Lecture 09:
Instrumental variables II

PPHA 34600
Prof. Fiona Burlig

Harris School of Public Policy
University of Chicago

From last time: introduction to IV

Recall that we want to split $D_i = B_i \varepsilon_i + C_i$ into the C_i and other parts

An instrumental variable...:

...Generates variation in C_i but is uncorrelated with ε_i

Z_i is a valid instrument for D_i when the following are satisfied:

1. **First stage**: $Cov(Z_i, D_i) \neq 0$
    - Z_i and D_i are related
    - Without this, you're capturing nothing
    - This is actually testable!
2. **Exclusion restriction**: $Cov(Z_i, \varepsilon_i) = 0$
    - Z_i and ε_i are **not** related
    - Z_i only affects Y_i through D_i
    - Fundamentally untestable!

What makes IV so useful?

IV can be used in many ways:

- Causal inference (see last time)
- (Omitted variable bias)
- Measurement error

An omitted variable bias refresher

Suppose the true data generating process is:

$$Y_i = \alpha + \tau D_i + \beta X_i + \varepsilon_i$$

We'll assume:

- D_i and X_i are uncorrelated with ε_i
- D_i and X_i are correlated with each other
  - $\rightarrow$ $Cov(D_i, X_i) \neq 0$
- We don't observe X_i ()
  - $\rightarrow$ Now we have to run:

    $$Y_i = \alpha + \tau D_i + \nu_i$$

    where

    $$\nu_i = \varepsilon_i + \beta X_i$$

An omitted variable bias refresher

$$\hat{\tau} = \frac{Cov(Y_i, D_i)}{Var(D_i)}$$

$$= \underbrace{\frac{Cov(\alpha + \tau D_i + \beta X_i + \varepsilon_i, D_i)}{Var(D_i)}}_{\text{plug in for } Y_i}$$

An omitted variable bias refresher

$$\hat{\tau} = \frac{Cov(Y_i, D_i)}{Var(D_i)}$$

$$= \underbrace{\frac{Cov(\alpha + \tau D_i + \beta X_i + \varepsilon_i, D_i)}{Var(D_i)}}_{\text{plug in for } Y_i}$$

$$= \underbrace{\frac{Cov(\alpha, D_i,) + \tau Cov(D_i, D_i) + \beta Cov(D_i, X_i) + Cov(\varepsilon_i, D_i)}{Var(D_i)}}_{\text{laws of } Cov()}$$

An omitted variable bias refresher

$$\hat{\tau} = \frac{Cov(Y_i, D_i)}{Var(D_i)}$$

$$= \underbrace{\frac{Cov(\alpha + \tau D_i + \beta X_i + \varepsilon_i, D_i)}{Var(D_i)}}_{\text{plug in for } Y_i}$$

$$= \underbrace{\frac{Cov(\alpha, D_i,) + \tau Cov(D_i, D_i) + \beta Cov(D_i, X_i) + Cov(\varepsilon_i, D_i)}{Var(D_i)}}_{\text{laws of } Cov()}$$

$$= \underbrace{\frac{0 + \tau Var(D_i) + \beta Cov(D_i, X_i) + 0}{Var(D_i)}}_{\text{definitions}}$$

An omitted variable bias refresher

$$
\begin{aligned}
\hat{\tau} &= \frac{Cov(Y_i, D_i)}{Var(D_i)} \\
&= \underbrace{\frac{Cov(\alpha + \tau D_i + \beta X_i + \varepsilon_i, D_i)}{Var(D_i)}}_{\text{plug in for } Y_i} \\
&= \underbrace{\frac{Cov(\alpha, D_i,) + \tau Cov(D_i, D_i) + \beta Cov(D_i, X_i) + Cov(\varepsilon_i, D_i)}{Var(D_i)}}_{\text{laws of } Cov()} \\
&= \underbrace{\frac{0 + \tau Var(D_i) + \beta Cov(D_i, X_i) + 0}{Var(D_i)}}_{\text{definitions}} \\
&= \underbrace{\tau + \beta \frac{Cov(D_i, X_i)}{Var(D_i)}}_{\text{simplify}}
\end{aligned}
$$

An omitted variable bias refresher

With omitted variable bias, we instead have

$$\hat{\tau} = \tau + \beta \frac{Cov(D_i, X_i)}{Var(D_i)} \neq \tau$$

An omitted variable bias refresher

With omitted variable bias, we instead have

$$\hat{\tau} = \tau + \beta \frac{Cov(D_i, X_i)}{Var(D_i)} \neq \tau$$

Let's take a trip to the instrument store:

- Suppose we have an instrument, Z_i
- Our instrument, Z_i, moves D_i, but is uncorrelated with the error term

An omitted variable bias refresher

With omitted variable bias, we instead have

$$\hat{\tau} = \tau + \beta \frac{Cov(D_i, X_i)}{Var(D_i)} \neq \tau$$

Let's take a trip to the instrument store:

- Suppose we have an instrument, Z_i
- Our instrument, Z_i, moves D_i, but is uncorrelated with the error term
- In this case, the error term is $\nu_i = \varepsilon_i + \beta X_i$
  - → In other words, $Cov(Z_i, D_i) \neq 0$ and $Cov(Z_i, \nu_i) = 0$

An omitted variable bias refresher

With omitted variable bias, we instead have

$$\hat{\tau} = \tau + \beta \frac{Cov(D_i, X_i)}{Var(D_i)} \neq \tau$$

Let's take a trip to the instrument store:

- Suppose we have an instrument, Z_i
- Our instrument, Z_i, moves D_i, but is uncorrelated with the error term
- In this case, the error term is $\nu_i = \varepsilon_i + \beta X_i$
  - $\rightarrow$ In other words, $Cov(Z_i, D_i) \neq 0$ and $Cov(Z_i, \nu_i) = 0$
- The second-stage IV estimate is then:

$$\hat{\tau}^{2SLS} = \tau + \beta \frac{Cov(\hat{D}_i, X_i)}{Var(\hat{D}_i)}$$

An omitted variable bias refresher

With omitted variable bias, we instead have

$$\hat{\tau} = \tau + \beta \frac{Cov(D_i, X_i)}{Var(D_i)} \neq \tau$$

<u>Let's take a trip to the instrument store:</u>

- Suppose we have an instrument, Z_i
- Our instrument, Z_i, moves D_i, but is uncorrelated with the error term
- In this case, the error term is $\nu_i = \varepsilon_i + \beta X_i$
  - → In other words, $Cov(Z_i, D_i) \neq 0$ and $Cov(Z_i, \nu_i) = 0$
- The second-stage IV estimate is then:

$$\hat{\tau}^{2SLS} = \tau + \beta \frac{Cov(\hat{D}_i, X_i)}{Var(\hat{D}_i)}$$

$$\hat{\tau}^{2SLS} = \tau + \underbrace{0}_{\text{exclusion restriction}}$$

Measurement error

We often worry about measurement error:

- What happens if we don't perfectly observe D_i or Y_i?
- This is extremely common!

Measurement error

We often worry about measurement error:

- What happens if we don't perfectly observe D_i or Y_i?
- This is extremely common!
- The answer is...

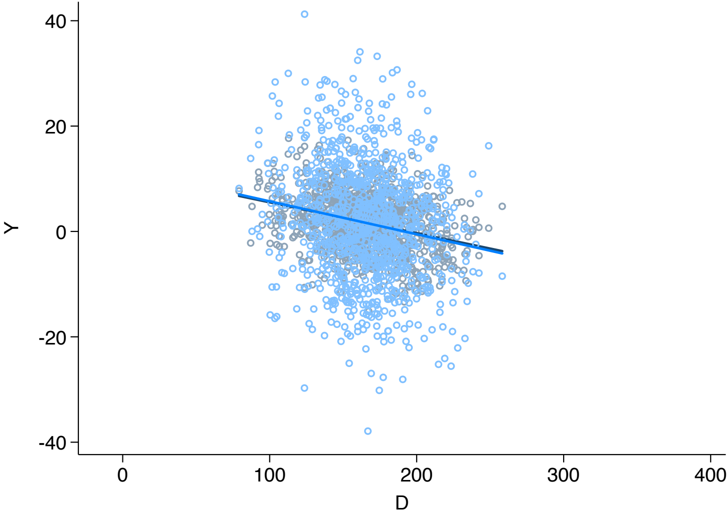
Measurement error

We often worry about measurement error:

- What happens if we don't perfectly observe D_i or Y_i?
- This is extremely common!
- The answer is...
- It depends!

Consider a true relationship

Measurement error in Y is fine



Estimated relationship: $\hat{\tau} = -0.061$

Why does this work?

We don't observe Y_i, but rather $\tilde{Y}_i = Y_i + \gamma_i$

$\rightarrow$ Assume $Cov(\gamma_i, \varepsilon_i) = 0$ and $Cov(\gamma_i, D_i) = 0$

If we run:

$$\tilde{Y}_i = \alpha + \tau D_i + \varepsilon_i$$

We'll estimate:

$$\hat{\tau} = \underbrace{\frac{Cov(\tilde{Y}_i, D_i)}{Var(D_i)}}_{\text{def'n of OLS}}$$

Why does this work?

We don't observe Y_i, but rather $\tilde{Y}_i = Y_i + \gamma_i$

$\rightarrow$ Assume $Cov(\gamma_i, \varepsilon_i) = 0$ and $Cov(\gamma_i, D_i) = 0$

If we run:

$$\tilde{Y}_i = \alpha + \tau D_i + \varepsilon_i$$

We'll estimate:

$$\hat{\tau} = \underbrace{\frac{Cov(\tilde{Y}_i, D_i)}{Var(D_i)}}_{\text{def'n of OLS}}$$

$$= \underbrace{\frac{Cov(Y_i + \gamma_i, D_i)}{Var(D_i)}}_{\text{def'n of } \tilde{D}_i}$$

Why does this work?

We don't observe Y_i, but rather $\tilde{Y}_i = Y_i + \gamma_i$

$\rightarrow$ Assume $Cov(\gamma_i, \varepsilon_i) = 0$ and $Cov(\gamma_i, D_i) = 0$

If we run:

$$\tilde{Y}_i = \alpha + \tau D_i + \varepsilon_i$$

We'll estimate:

$$\hat{\tau} = \underbrace{\frac{Cov(\tilde{Y}_i, D_i)}{Var(D_i)}}_{\text{def'n of OLS}}$$

$$= \underbrace{\frac{Cov(Y_i + \gamma_i, D_i)}{Var(D_i)}}_{\text{def'n of } \tilde{D}_i}$$

$$= \underbrace{\frac{Cov(\alpha + \tau D_i + \varepsilon_i + \gamma_i, D_i)}{Var(D_i)}}_{\text{def'n of } Y_i}$$

Why does this work?

$$\hat{\tau} = \frac{Cov(\alpha + \tau D_i + \varepsilon_i + \gamma_i, D_i)}{Var(D_i)}$$

Why does this work?

$$\hat{\tau} = \frac{Cov(\alpha + \tau D_i + \varepsilon_i + \gamma_i, D_i)}{Var(D_i)}$$

$$= \underbrace{\frac{Cov(\alpha, D_i) + \tau Cov(D_i, D_i) + Cov(\varepsilon_i, D_i) + Cov(\gamma_i, D_i)}{Var(D_i)}}_{\text{covariance rules}}$$

Why does this work?

$$\hat{\tau} = \frac{Cov(\alpha + \tau D_i + \varepsilon_i + \gamma_i, D_i)}{Var(D_i)}$$

$$= \underbrace{\frac{Cov(\alpha, D_i) + \tau Cov(D_i, D_i) + Cov(\varepsilon_i, D_i) + Cov(\gamma_i, D_i)}{Var(D_i)}}_{\text{covariance rules}}$$

$$= \underbrace{\frac{\tau Cov(D_i, D_i)}{Var(D_i)}}_{\text{assumptions}}$$

Why does this work?

$$\hat{\tau} = \frac{Cov(\alpha + \tau D_i + \varepsilon_i + \gamma_i, D_i)}{Var(D_i)}$$

$$= \underbrace{\frac{Cov(\alpha, D_i) + \tau Cov(D_i, D_i) + Cov(\varepsilon_i, D_i) + Cov(\gamma_i, D_i)}{Var(D_i)}}_{\text{covariance rules}}$$

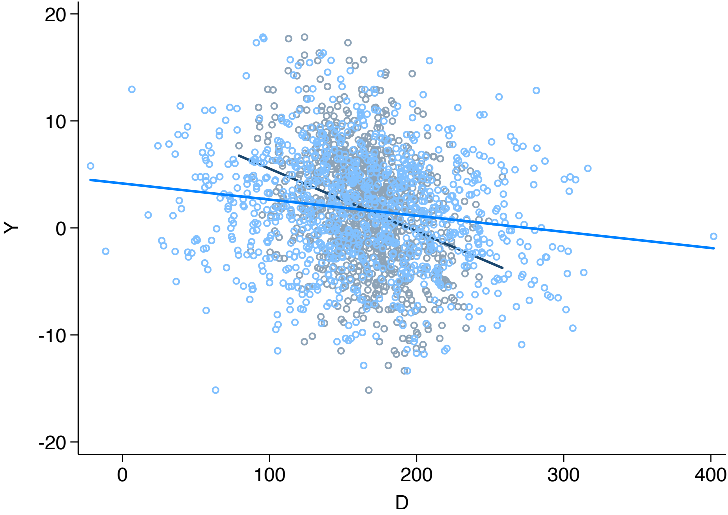
$$= \underbrace{\frac{\tau Cov(D_i, D_i)}{Var(D_i)}}_{\text{assumptions}}$$

$$= \tau$$

Success!

Classical measurement error in D_i is bad

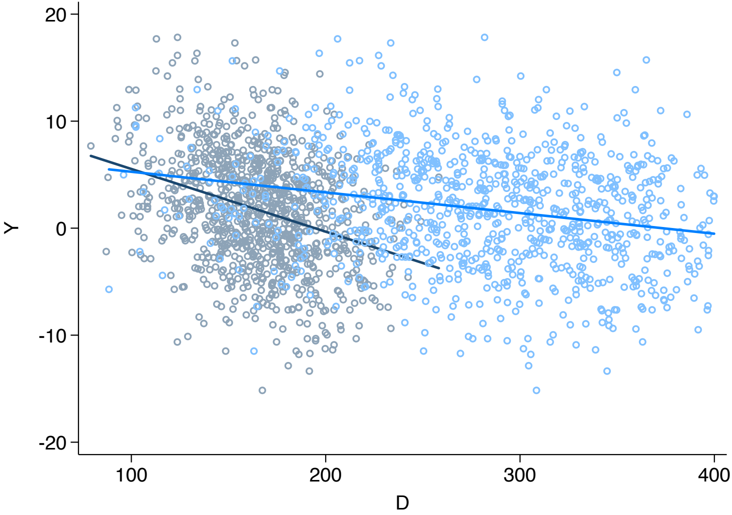
Classical measurement error in D_i is bad



Estimated relationship: $\hat{\tau} = -0.015$

Non-classical measurement error in D_i is bad

Non-classical measurement error in D_i is bad



Estimated relationship: $\hat{\tau} = -0.019$

What's going wrong?

We want to estimate:

$$Y_i = \alpha + \tau D_i + \varepsilon_i$$

What's going wrong?

We want to estimate:

$$Y_i = \alpha + \tau D_i + \varepsilon_i$$

Enter measurement error:

We don't observe D_i, but rather

$$\tilde{D}_i = D_i + \gamma_i$$

Assume:

- $Cov(D_i, \varepsilon_i) = 0$: Treatment is (as good as) random
- $Cov(\gamma_i, D_i) = 0$: Measurement error is uncorrelated with treatment
- $Cov(\gamma_i, \varepsilon_i) = 0$: Measurement error is not in our original error term

What's going wrong?

If we run:

$$Y_i = \alpha + \tau \tilde{D}_i + \varepsilon_i$$

We'll estimate:

$$\hat{\tau} = \underbrace{\frac{Cov(Y_i, \tilde{D}_i)}{Var(\tilde{D}_i)}}_{\text{def'n of OLS}}$$

What's going wrong?

If we run:

$$Y_i = \alpha + \tau \tilde{D}_i + \varepsilon_i$$

We'll estimate:

$$\hat{\tau} = \underbrace{\frac{Cov(Y_i, \tilde{D}_i)}{Var(\tilde{D}_i)}}_{\text{def'n of OLS}}$$

$$= \underbrace{\frac{Cov(Y_i, D_i + \gamma_i)}{Var(D_i + \gamma_i)}}_{\text{def'n of } \tilde{D}_i}$$

What's going wrong?

If we run:

$$Y_i = \alpha + \tau \tilde{D}_i + \varepsilon_i$$

We'll estimate:

$$\hat{\tau} = \underbrace{\frac{Cov(Y_i, \tilde{D}_i)}{Var(\tilde{D}_i)}}_{\text{def'n of OLS}}$$

$$= \underbrace{\frac{Cov(Y_i, D_i + \gamma_i)}{Var(D_i + \gamma_i)}}_{\text{def'n of } \tilde{D}_i}$$

$$= \underbrace{\frac{Cov(\alpha + \tau D_i + \varepsilon_i, D_i + \gamma_i)}{Var(D_i + \gamma_i)}}_{\text{def'n of } Y_i}$$

What's going wrong?

$$\hat{\tau} = \frac{Cov(\alpha + \tau D_i + \varepsilon_i, D_i + \gamma_i)}{Var(D_i + \gamma_i)}$$

What's going wrong?

$$\hat{\tau} = \frac{Cov(\alpha + \tau D_i + \varepsilon_i, D_i + \gamma_i)}{Var(D_i + \gamma_i)}$$

$$= \underbrace{\frac{Cov(\alpha + \tau D_i + \varepsilon_i, D_i + \gamma_i)}{Var(D_i) + Var(\gamma_i)}}_{\text{variance rules}}$$

What's going wrong?

$$\hat{\tau} = \frac{Cov(\alpha + \tau D_i + \varepsilon_i, D_i + \gamma_i)}{Var(D_i + \gamma_i)}$$

$$= \underbrace{\frac{Cov(\alpha + \tau D_i + \varepsilon_i, D_i + \gamma_i)}{Var(D_i) + Var(\gamma_i)}}_{\text{variance rules}}$$

$$= \underbrace{\frac{Cov(\alpha, D_i) + Cov(\alpha, \gamma_i) + \tau Cov(D_i, D_i)}{Var(D_i) + Var(\gamma_i)} + \frac{\tau Cov(D_i, \gamma_i) + Cov(D_i, \varepsilon_i) + Cov(\gamma_i, \varepsilon_i)}{Var(D_i) + Var(\gamma_i)}}_{\text{covariance rules}}$$

What's going wrong?

$$\hat{\tau} = \frac{Cov(\alpha + \tau D_i + \varepsilon_i, D_i + \gamma_i)}{Var(D_i + \gamma_i)}$$

$$= \underbrace{\frac{Cov(\alpha + \tau D_i + \varepsilon_i, D_i + \gamma_i)}{Var(D_i) + Var(\gamma_i)}}_{\text{variance rules}}$$

$$= \frac{Cov(\alpha, D_i) + Cov(\alpha, \gamma_i) + \tau Cov(D_i, D_i)}{Var(D_i) + Var(\gamma_i)}$$

$$+ \underbrace{\frac{\tau Cov(D_i, \gamma_i) + Cov(D_i, \varepsilon_i) + Cov(\gamma_i, \varepsilon_i)}{Var(D_i) + Var(\gamma_i)}}_{\text{covariance rules}}$$

$$= \frac{\tau Var(D_i)}{\underbrace{Var(D_i) + Var(\gamma_i)}_{\text{assumptions}}}$$

What's going wrong?

$$\hat{\tau} = \tau \left(\frac{Var(D_i)}{Var(D_i) + Var(\gamma_i)} \right)$$

This is the classic **attenuation bias**

- $\hat{\tau}$ is biased towards zero
- Note we assumed the most innocuous form of measurement error
- If measurement error is correlated with treatment, we get OVB

A second trip to the instrument store

To solve the measurement error problem, we'll use a clever instrument:

- We will instrument for $\tilde{D}_i$ with Z_i, a different noisy measure of D_i:

$$Z_i \equiv \mathring{D}_i = D_i + \zeta_i$$

Assume:

- $Cov(\zeta_i, D_i) = 0$: Measurement error is uncorrelated with treatment
- $Cov(\zeta_i, \gamma_i) = 0$: Measurement error in Z_i is uncorrelated w error in $\tilde{D}_i$
- $Cov(\zeta_i, \varepsilon_i) = 0$: Measurement error is uncorrelated with original error

Does this meet our two assumptions?

1. **First stage:** Yes! $Cov(Z_i, \tilde{D}_i) \neq 0$
2. **Exclusion restriction:** Yes! $Cov(Z_i, \varepsilon_i) = 0$

What do these assumptions buy us?

Remember that:

$$\hat{\tau}^{IV} = \frac{Cov(Y_i, Z_i)}{Cov(\tilde{D}_i, Z_i)}$$

What do these assumptions buy us?

Remember that:

$$\hat{\tau}^{IV} = \frac{Cov(Y_i, Z_i)}{Cov(\tilde{D}_i, Z_i)}$$

$$= \underbrace{\frac{Cov(\tau D_i + \varepsilon_i, Z_i)}{Cov(D_i + \gamma_i, Z_i)}}_{\text{definition of } Y_i,\ \tilde{D}_i}$$

What do these assumptions buy us?

Remember that:

$$\hat{\tau}^{IV} = \frac{Cov(Y_i, Z_i)}{Cov(\tilde{D}_i, Z_i)}$$

$$= \underbrace{\frac{Cov(\tau D_i + \varepsilon_i, Z_i)}{Cov(D_i + \gamma_i, Z_i)}}_{\text{definition of } Y_i,\ \tilde{D}_i}$$

$$= \underbrace{\frac{\tau Cov(D_i, Z_i) + Cov(\varepsilon_i, Z_i)}{Cov(D_i, Z_i) + Cov(\gamma_i, Z_i)}}_{\text{variance rules}}$$

What do these assumptions buy us?

Remember that:

$$\hat{\tau}^{IV} = \frac{Cov(Y_i, Z_i)}{Cov(\tilde{D}_i, Z_i)}$$

$$= \underbrace{\frac{Cov(\tau D_i + \varepsilon_i, Z_i)}{Cov(D_i + \gamma_i, Z_i)}}_{\text{definition of } Y_i,\ \tilde{D}_i}$$

$$= \underbrace{\frac{\tau Cov(D_i, Z_i) + Cov(\varepsilon_i, Z_i)}{Cov(D_i, Z_i) + Cov(\gamma_i, Z_i)}}_{\text{variance rules}}$$

$$= \tau \underbrace{\left(\frac{Cov(D_i, Z_i)}{Cov(D_i, Z_i)}\right)}_{\text{assumptions}}$$

What do these assumptions buy us?

Remember that:

$$\hat{\tau}^{IV} = \frac{Cov(Y_i, Z_i)}{Cov(\tilde{D}_i, Z_i)}$$

$$= \underbrace{\frac{Cov(\tau D_i + \varepsilon_i, Z_i)}{Cov(D_i + \gamma_i, Z_i)}}_{\text{definition of } Y_i,\ \tilde{D}_i}$$

$$= \underbrace{\frac{\tau Cov(D_i, Z_i) + Cov(\varepsilon_i, Z_i)}{Cov(D_i, Z_i) + Cov(\gamma_i, Z_i)}}_{\text{variance rules}}$$

$$= \underbrace{\tau \left(\frac{Cov(D_i, Z_i)}{Cov(D_i, Z_i)} \right)}_{\text{assumptions}}$$

$$= \tau$$

Success!

IV solves measurement error

What's the intuition?

- $\tilde{D}_i = D_i + \gamma_i$
- $Z_i = \mathring{D}_i = D_i + \zeta_i$

$\rightarrow$ Z_i and $\tilde{D}_i$ only have the **true** D_i in common

$\rightarrow$ We've assumed that $Cov(\gamma_i, \zeta_i) = 0$

IV solves measurement error

What's the intuition?

- $\tilde{D}_i = D_i + \gamma_i$
- $Z_i = \mathring{D}_i = D_i + \zeta_i$

→ Z_i and $\tilde{D}_i$ only have the **true** D_i in common

→ We've assumed that $Cov(\gamma_i, \zeta_i) = 0$

The first stage is:

$$\tilde{D}_i = \alpha + \pi \mathring{D}_i + \epsilon_i$$

→ We're only using the variation from D_i (not from ζ_i or γ_i)!

IV solves measurement error

What's the intuition?

- $\tilde{D}_i = D_i + \gamma_i$
- $Z_i = \mathring{D}_i = D_i + \zeta_i$

→ Z_i and $\tilde{D}_i$ only have the **true** D_i in common

→ We've assumed that $Cov(\gamma_i, \zeta_i) = 0$

The first stage is:

$$\tilde{D}_i = \alpha + \pi \mathring{D}_i + \epsilon_i$$

→ We're only using the variation from D_i (not from ζ_i or γ_i)!

→ **Important caveat:** This does not work with binary D_i!

- If true $D_i = 1$, measurement error can only be -1 or 0
- If true $D_i = 0$, measurement error can only be 0 or 1

→ Measurement error in $\tilde{D}_i$ and $\mathring{D}_i$ will be correlated

Non-classical measurement error

We want to estimate:

$$Y_i = \alpha + \tau D_i + \varepsilon_i$$

Non-classical measurement error

We want to estimate:

$$Y_i = \alpha + \tau D_i + \varepsilon_i$$

<u>Enter measurement error:</u>

We don't observe D_i, but rather

$$\tilde{D}_i = D_i + \gamma_i$$

Non-classical measurement error

We want to estimate:

$$Y_i = \alpha + \tau D_i + \varepsilon_i$$

Enter measurement error:

We don't observe D_i, but rather

$$\tilde{D}_i = D_i + \gamma_i$$

Assume:

- $Cov(D_i, \varepsilon_i) = 0$: Treatment is (as good as) random
- $Cov(\gamma_i, \varepsilon_i) = 0$: Measurement error is not in our original error term

Relax the orthogonality assumption:

- Allow $Cov(D_i, \gamma_i) \neq 0$: Measurement error can be correlated with treatment

IV with non-classical measurement error

Again, we'll now have:

$$\hat{\tau} = \frac{Cov(Y_i, D_i + \gamma_i)}{Var(D_i + \gamma_i)}$$

IV with non-classical measurement error

Again, we'll now have:

$$\hat{\tau} = \frac{Cov(Y_i, D_i + \gamma_i)}{Var(D_i + \gamma_i)}$$

$$= \underbrace{\frac{Cov(\alpha + \tau D_i + \varepsilon_i, D_i + \gamma_i)}{Var(D_i + \gamma_i)}}_{\text{def'n of } Y_i, \tilde{D}_i}$$

IV with non-classical measurement error

Again, we'll now have:

$$\hat{\tau} = \frac{Cov(Y_i, D_i + \gamma_i)}{Var(D_i + \gamma_i)}$$

$$= \underbrace{\frac{Cov(\alpha + \tau D_i + \varepsilon_i, D_i + \gamma_i)}{Var(D_i + \gamma_i)}}_{\text{def'n of } Y_i, \tilde{D}_i}$$

$$= \underbrace{\frac{Cov(\alpha + \tau D_i + \varepsilon_i, D_i + \gamma_i)}{Var(D_i) + Var(\gamma_i) + 2Cov(D_i, \gamma_i)}}_{\text{variance rules}}$$

IV with non-classical measurement error

Again, we'll now have:

$$\hat{\tau} = \frac{Cov(Y_i, D_i + \gamma_i)}{Var(D_i + \gamma_i)}$$

$$= \underbrace{\frac{Cov(\alpha + \tau D_i + \varepsilon_i, D_i + \gamma_i)}{Var(D_i + \gamma_i)}}_{\text{def'n of } Y_i, \tilde{D}_i}$$

$$= \underbrace{\frac{Cov(\alpha + \tau D_i + \varepsilon_i, D_i + \gamma_i)}{Var(D_i) + Var(\gamma_i) + 2Cov(D_i, \gamma_i)}}_{\text{variance rules}}$$

$$= \underbrace{\frac{Cov(\alpha, D_i) + Cov(\alpha, \gamma_i) + \tau Cov(D_i, D_i)}{Var(D_i) + Var(\gamma_i) + 2Cov(D_i, \gamma_i)} + \frac{\tau Cov(D_i, \gamma_i) + Cov(D_i, \varepsilon_i) + Cov(\gamma_i, \varepsilon_i)}{Var(D_i) + Var(\gamma_i) + 2Cov(D_i, \gamma_i)}}_{\text{covariance rules}}$$

IV with non-classical measurement error

$$\hat{\tau} = \frac{Cov(\alpha, D_i) + Cov(\alpha, \gamma_i) + \tau Cov(D_i, D_i)}{Var(D_i) + Var(\gamma_i) + 2Cov(D_i, \gamma_i)}$$

$$+ \frac{\tau Cov(D_i, \gamma_i) + Cov(D_i, \varepsilon_i) + Cov(\gamma_i, \varepsilon_i)}{Var(D_i) + Var(\gamma_i) + 2Cov(D_i, \gamma_i)}$$

IV with non-classical measurement error

$$\hat{\tau} = \frac{Cov(\alpha, D_i) + Cov(\alpha, \gamma_i) + \tau Cov(D_i, D_i)}{Var(D_i) + Var(\gamma_i) + 2Cov(D_i, \gamma_i)}$$

$$+ \frac{\tau Cov(D_i, \gamma_i) + Cov(D_i, \varepsilon_i) + Cov(\gamma_i, \varepsilon_i)}{Var(D_i) + Var(\gamma_i) + 2Cov(D_i, \gamma_i)}$$

$$= \underbrace{\tau \left(\frac{Var(D_i) + Cov(D_i, \gamma_i)}{Var(D_i) + Var(\gamma_i) + 2Cov(D_i, \gamma_i)} \right)}_{\text{rearrange}}$$

IV with non-classical measurement error

$$\hat{\tau} = \frac{Cov(\alpha, D_i) + Cov(\alpha, \gamma_i) + \tau Cov(D_i, D_i)}{Var(D_i) + Var(\gamma_i) + 2Cov(D_i, \gamma_i)}$$

$$+ \frac{\tau Cov(D_i, \gamma_i) + Cov(D_i, \varepsilon_i) + Cov(\gamma_i, \varepsilon_i)}{Var(D_i) + Var(\gamma_i) + 2Cov(D_i, \gamma_i)}$$

$$= \underbrace{\tau \left(\frac{Var(D_i) + Cov(D_i, \gamma_i)}{Var(D_i) + Var(\gamma_i) + 2Cov(D_i, \gamma_i)} \right)}_{\text{rearrange}}$$

→ Again, we get **bias**

→ Note that this **need not attenuate** $\hat{\tau}$

→ This can actually **flip the sign** of $\hat{\tau}$ relative to τ

→ (This depends on the sign of $Cov(D_i, \gamma_i)$)

What about IV?

Just like before...:

$$\hat{\tau}^{IV} = \frac{Cov(Y_i, Z_i)}{Cov(\tilde{D}_i, Z_i)}$$

$$= \underbrace{\frac{Cov(\tau D_i + \varepsilon_i, Z_i)}{Cov(D_i + \gamma_i, Z_i)}}_{\text{definition of } Y_i,\ \tilde{D}_i}$$

$$= \underbrace{\frac{\tau Cov(D_i, Z_i) + Cov(\varepsilon_i, Z_i)}{Cov(D_i, Z_i) + Cov(\gamma_i, Z_i)}}_{\text{variance rules}}$$

$$= \underbrace{\tau \left(\frac{Cov(D_i, Z_i)}{Cov(D_i, Z_i)} \right)}_{\text{assumptions}}$$

$$= \tau$$

Success!

An example: Early-life rainfall and health

Policy issue:

- Early-life shocks may be very important
- With bad harvests, kids may not get the proper nutrition

Approach:

- (We're not actually evaluating a program here)
- We want to estimate the effect of rainfall on health
- **Measurement of rainfall is poor in Indonesia**
- Instrument of choice: rainfall at weather stations $j \neq i$

Estimating the effects of rainfall on health

The authors will run a (simplified) version of:

$$Y_i = \tau Rainfall_i + \varepsilon_i$$

Where:

Y_i is a health outcome of interest

$Rainfall_i$ is rain in location i

- (They'll actually do this in a series of lags)

ε_i is an error term

Estimating the effects of rainfall on health

The authors will run a (simplified) version of:

$$Y_i = \tau Rainfall_i + \varepsilon_i$$

Where:

Y_i is a health outcome of interest

$Rainfall_i$ is rain in location i

- (They'll actually do this in a series of lags)

ε_i is an error term

A big concern

- $Rainfall_i$ is measured with error
- We are likely to understate the true effect
- **Solution:** $Z_i = Rainfall\ Nearby_i$!

First stage estimates

Dependent variable: Rainfall in birthyear and birthdistrict (deviation of log rainfall in birth district from log of 1953-1999 district mean rainfall)

	Women	Men
Birthyear/birthdistrict rainfall, 2nd-closest station	0.138 (0.024)***	0.120 (0.023)***
Birthyear/birthdistrict rainfall, 3rd-closest station	0.144 (0.039)***	0.158 (0.035)***
Birthyear/birthdistrict rainfall, 4th-closest station	0.088 (0.053)	0.081 (0.044)*
Birthyear/birthdistrict rainfall, 5th-closest station	0.125 (0.025)***	0.158 (0.039)***
Number of observations	4,615	4,277
R-squared	0.59	0.59
F-statistic: Joint significance of all four rainfall variables	31.61	28.80
P-value	0.000	0.000

Placebo test estimates

Coefficients (std. errors) in regression of outcome on *child's* birthyear rainfall.

	Women	Men
Mother's characteristics		
Completed grades of schooling	0.204	0.132
	(1.136)	(0.947)
	[2,447]	[2,258]
Currently alive (indicator)	0.084	0.029
	(0.083)	(0.108)
	[4,542]	[4,039]
Father's characteristics		
Completed grades of schooling	0.273	0.166
	(1.172)	(1.309)
	[2,810]	[2,621]
Currently alive (indicator)	0.010	-0.093
	(0.080)	(0.169)
	[4,541]	[4,040]

Placebo test estimates

	Women	Men
Self-rep. health status very good (indic.)	0.123 (0.099) [1,239]	-0.115 (0.078) [1,264]
Self-rep. health status poor/very poor (indic.)	0.090 (0.154) [1,239]	0.106 (0.134) [1,264]
Ln (lung capacity)	-0.067 (0.034)* [1,195]	0.008 (0.089) [1,130]
Height (cm.)	-1.165 (1.660) [1,207]	3.054 (2.017) [1,132]
Days absent due to illness (last 4 weeks)	0.669 (0.688) [1,240]	3.075 (1.505)* [1,261]
Completed grades of schooling	0.958 (1.274) [1,240]	-1.441 (1.947) [1,260]
Ln (expenditures per cap. in hh)	-0.193 (0.284) [1,240]	-0.329 (0.189) [1,264]
Asset index	-0.773 (0.497) [1,240]	0.166 (0.353) [1,264]
Ln (annual earnings)	0.202 (0.333) [631]	-0.612 (0.344) [1,142]

2SLS estimates

TABLE 2—EFFECT OF BIRTH YEAR RAINFALL ON ADULT OUTCOMES: WOMEN AND MEN BORN 1953–1974
(Instrumental variables estimates. Coefficients (standard errors) in regression of outcome on rainfall in individual's birth year and birth district. Instrumental variables for birth year/birth district rainfall are rainfall measured at second- through fifth-closest rainfall stations to respondent's birth district.)

	Women	Men
Self-reported health status very good (indicator)	0.101	−0.029
	(0.058)*	(0.072)
	[4,613]	[4,270]
Self-reported health status poor/very poor (indicator)	−0.192	−0.100
	(0.082)**	(0.098)
	[4,613]	[4,270]
Ln (lung capacity)	−0.044	−0.073
	(0.049)	(0.062)
	[4,454]	[3,907]
Height (centimeters)	2.832	0.998
	(0.821)***	(1.795)
	[4,495]	[3,924]
Days absent due to illness (last four weeks)	−1.175	0.515
	(0.831)	(0.779)
	[4,611]	[4,267]
Completed grades of schooling	1.086	−0.474
	(0.453)**	(1.490)
	[4,598]	[4,259]
Ln (expenditures per capita in household)	0.095	−0.274
	(0.204)	(0.301)
	[4,615]	[4,277]
Asset index	0.876	−0.279
	(0.324)**	(0.507)
	[4,613]	[4,276]
Ln (annual earnings)	0.065	−0.202
	(0.988)	(0.350)
	[2,332]	[3,963]

2SLS estimates

TABLE 3—EFFECT OF RAINFALL IN YEARS BEFORE AND AFTER BIRTH: WOMEN BORN 1953–1974
(Instrumental variables estimates. Rainfall in individual's birth year and birth district instrumented with rainfall measured at second- through fifth-closest rainfall stations to respondent's birth district.)

Dependent variable	Self-reported health status very good (indicator)	Self-reported health status poor/very poor (indicator)	Height (centimeters)	Completed grades of schooling	Asset index
Coefficient on rainfall in:					
Year −3	0.025	−0.114	1.505	−0.065	0.003
	(0.084)	(0.120)	(1.572)	(0.992)	(0.424)
Year −2	−0.037	−0.013	0.854	−0.852	−0.426
	(0.103)	(0.075)	(1.813)	(1.670)	(0.721)
Year −1	−0.080	−0.045	3.338	0.104	−0.380
	(0.123)	(0.088)	(2.155)	(1.332)	(0.530)
Year 0	0.090	−0.179	3.833	1.598	0.750
	(0.067)	(0.093)*	(1.420)**	(0.675)**	(0.399)*
Year 1	−0.008	−0.096	0.676	1.083	0.203
	(0.053)	(0.067)	(1.592)	(0.769)	(0.272)
Year 2	−0.041	−0.015	1.666	0.117	−0.229
	(0.043)	(0.068)	(0.984)	(0.840)	(0.452)
Year 3	−0.020	−0.104	1.996	−0.135	0.088
	(0.116)	(0.067)	(1.774)	(0.802)	(0.232)
Observations	4,613	4,613	4,495	4,598	4,613

Recap

TL;DR:

1. Instrumental variables are very powerful
2. With the right assumptions...
3. ...we can handle OVB and ME (and simultaneity)

For next class

Topics:

- Instrumental variables III

Reading: Day off!

- Take another look at Fowlie, Wolfram, et al.