Let's consider a basic linear model with cross-sectional data:

Let's consider a basic linear model with cross-sectional data:

$$y = \alpha d + X\beta + \varepsilon \tag{1}$$

where d is a treatment indicator; X's are other observable covariates

Let's consider a basic linear model with cross-sectional data:

$$y = \alpha d + X\beta + \varepsilon \tag{1}$$

where d is a treatment indicator; X's are other observable covariates

Assume that $Corr(X, \varepsilon) = 0$ but that $Corr(d, \varepsilon) \neq 0$

How can $Corr(d, \varepsilon) \neq 0$? Rewrite (1):

How can $Corr(d, \varepsilon) \neq 0$? Rewrite (1):

$$y = \alpha d + X\beta + \underbrace{W\gamma + \nu}_{\varepsilon} \tag{2}$$

with $Corr(d, W) \neq 0$ but $Corr(X, W) = 0$; ν is purely idiosyncratic

How can $Corr(d, \varepsilon) \neq 0$? Rewrite (1):

$$y = \alpha d + X\beta + \underbrace{W\gamma + \nu}_{\varepsilon} \tag{2}$$

with $Corr(d, W) \neq 0$ but $Corr(X, W) = 0$; ν is purely idiosyncratic

W's are variables that are unobservable or excluded from the model

Some terminology:

Some terminology:

- $Corr(d, X\beta)$ is called selection on observables

Some terminology:

- $Corr(d, X\beta)$ is called selection on observables
- $Corr(d, \varepsilon)$ is called selection on unobservables

Some terminology:

- $Corr(d, X\beta)$ is called selection on observables
- $Corr(d, \varepsilon)$ is called selection on unobservables
- Selection on unobservables is what confounds our original regression model

Some terminology:

- $Corr(d, X\beta)$ is called selection on observables
- $Corr(d, \varepsilon)$ is called selection on unobservables
- Selection on unobservables is what confounds our original regression model
- Can always measure $Corr(d, X\beta)$ in the data; can never measure $Corr(d, \varepsilon)$

Some terminology:

- $Corr(d, X\beta)$ is called selection on observables
- $Corr(d, \varepsilon)$ is called selection on unobservables
- Selection on unobservables is what confounds our original regression model
- Can always measure $Corr(d, X\beta)$ in the data; can never measure $Corr(d, \varepsilon)$

Can we use selection on observables to tell us amount of selection on unobservables?

$$y = \alpha d + X\beta + \underbrace{W\gamma + \nu}_{\varepsilon} \tag{2}$$

Key question: Can we use $Corr(d, X\beta)$ to learn about the causal effect of d on y?

$$y = \alpha d + X\beta + \underbrace{W\gamma + \nu}_{\varepsilon} \tag{2}$$

Key question: Can we use $Corr(d, X\beta)$ to learn about the causal effect of d on y?

- One extreme: Assume unconfoundedness, i.e. $Corr(d, \varepsilon) = 0$

$$y = \alpha d + X\beta + \underbrace{W\gamma + \nu}_{\varepsilon} \tag{2}$$

Key question: Can we use $Corr(d, X\beta)$ to learn about the causal effect of d on y?

- One extreme: Assume unconfoundedness, i.e. $Corr(d, \varepsilon) = 0$
- Another extreme: Assume $Corr(d, \varepsilon) = Corr(d, X\beta)$, e.g. if $X \underset{\text{random}}{\subset} W$

$$y = \alpha d + X\beta + \underbrace{W\gamma + \nu}_{\varepsilon} \tag{2}$$

Key question: Can we use $Corr(d, X\beta)$ to learn about the causal effect of d on y?

- One extreme: Assume unconfoundedness, i.e. $Corr(d, \varepsilon) = 0$
- Another extreme: Assume $Corr(d, \varepsilon) = Corr(d, X\beta)$, e.g. if $X \underset{\text{random}}{\subset} W$
- Combine $Corr(d, X\beta)$ with assumption about $Corr(d, \varepsilon)$ to properly "shade" $\hat{\alpha}$

We can follow these steps to obtain a plausibly causal $\hat{\alpha}$:

We can follow these steps to obtain a plausibly causal $\hat{\alpha}$:

1. Regress y on d

We can follow these steps to obtain a plausibly causal $\hat{\alpha}$:

1. Regress y on d

2. Regress y on d and X

We can follow these steps to obtain a plausibly causal $\hat{\alpha}$:

1. Regress y on d

2. Regress y on d and X

3. Observe how $\hat{\alpha}$ and R^2 change across steps 1 and 2

We can follow these steps to obtain a plausibly causal $\hat{\alpha}$:

1. Regress y on d
2. Regress y on d and X
3. Observe how $\hat{\alpha}$ and R^2 change across steps 1 and 2
4. Make an assumption on λ such that $Corr(d, \varepsilon) = \lambda Corr(d, X\beta)$

We can follow these steps to obtain a plausibly causal $\hat{\alpha}$:

1. Regress y on d

2. Regress y on d and X

3. Observe how $\hat{\alpha}$ and R^2 change across steps 1 and 2

4. Make an assumption on λ such that $Corr(d, \varepsilon) = \lambda Corr(d, X\beta)$

5. Use all of the above information to get an interval on plausible $\hat{\alpha}$ TE