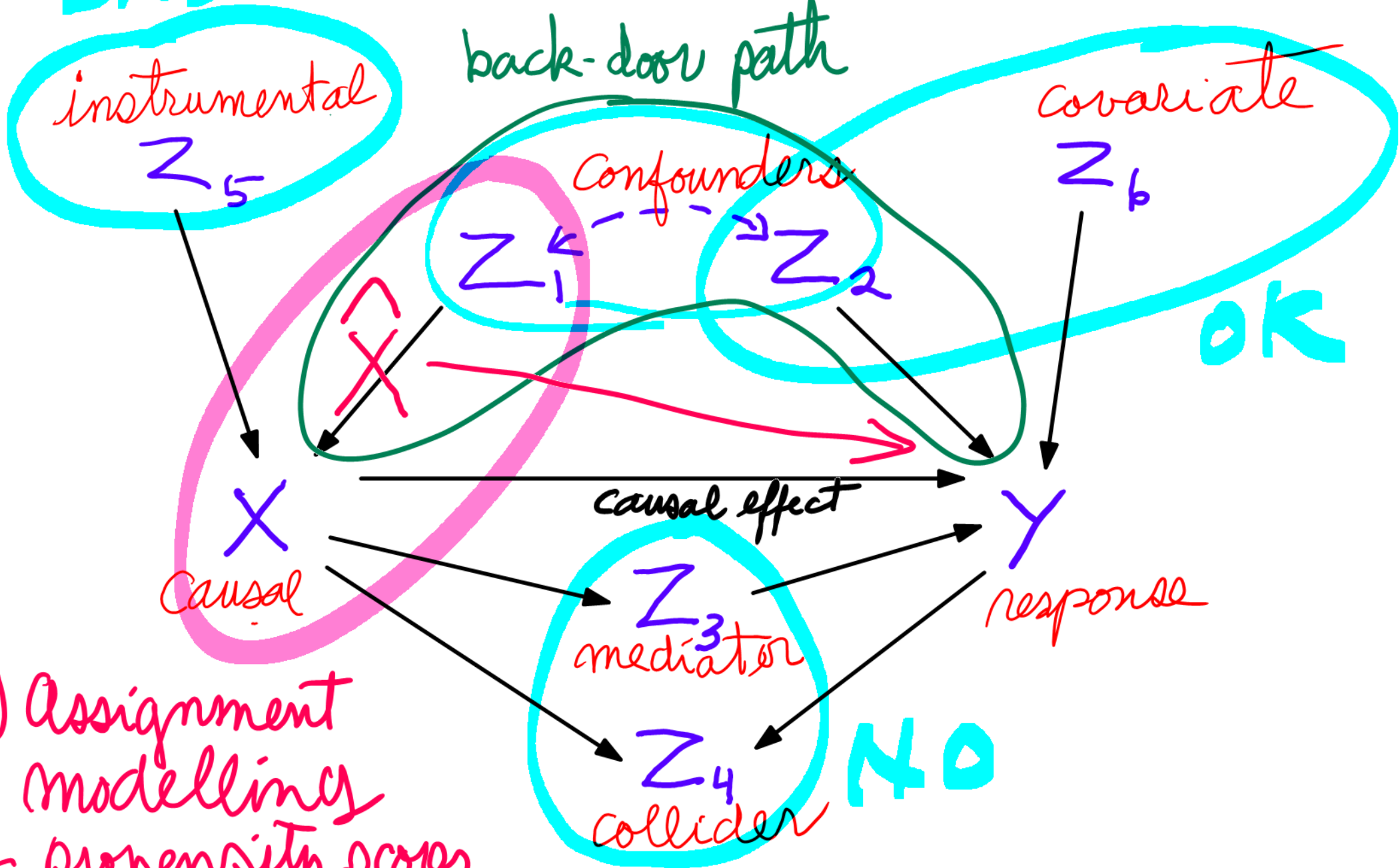


3) Assignment modelling  
 - propensity scores



3) Assignment modelling  
 - propensity scores

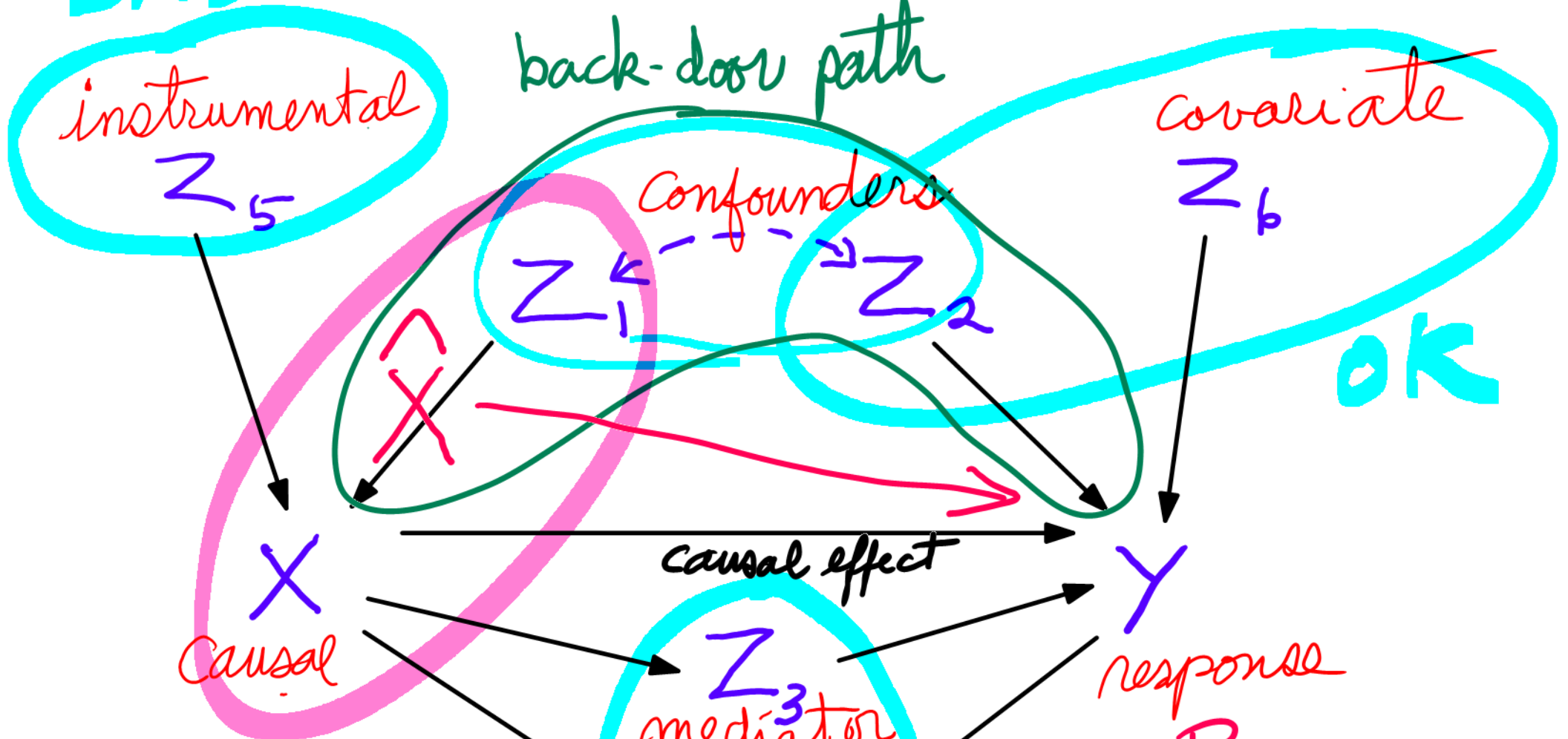
**BAD**



3) Assignment modelling - propensity scores



**BAD**



**OK**

**NO**

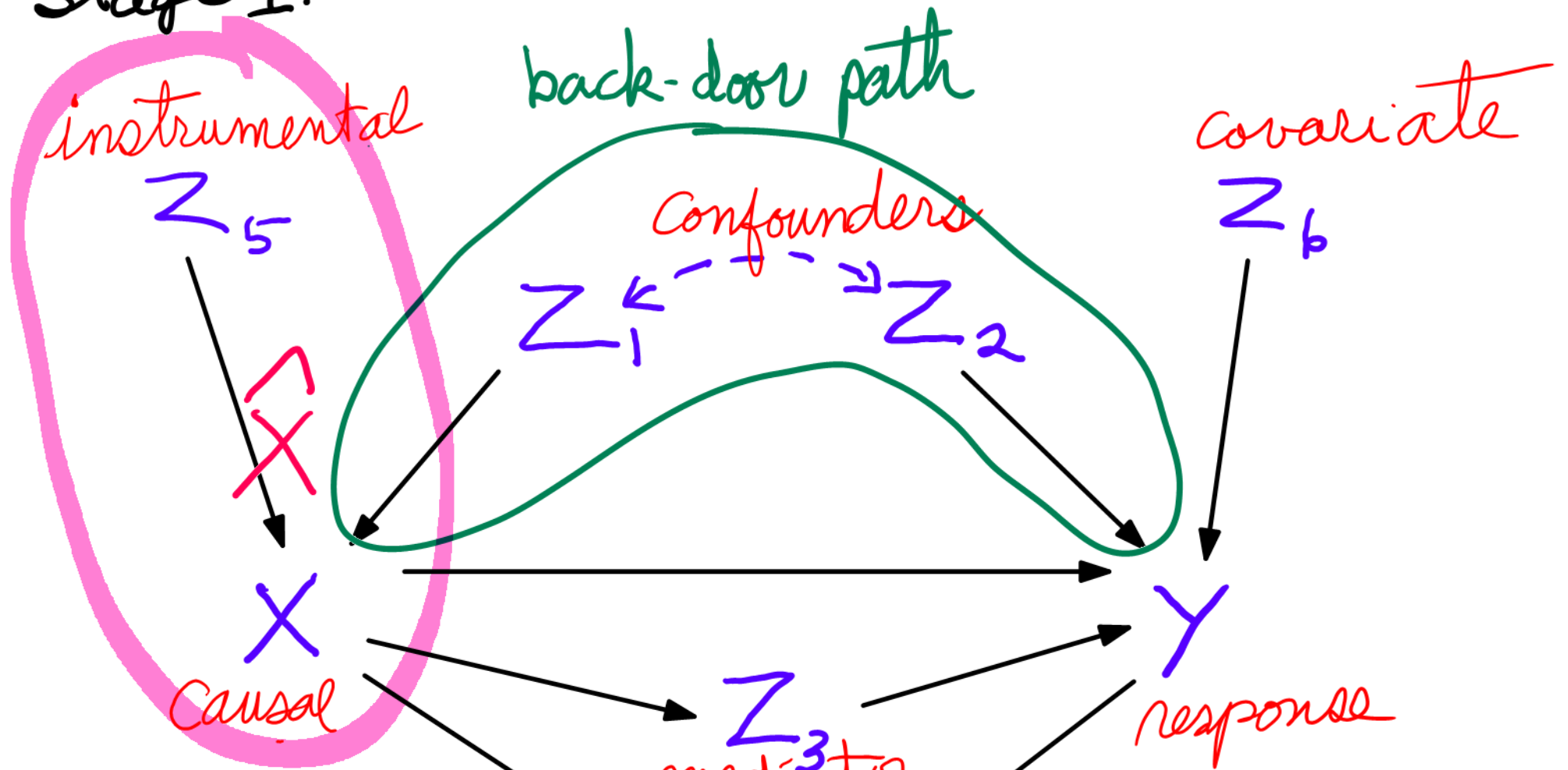
Regress:

$$Y \sim X + \hat{X} + \dots$$

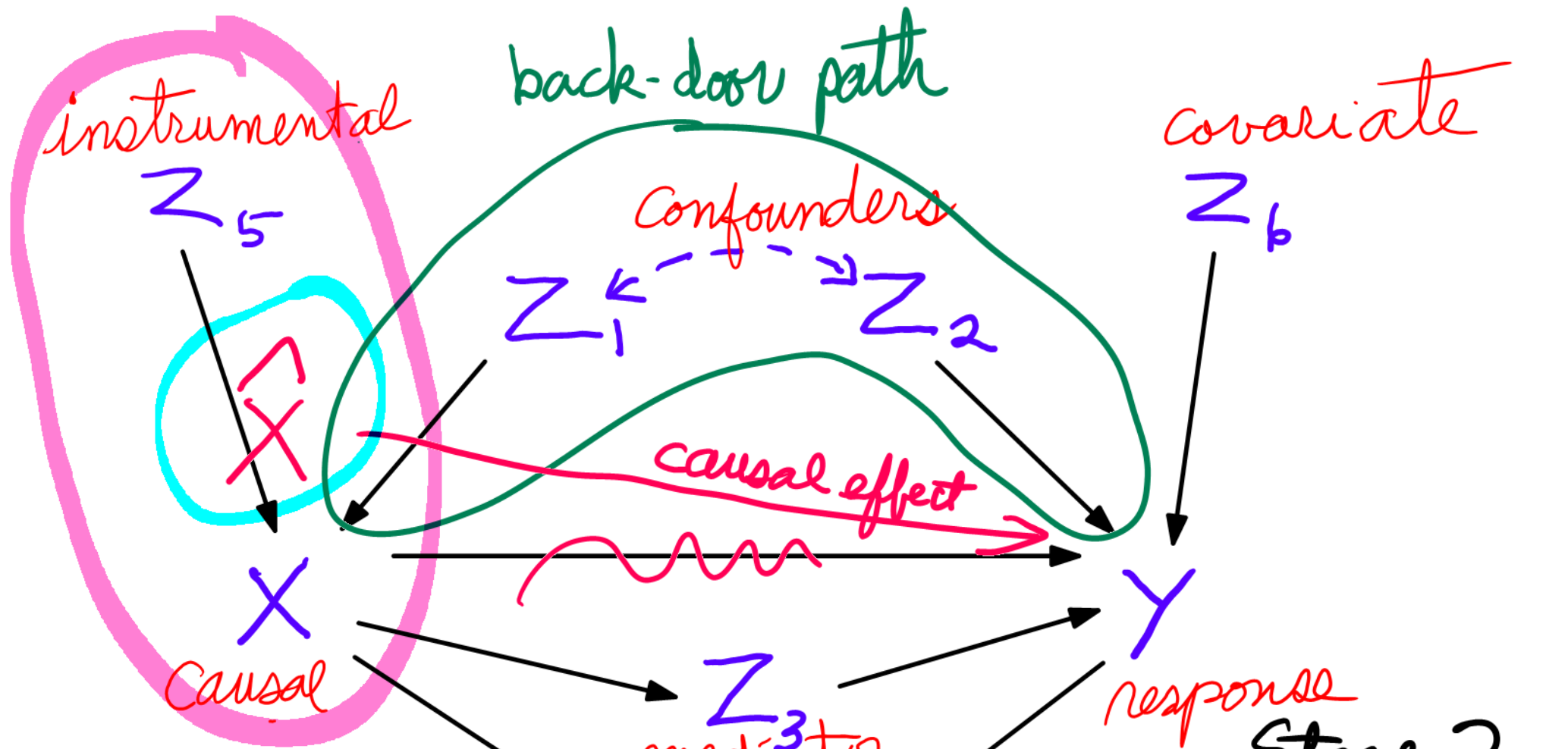
Causal effect is  $\beta_x$

3) Assignment modelling  
- propensity scores

Stage 1:



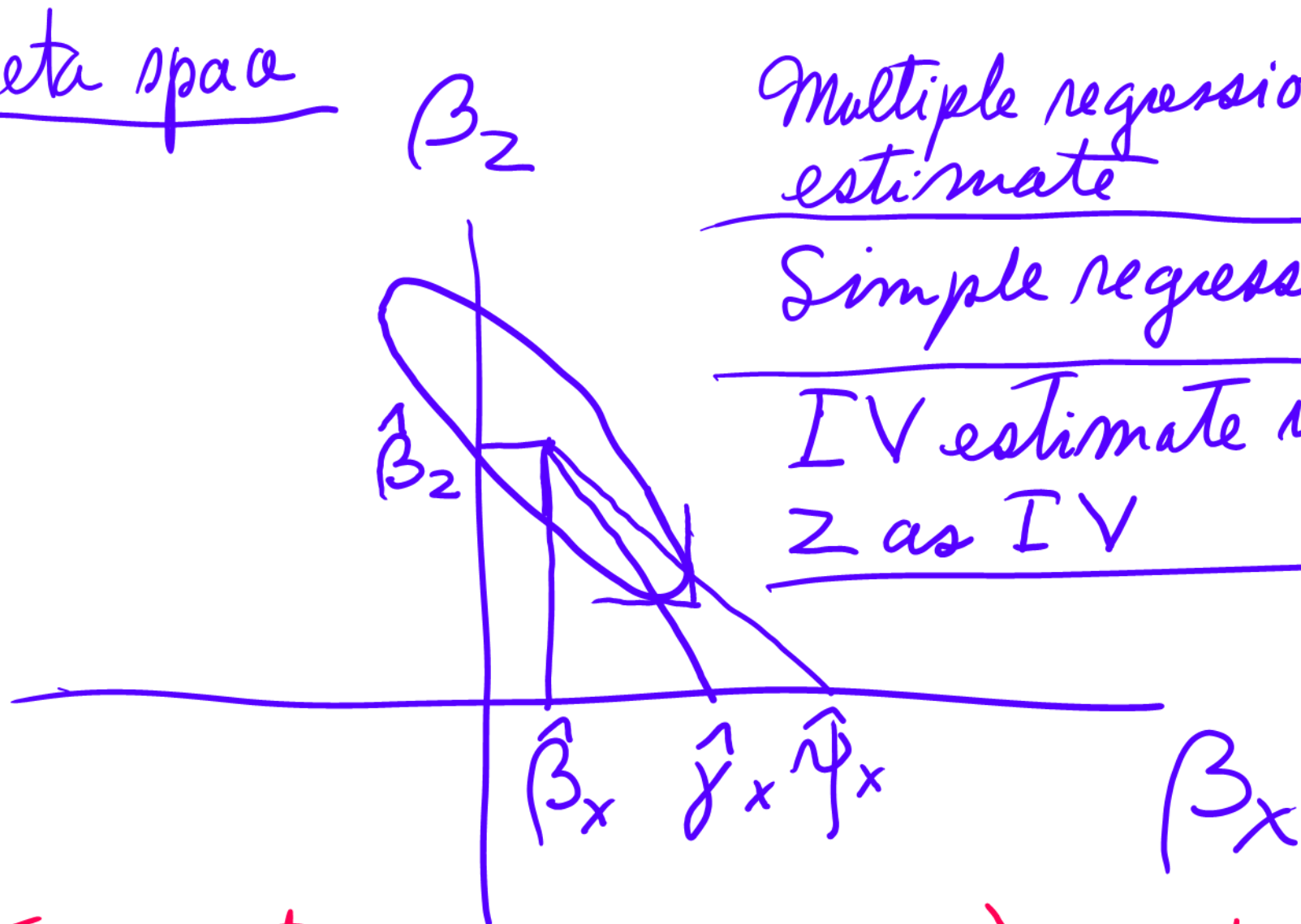
4) 2-stage least-squares  
instrumental  
variables



4) 2-stage least-squares  
instrumental  
variables

Stage 2  
Regress:  
 $Y \sim \hat{X}$   
Causal effect is  $\beta_{\hat{X}}$

# Beta space



Multiple regression estimate

$\hat{\beta}_x$

Simple regression

$\hat{\delta}_x$

IV estimate with  $Z$  as IV

$\hat{\psi}_x$

For a strong IV,  $\text{Corr}(X, Z)$  close to 1 and exclusion restriction  $\Rightarrow \beta_2$  close to 0

So  $\hat{\psi}_x$  close to  $\hat{\delta}_x$

Note that we can't test the assumption of "exclusion restriction" by looking at the coefficient of the instrumental variable  $I$  in the regression

$$Y \sim I + X$$

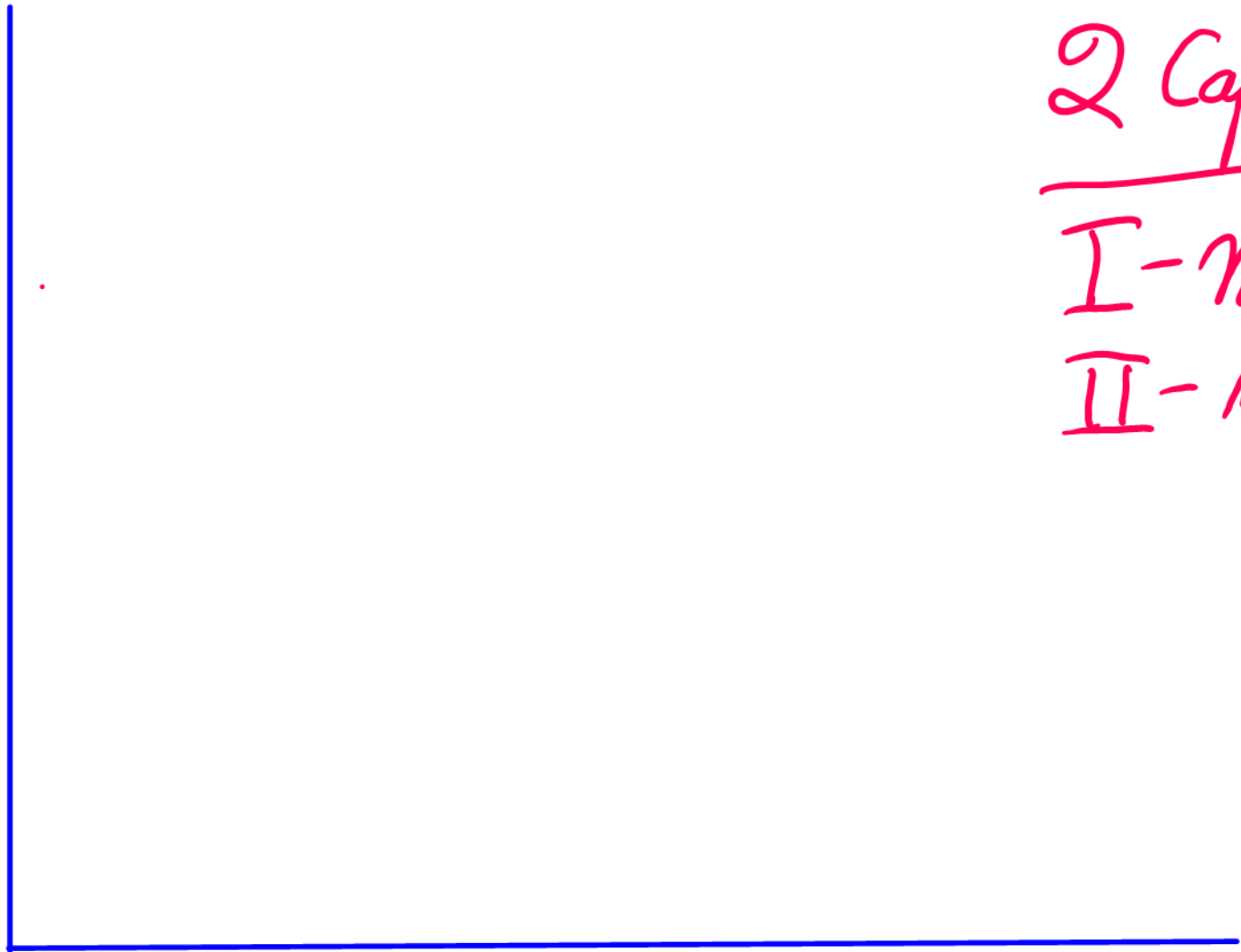
since  $X$  is a collider if there is an omitted confounder  $C$



and  $\hat{\beta}_I$  should not be 0 even if  $I$  is a good instrument.

# Lord's Paradox (Wainer version)

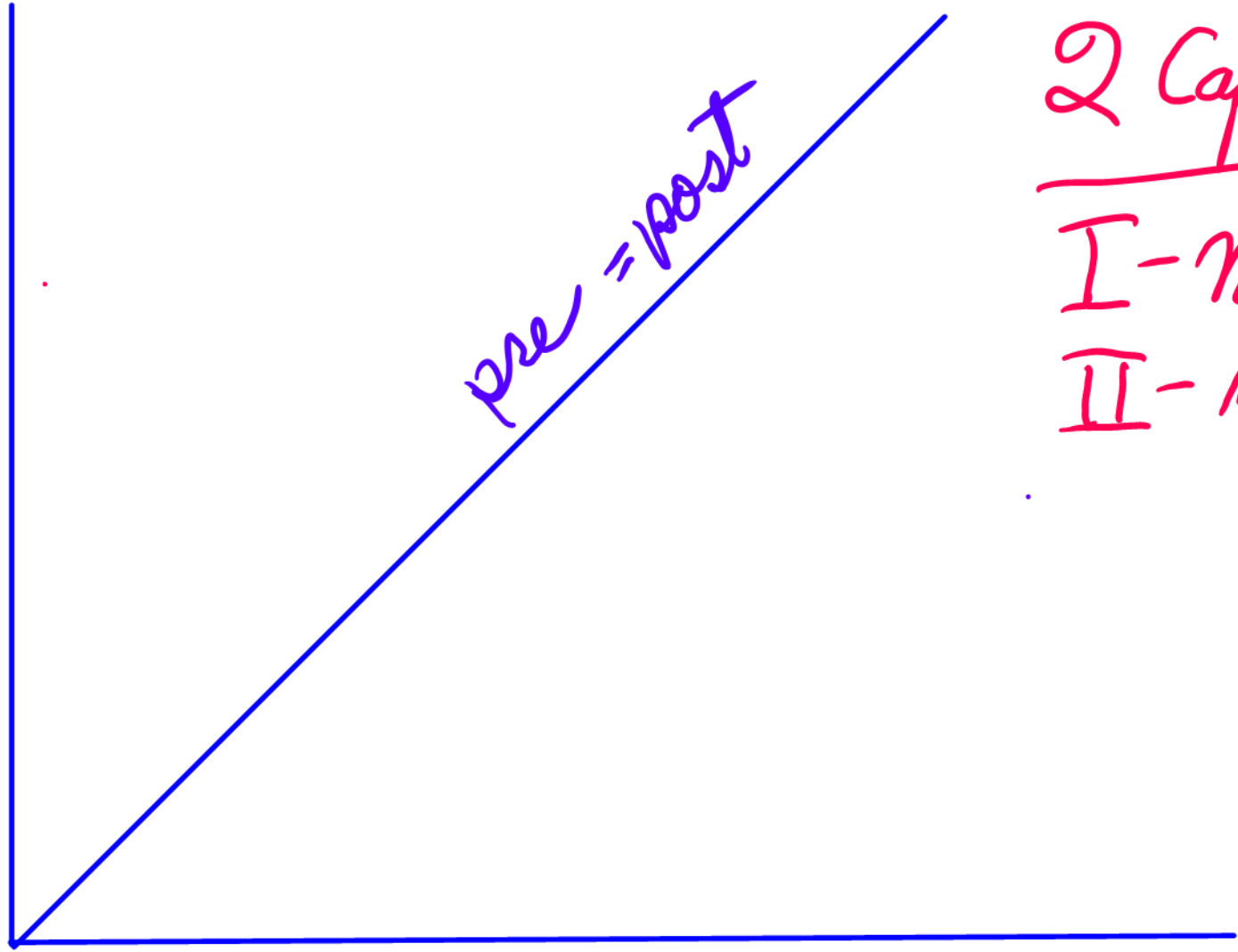
$Y_2$   
post



2 Cafeterias  
I - normal  
II - weight loss

$Y_1$  pre

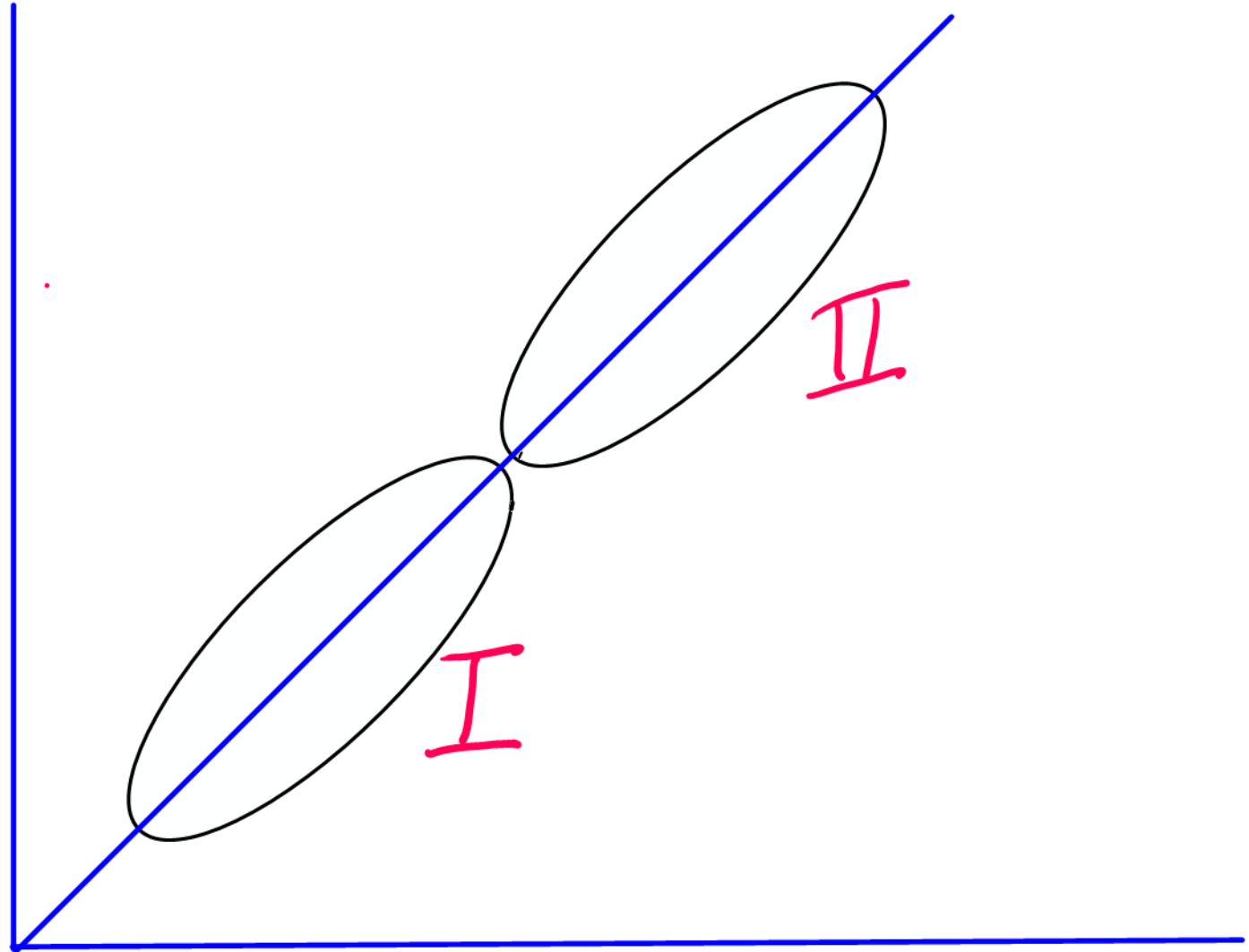
$y_2$   
post



2 Cafeterias  
I - normal  
II - weight loss

$y_1$  pre

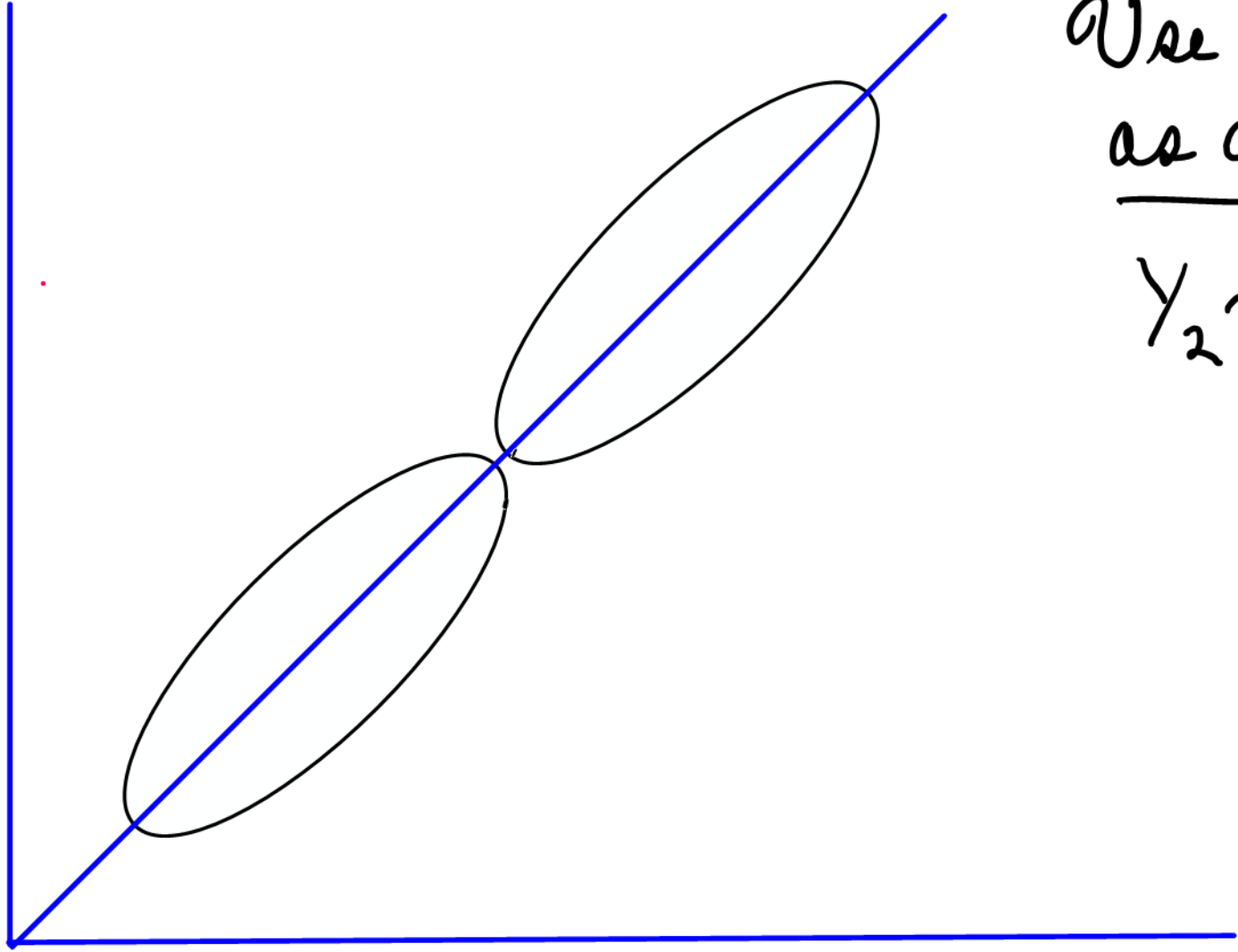
$y_2$   
post



$y_1$  pre



$Y_2$   
post



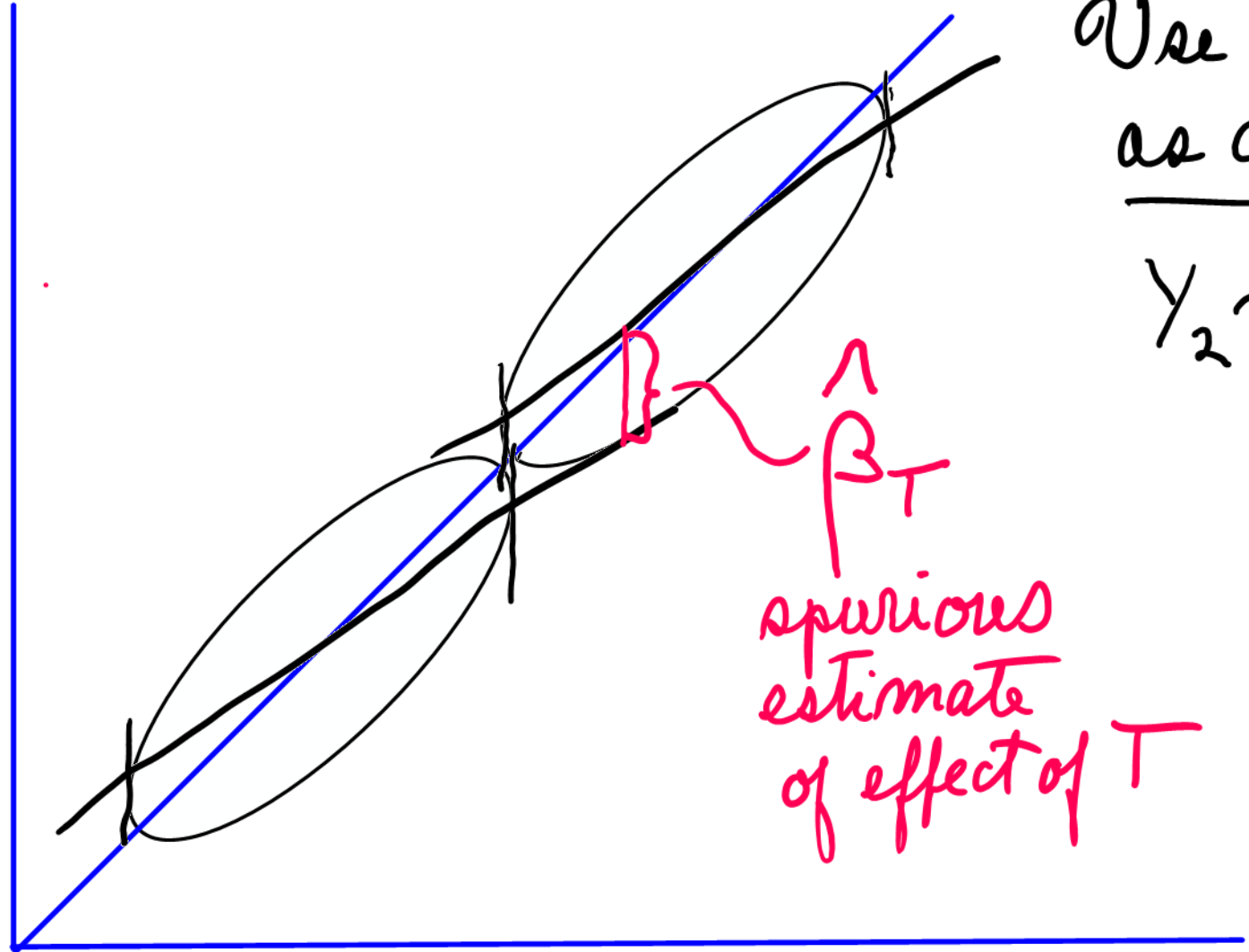
Use pretest  
as covariate

---

$$Y_2 \sim T + Y_1$$

$Y_1$  pre

$Y_2$   
post



Use pretest  
as covariate

$$Y_2 \sim T + Y_1$$

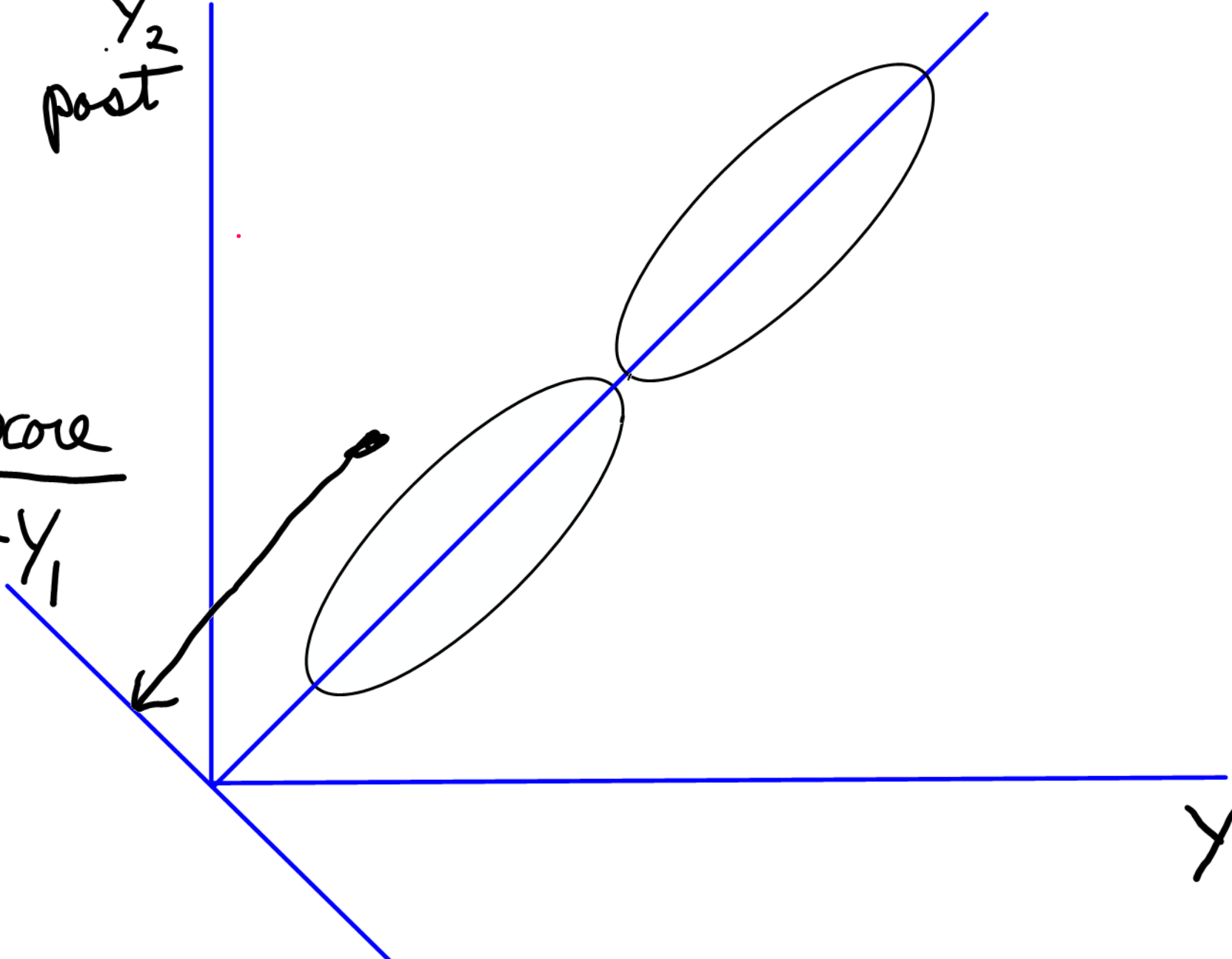
$\hat{\beta}_T$   
spurious  
estimate  
of effect of T

$Y_1$  pre

$y_2$   
post

gain score

$$G = y_2 - y_1$$



$y_1$  pre

$y_2$   
post

gain score

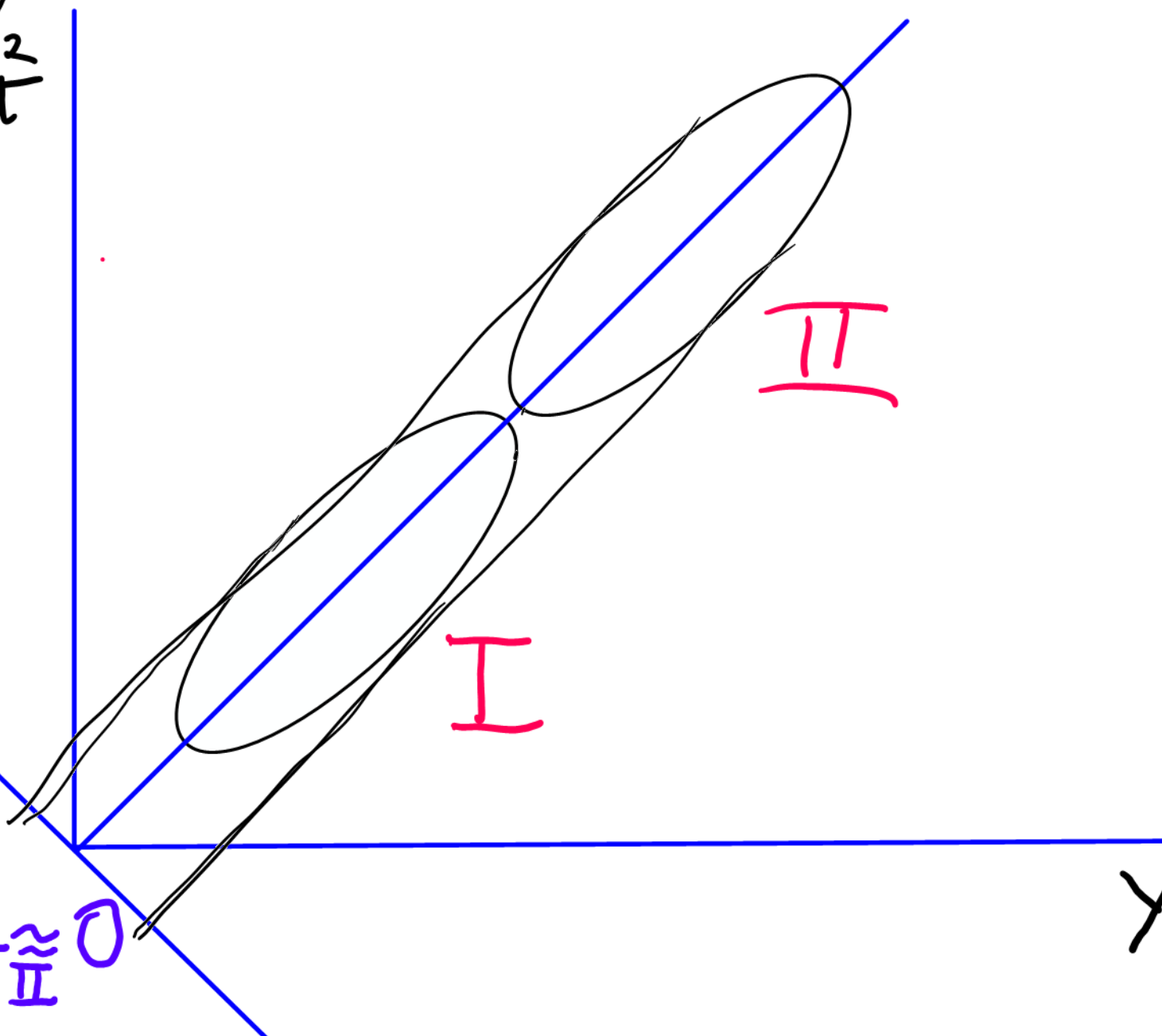
$$G = y_2 - y_1$$

II

I

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

$y_1$  pre



$y_2$   
post

gain score

$$G = y_2 - y_1$$

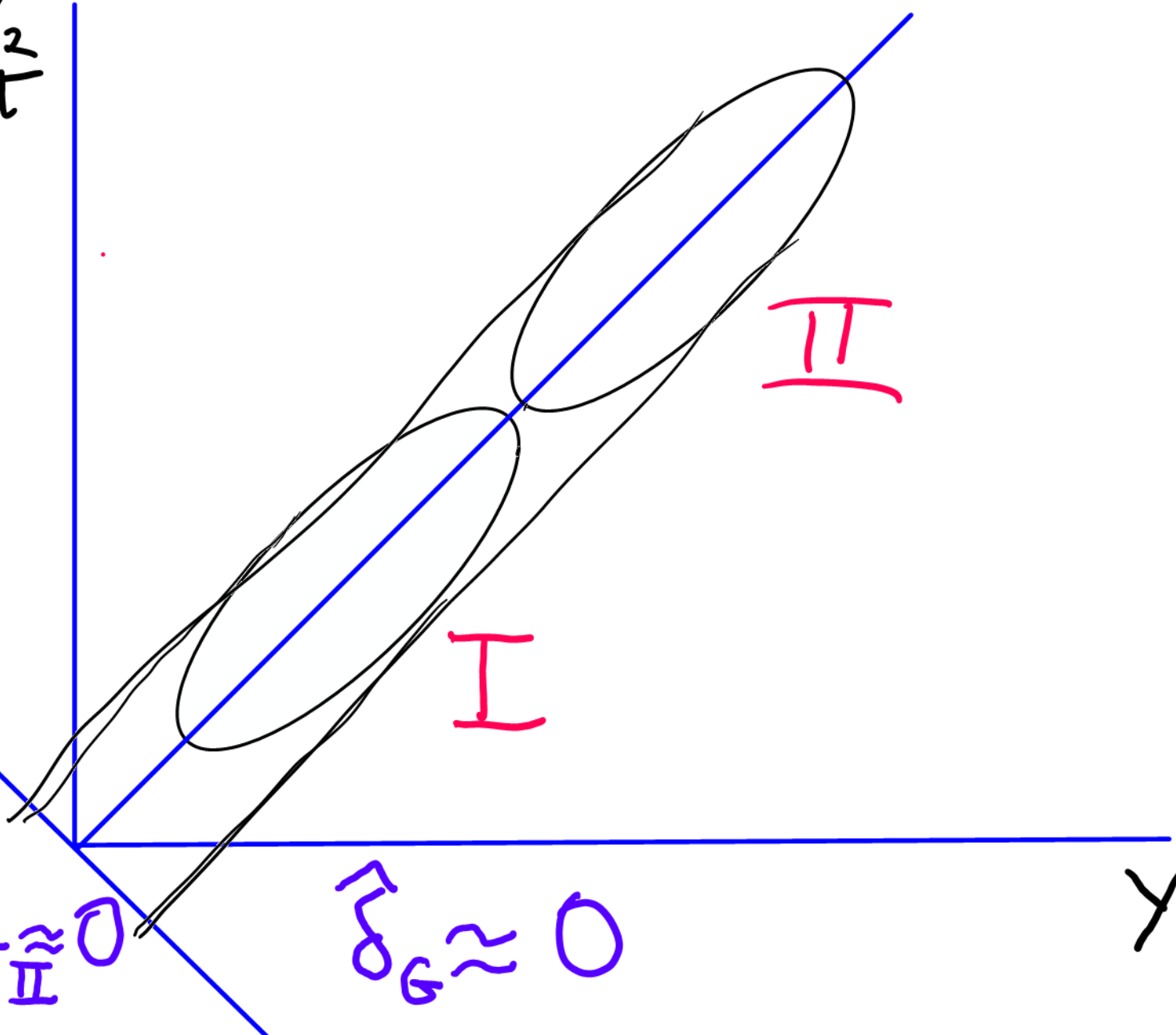
II

I

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

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

$y_1$  pre



$y_2$   
post

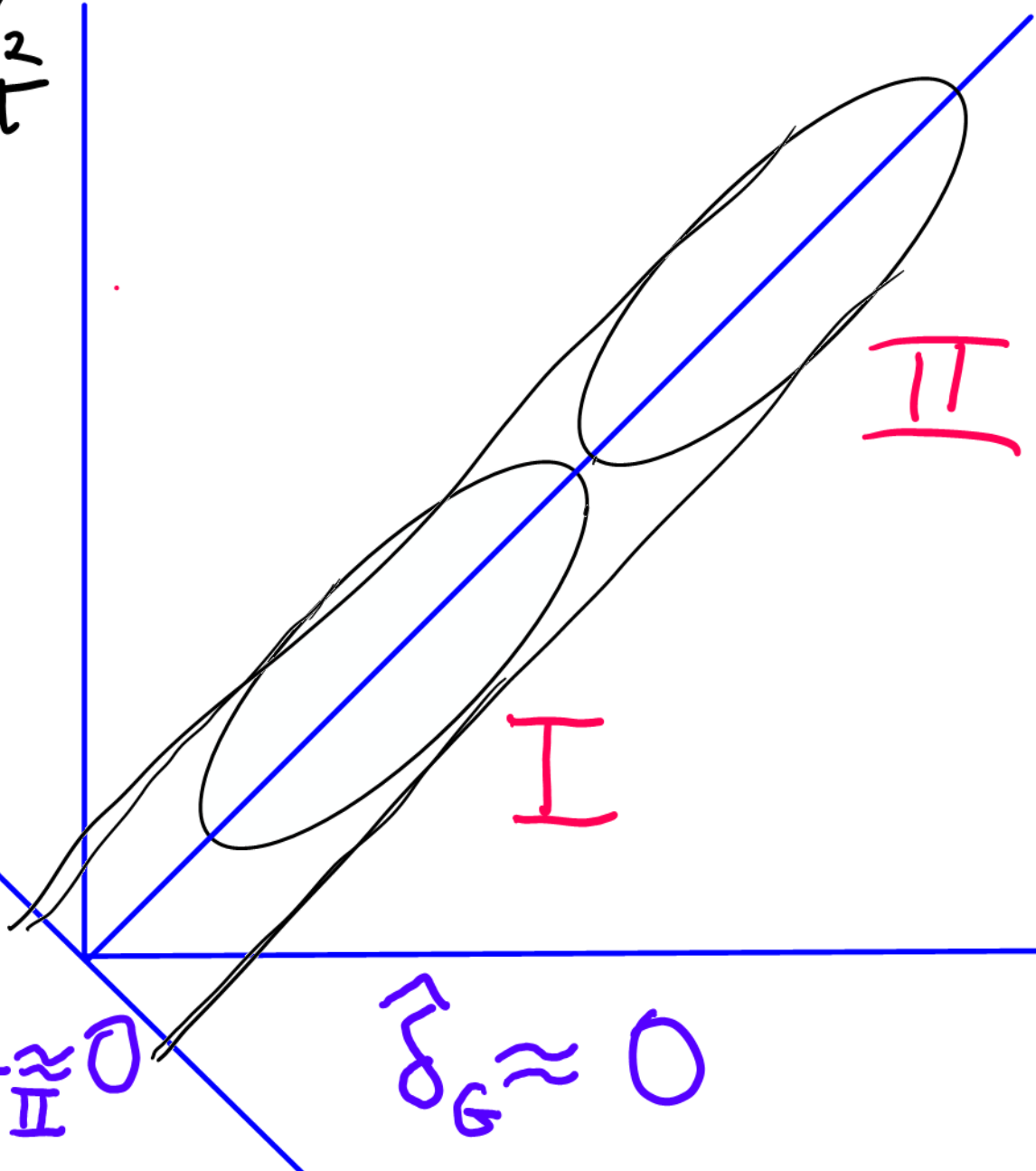
gain score

$$G = y_2 - y_1$$

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

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

$y_1$  pre



II

I

Longitudinal  
gain score  
provides  
a correct  
comparison

## 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 back doors - NOT mediators & colliders

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

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

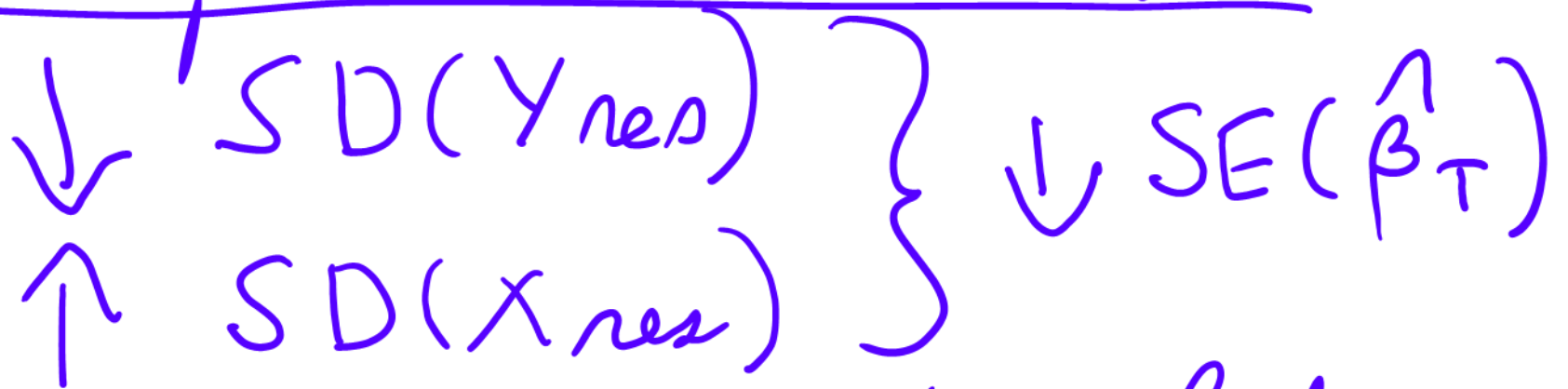
3) Honest SE

4) Robust Propensity scores - focus on X

Use the AVP to compare models.



Using confounders close to Y



But may not have knowledge about structure of model for Y

Using confounders close to X



But may have better understanding of assignment model.

Propensity score methods focus on predicting  $X$  with  $\hat{X}$

- no need to understand model for  $Y$

- except to avoid mediators & colliders

Then regress  $Y$  on  $X$  and  $\hat{X}$  (often grouped into intervals)

"Doubly robust:" throw in some  $Z$ 's close to  $Y$  and covariates.