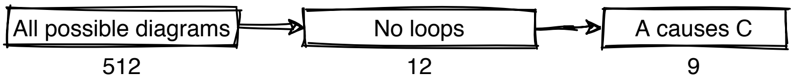
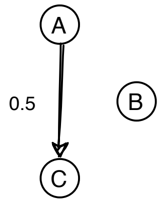


1



B not associated with A or C

=> **No need to condition;** makes little difference in practice.

Note sampling might induce A-B correlation which could slightly increase standard errors.

Legend

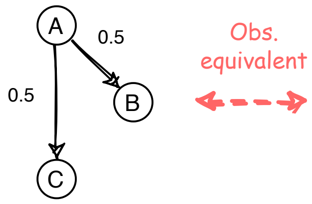
 Variable or quantity of interest

 Causal or functional link

 Observational equivalence

(Two diagrams generate the same observed distributions of data)

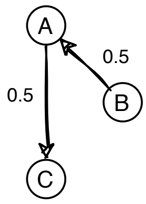
2



B not associated with C (no need to condition)
B correlated with A leading to increase in marginal standard error

=> **conditioning complicates interpretation**

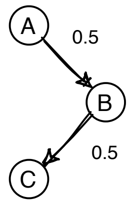
3



B not associated with C.
B associated with A => increased in marginal standard error =>

=> **conditioning complicates interpretation**

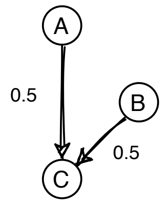
4



Causal pathway from A to C goes via B.

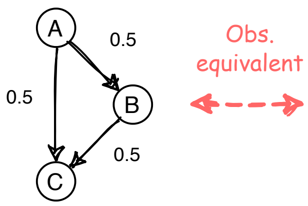
=> **conditioning destroys signal for A, loads on B instead.**

5



B not associated with A.
=> **no need to condition but does not hurt.**
If variables not fully gaussian, conditioning on B might help explain variation in C leading to better model fit.

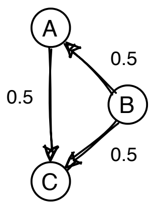
6



B associated with A and C. Two causal pathways A->C.
Not conditioning: get total effect from sum of both pathways.
Conditioning: estimate only direct effect.

=> **depends on aim.**

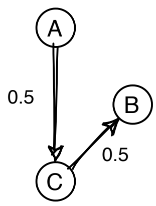
7



B is a **confounder**. c.f. also Simpson's paradox.
True effect of A on C is 0.5 but estimated as ~0.75 without conditioning

=> **mandatory to condition**

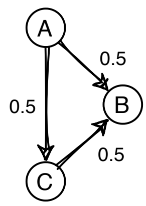
8



B is correlated with A through C (covariance=0.25)
Conditioning on B messes up estimate of A's effect

=> **must not condition**

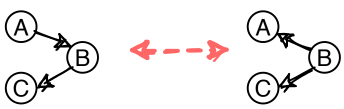
9



B is a **collider** (arrows meet at B).
Conditioning on B induces dependence between A and C

=> **must not condition**

See also: Simpson's paradox



See also: Berkson's paradox, collider bias