

Instrumental Variables

EC 607, Set 9

Edward Rubin

Prologue

Schedule

Last time

Matching and propensity-score methods

- Conditional independence
- Overlap

Today

Instrumental variables (and two-stage least squares)

Upcoming

Assignment 2

Research designs

Research designs

Selection on observables and/or unobservables

We've been focusing on **selection-on-observables designs**, *i.e.*,

$$(Y_{0i}, Y_{1i}) \perp\!\!\!\perp D_i | X_i$$

for **observable** variables X_i .

Research designs

Selection on observables and/or unobservables

We've been focusing on **selection-on-observables designs**, i.e.,

$$(Y_{0i}, Y_{1i}) \perp\!\!\!\perp D_i | X_i$$

for **observable** variables X_i .

Selection-on-unobservable designs replace this assumption with two new (but related) assumptions

1. $(Y_{0i}, Y_{1i}) \perp Z_i$
2. $\text{Cov}(Z_i, D_i) \neq 0$

Research designs

Selection on observables and/or unobservables

Our main goal in causal-inference minded (applied) econometrics boils down to isolating **"good" variation** in D_i (exogenous/as-good-as-random) from **"bad" variation** (the part of D_i correlated with Y_{0i} and Y_{1i}).

Research designs

Selection on observables and/or unobservables

Our main goal in causal-inference minded (applied) econometrics boils down to isolating **"good" variation** in D_i (exogenous/as-good-as-random) from **"bad" variation** (the part of D_i correlated with Y_{0i} and Y_{1i}).

(We want to avoid selection bias.)

Research designs

Selection on observables and/or unobservables

Our main goal in causal-inference minded (applied) econometrics boils down to isolating **"good" variation** in D_i (exogenous/as-good-as-random) from **"bad" variation** (the part of D_i correlated with Y_{0i} and Y_{1i}).

(We want to avoid selection bias.)

- **Selection-on-observables designs** assume that we can control for all *bad variation* (selection) in D_i through a known (observed) X_i .

Research designs

Selection on observables and/or unobservables

Our main goal in causal-inference minded (applied) econometrics boils down to isolating **"good" variation** in D_i (exogenous/as-good-as-random) from **"bad" variation** (the part of D_i correlated with Y_{0i} and Y_{1i}).

(We want to avoid selection bias.)

- **Selection-on-observables designs** assume that we can control for all *bad variation* (selection) in D_i through a known (observed) X_i .
- **Selection-on-unobservables designs** assume that we can extract **part of** the *good variation* in D_i (generally using some Z_i) and then use this *good part* of D_i to estimate the effect of D_i on Y_i .

Research designs

Selection on observables and/or unobservables

Our main goal in causal-inference minded (applied) econometrics boils down to isolating **"good" variation** in D_i (exogenous/as-good-as-random) from **"bad" variation** (the part of D_i correlated with Y_{0i} and Y_{1i}).

(We want to avoid selection bias.)

- **Selection-on-observables designs** assume that we can control for all *bad variation* (selection) in D_i through a known (observed) X_i .
- **Selection-on-unobservables designs** assume that we can extract **part of** the *good variation* in D_i (generally using some Z_i) and then use this *good part* of D_i to estimate the effect of D_i on Y_i . We throw away the rest of D_i (it includes *bad variation*).

Research designs

Which route?

Which set of research designs is more palatable?

Research designs

Which route?

Which set of research designs is more palatable?

1. There are plenty of bad applications of both sets.
Violated assumptions, bad controls, etc.

Research designs

Which route?

Which set of research designs is more palatable?

1. There are plenty of bad applications of both sets.

Violated assumptions, bad controls, etc.

2. **Selection on observables** assumes we know *everything* about selection into treatment—we can identify *all* of the good (or bad) variation in \mathbf{D}_i .

Research designs

Which route?

Which set of research designs is more palatable?

1. There are plenty of bad applications of both sets.
Violated assumptions, bad controls, etc.
2. **Selection on observables** assumes we know *everything* about selection into treatment—we can identify *all* of the good (or bad) variation in \mathbf{D}_i .
Tough in non-experimental settings. Difficult to validate in practice.

Research designs

Which route?

Which set of research designs is more palatable?

1. There are plenty of bad applications of both sets.
Violated assumptions, bad controls, etc.
2. **Selection on observables** assumes we know *everything* about selection into treatment—we can identify *all* of the good (or bad) variation in \mathbf{D}_i .
Tough in non-experimental settings. Difficult to validate in practice.
3. **Selection on unobservables** assumes we can isolate *some* good/clean variation in \mathbf{D}_i , which we then use to estimate the effect of \mathbf{D}_i on \mathbf{Y}_i .

Research designs

Which route?

Which set of research designs is more palatable?

1. There are plenty of bad applications of both sets.
Violated assumptions, bad controls, etc.
2. **Selection on observables** assumes we know *everything* about selection into treatment—we can identify *all* of the good (or bad) variation in \mathbf{D}_i .
Tough in non-experimental settings. Difficult to validate in practice.
3. **Selection on unobservables** assumes we can isolate *some* good/clean variation in \mathbf{D}_i , which we then use to estimate the effect of \mathbf{D}_i on \mathbf{Y}_i .
Seems more plausible. Possible to validate. May be underpowered.

Instrumental variables

Introduction

Instrumental variables (IV)[†] is the canonical selection-on-unobservables design—isolating *good variation* in \mathbf{D}_i via some magical **instrument** \mathbf{Z}_i .

[†] For the moment, we're lumping together IV and two-stage least squares (2SLS) together—as many people do—even though they are technically different.

Instrumental variables

Introduction

Instrumental variables (IV)[†] is the canonical selection-on-unobservables design—isolating *good variation* in \mathbf{D}_i via some magical **instrument** \mathbf{Z}_i .

Consider some model (structural equation)

$$\mathbf{Y}_i = \beta_0 + \beta_1 \mathbf{D}_i + \varepsilon_i \quad (1)$$

To guarantee consistent OLS estimates for β_1 , want $\text{Cov}(\mathbf{D}_i, \varepsilon_i) = 0$.
In general, this is a heroic assumption.

[†] For the moment, we're lumping together IV and two-stage least squares (2SLS) together—as many people do—even though they are technically different.

Instrumental variables

Introduction

Instrumental variables (IV)[†] is the canonical selection-on-unobservables design—isolating *good variation* in \mathbf{D}_i via some magical **instrument** \mathbf{Z}_i .

Consider some model (structural equation)

$$\mathbf{Y}_i = \beta_0 + \beta_1 \mathbf{D}_i + \varepsilon_i \quad (1)$$

To guarantee consistent OLS estimates for β_1 , want $\text{Cov}(\mathbf{D}_i, \varepsilon_i) = 0$.
In general, this is a heroic assumption.

Alternative: Estimate β_1 via instrumental variables.

[†] For the moment, we're lumping together IV and two-stage least squares (2SLS) together—as many people do—even though they are technically different.

Instrumental variables

Definition

For our model

$$Y_i = \beta_0 + \beta_1 D_i + \varepsilon_i \quad (1)$$

A valid **instrument** is a variable Z_i such that

1. $\text{Cov}(Z_i, D_i) \neq 0$

Instrumental variables

Definition

For our model

$$Y_i = \beta_0 + \beta_1 D_i + \varepsilon_i \quad (1)$$

A valid **instrument** is a variable Z_i such that

1. $\text{Cov}(Z_i, D_i) \neq 0$
our **instrument** correlates with treatment

Instrumental variables

Definition

For our model

$$Y_i = \beta_0 + \beta_1 D_i + \varepsilon_i \quad (1)$$

A valid **instrument** is a variable Z_i such that

1. $\text{Cov}(Z_i, D_i) \neq 0$

our **instrument** correlates with treatment (so we can keep part of D_i)

Instrumental variables

Definition

For our model

$$Y_i = \beta_0 + \beta_1 D_i + \varepsilon_i \quad (1)$$

A valid **instrument** is a variable Z_i such that

1. $\text{Cov}(Z_i, D_i) \neq 0$
our **instrument** correlates with treatment (so we can keep part of D_i)
2. $\text{Cov}(Z_i, \varepsilon_i) = 0$

Instrumental variables

Definition

For our model

$$Y_i = \beta_0 + \beta_1 D_i + \varepsilon_i \quad (1)$$

A valid **instrument** is a variable Z_i such that

1. $\text{Cov}(Z_i, D_i) \neq 0$
our **instrument** correlates with treatment (so we can keep part of D_i)
2. $\text{Cov}(Z_i, \varepsilon_i) = 0$
our **instrument** is uncorrelated with other (non- D_i) determinants of Y_i

Instrumental variables

Definition

For our model

$$Y_i = \beta_0 + \beta_1 D_i + \varepsilon_i \quad (1)$$

A valid **instrument** is a variable Z_i such that

1. $\text{Cov}(Z_i, D_i) \neq 0$
our **instrument** correlates with treatment (so we can keep part of D_i)
2. $\text{Cov}(Z_i, \varepsilon_i) = 0$
our **instrument** is uncorrelated with other (non- D_i) determinants of Y_i ,
i.e., Z_i is excludable from equation (1).

Instrumental variables

Definition

For our model

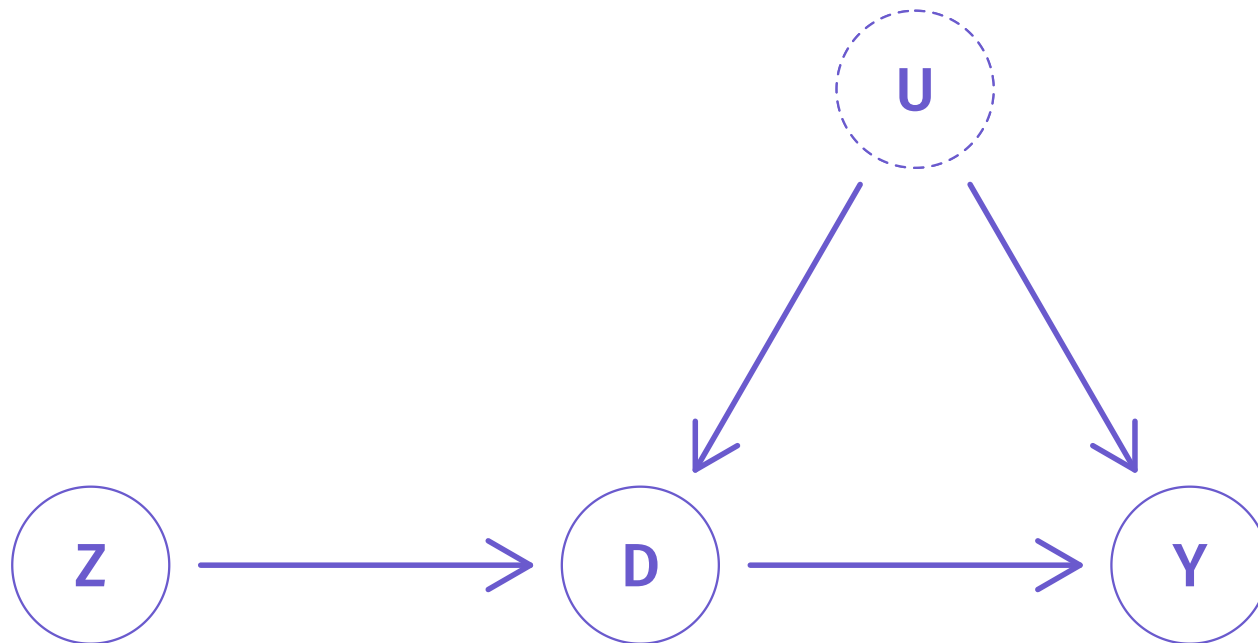
$$Y_i = \beta_0 + \beta_1 D_i + \varepsilon_i \quad (1)$$

A valid **instrument** is a variable Z_i such that

1. $\text{Cov}(Z_i, D_i) \neq 0$
our **instrument** correlates with treatment (so we can keep part of D_i)
2. $\text{Cov}(Z_i, \varepsilon_i) = 0$
our **instrument** is uncorrelated with other (non- D_i) determinants of Y_i ,
i.e., Z_i is excludable from equation (1). (**exclusion restriction**)

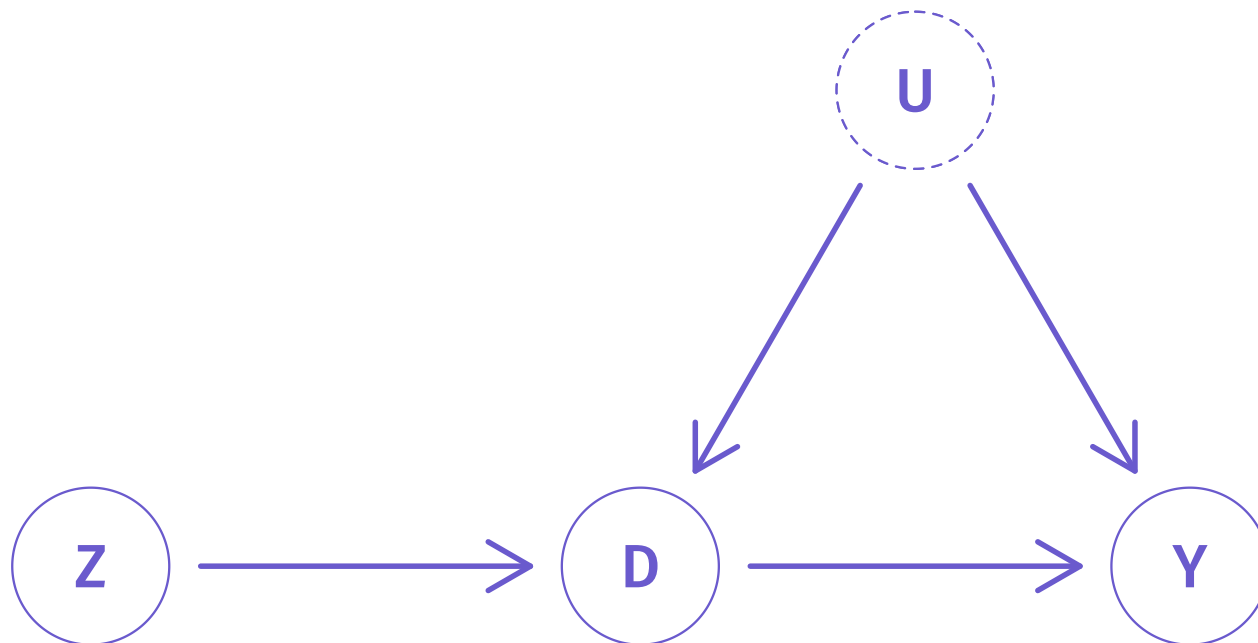
Instrumental variables

The DAG



Instrumental variables

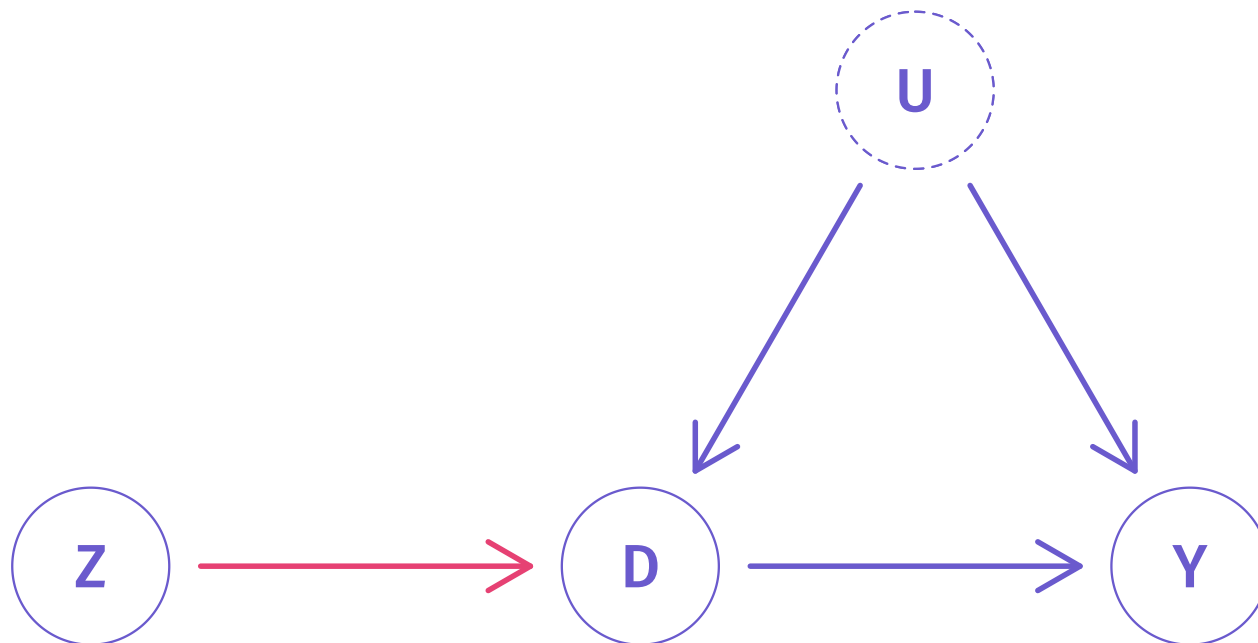
The DAG



Q How does this DAG illustrate the requirements and identification of IV?

Instrumental variables

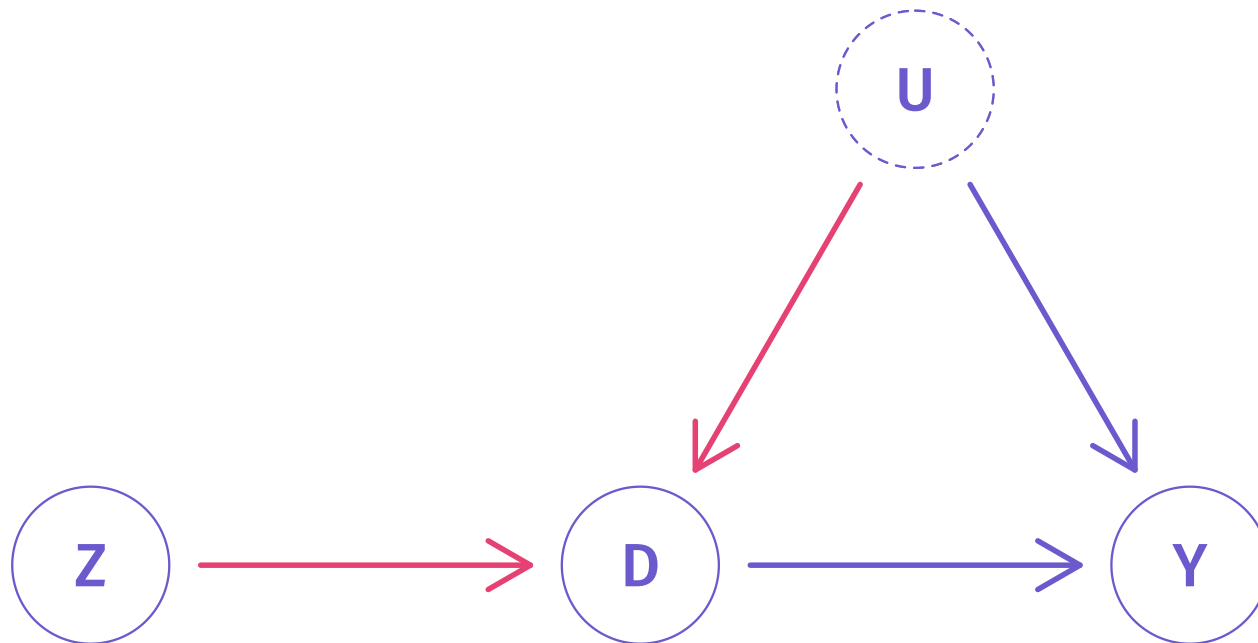
The DAG



Relevance: **Z** causes an effect in **D**.

Instrumental variables

The DAG

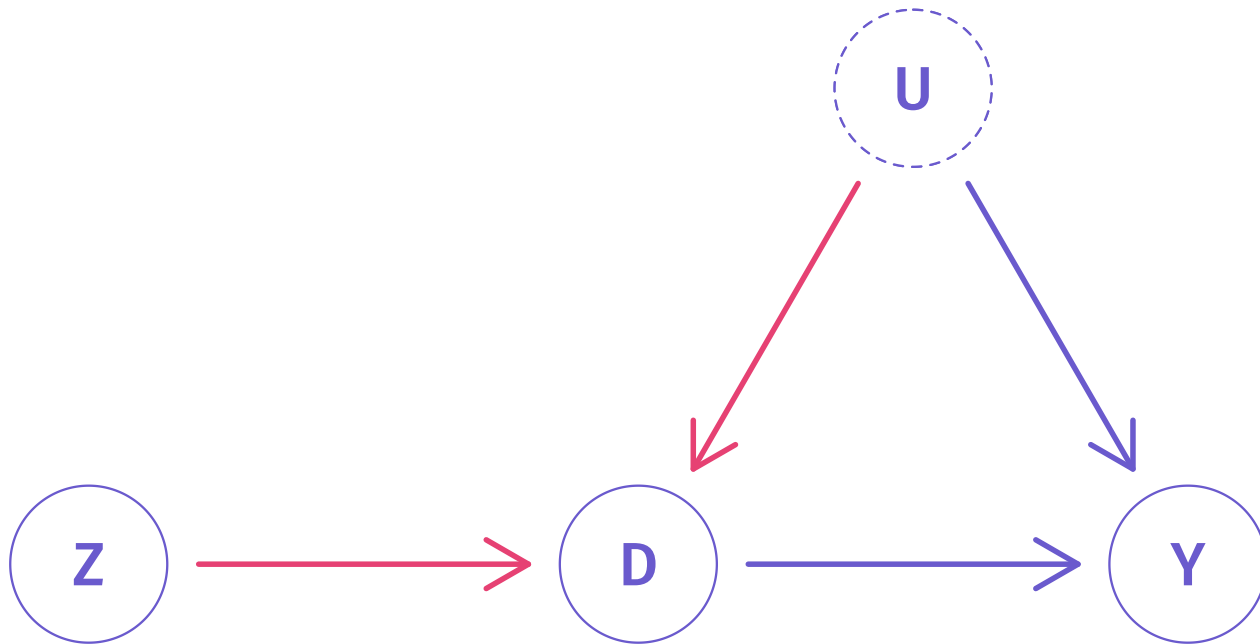


Exclusion restriction:

1. **Z** is **exogenous** (not associated with) **U** because

Instrumental variables

The DAG

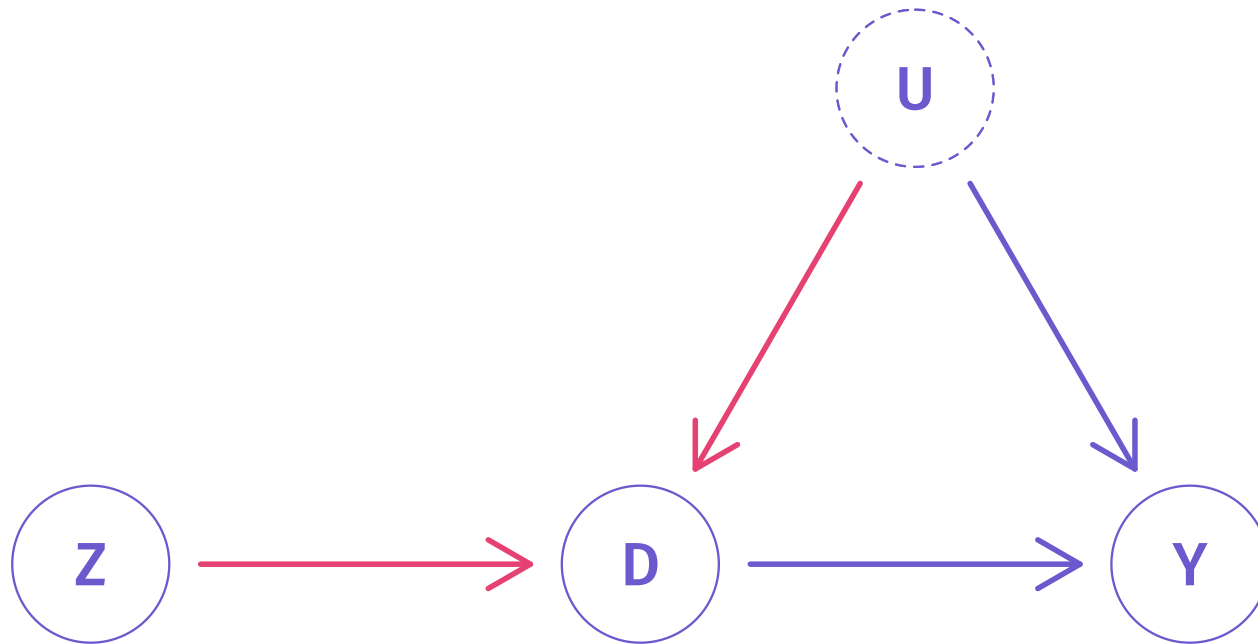


Exclusion restriction:

1. **Z** is **exogenous** (not associated with) **U** because **D** is a collider.

Instrumental variables

The DAG



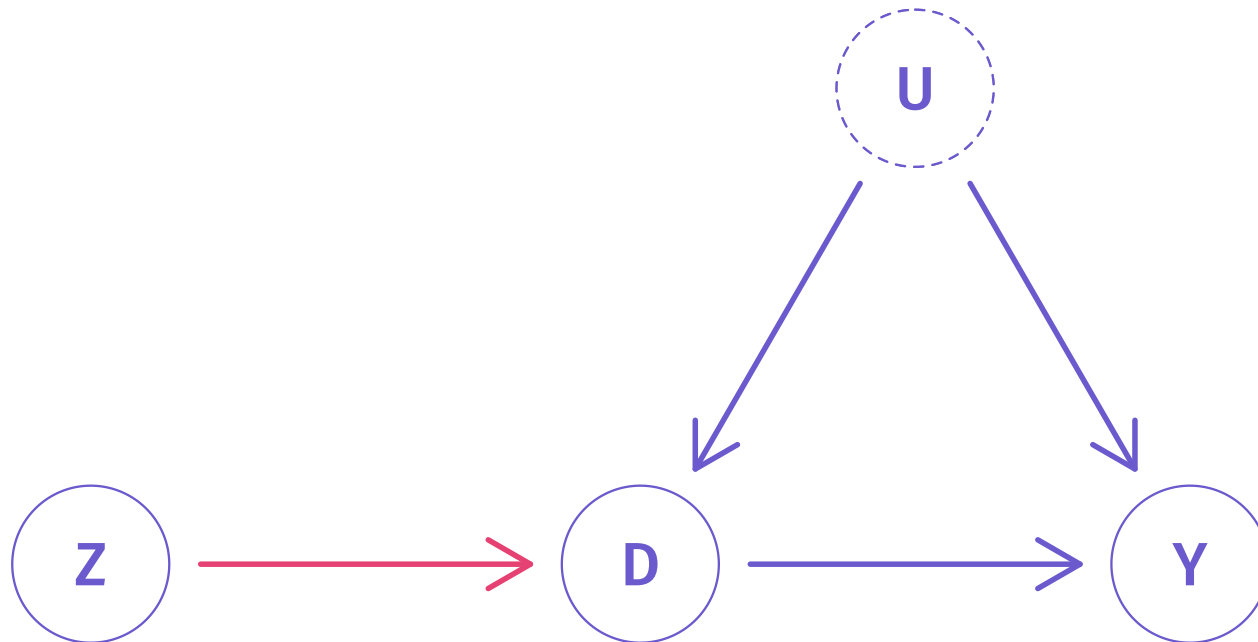
Exclusion restriction:

1. **Z** is **exogenous** (not associated with) **U** because **D** is a collider.

I.e., $Z \rightarrow D \leftarrow U \rightarrow Y$ is closed without conditioning on (unobservable) **U**.

Instrumental variables

The DAG



Exclusion restriction:

1. **Z** is **exogenous** (not associated with) **U** because **D** is a collider.
2. Also: **Z** does not directly cause **Y**.

Instrumental variables

Example

Back to the returns to a college degree,

$$\text{Income}_i = \beta_0 + \beta_1 \text{Grad}_i + \varepsilon_i$$

OLS is likely biased.

Instrumental variables

Example

Back to the returns to a college degree,

$$\text{Income}_i = \beta_0 + \beta_1 \text{Grad}_i + \varepsilon_i$$

OLS is likely biased.

What if that state conducts a (random) **lottery** for scholarships?

Instrumental variables

Example

Back to the returns to a college degree,

$$\text{Income}_i = \beta_0 + \beta_1 \text{Grad}_i + \varepsilon_i$$

OLS is likely biased.

What if that state conducts a (random) **lottery** for scholarships?

Let **Lottery**_{*i*} denote an indicator for whether *i* won a lottery scholarship.[†]

[†] We'll have to focus on families who were eligible/who applied.

Instrumental variables

Example

Back to the returns to a college degree,

$$\text{Income}_i = \beta_0 + \beta_1 \text{Grad}_i + \varepsilon_i$$

OLS is likely biased.

What if that state conducts a (random) **lottery** for scholarships?

Let **Lottery**_{*i*} denote an indicator for whether *i* won a lottery scholarship.[†]

1. $\text{Cov}(\text{Lottery}_i, \text{Grad}_i) \neq 0 (> 0)$ if scholarships increase grad. rates.

[†] We'll have to focus on families who were eligible/who applied.

Instrumental variables

Example

Back to the returns to a college degree,

$$\text{Income}_i = \beta_0 + \beta_1 \text{Grad}_i + \varepsilon_i$$

OLS is likely biased.

What if that state conducts a (random) **lottery** for scholarships?

Let **Lottery**_{*i*} denote an indicator for whether *i* won a lottery scholarship.[†]

1. $\text{Cov}(\text{Lottery}_i, \text{Grad}_i) \neq 0$ (> 0) if scholarships increase grad. rates.
2. $\text{Cov}(\text{Lottery}_i, \varepsilon_i) = 0$ since the lottery is randomized.

[†] We'll have to focus on families who were eligible/who applied.

Instrument variables

The IV estimator

The IV estimator for our model

$$Y_i = \beta_0 + \beta_1 D_i + \varepsilon_i \quad (1)$$

with (valid) instrument Z_i is

$$\hat{\beta}_{IV} = (Z'D)^{-1} (Z'Y)$$

Instrument variables

The IV estimator

The IV estimator for our model

$$Y_i = \beta_0 + \beta_1 D_i + \varepsilon_i \quad (1)$$

with (valid) instrument Z_i is

$$\hat{\beta}_{IV} = (Z'D)^{-1} (Z'Y)$$

If you have no covariates, then

$$\hat{\beta}_{IV} = \frac{\text{Cov}(Z_i, Y_i)}{\text{Cov}(Z_i, D_i)}$$

Instrument variables

The IV estimator

The IV estimator for our model

$$Y_i = \beta_0 + \beta_1 D_i + \varepsilon_i \quad (1)$$

with (valid) instrument Z_i is

$$\hat{\beta}_{IV} = (Z'D)^{-1} (Z'Y)$$

If you have additional (exogenous) covariates X_i , then

$$Z = [Z_i \quad X_i]$$

$$D = [D_i \quad X_i]$$

Instrumental variables

Proof: Consistency

With a valid instrument \mathbf{Z}_i , $\hat{\beta}_{IV}$ is a consistent estimator for β_1 in

$$\mathbf{Y}_i = \beta_0 + \beta_1 \mathbf{X}_i + \varepsilon_i \quad (1)$$

$$\text{plim}(\hat{\beta}_{IV})$$

Instrumental variables

Proof: Consistency

With a valid instrument Z_i , $\hat{\beta}_{IV}$ is a consistent estimator for β_1 in

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i \quad (1)$$

$$\text{plim}(\hat{\beta}_{IV})$$

$$= \text{plim}\left((Z'D)^{-1} (Z'Y)\right)$$

Instrumental variables

Proof: Consistency

With a valid instrument Z_i , $\hat{\beta}_{IV}$ is a consistent estimator for β_1 in

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i \quad (1)$$

$$\text{plim}(\hat{\beta}_{IV})$$

$$= \text{plim}\left((Z'D)^{-1} (Z'Y)\right)$$

$$= \text{plim}\left((Z'D)^{-1} (Z'D\beta + Z'\varepsilon)\right)$$

Instrumental variables

Proof: Consistency

With a valid instrument \mathbf{Z}_i , $\hat{\beta}_{IV}$ is a consistent estimator for β_1 in

$$\mathbf{Y}_i = \beta_0 + \beta_1 \mathbf{X}_i + \varepsilon_i \quad (1)$$

$$\text{plim}(\hat{\beta}_{IV})$$

$$= \text{plim}\left((\mathbf{Z}'\mathbf{D})^{-1} (\mathbf{Z}'\mathbf{Y})\right)$$

$$= \text{plim}\left((\mathbf{Z}'\mathbf{D})^{-1} (\mathbf{Z}'\mathbf{D}\beta + \mathbf{Z}'\varepsilon)\right)$$

$$= \text{plim}\left((\mathbf{Z}'\mathbf{D})^{-1} (\mathbf{Z}'\mathbf{D}) \beta\right) + \text{plim}\left(\frac{1}{N}\mathbf{Z}'\mathbf{D}\right)^{-1} \text{plim}\left(\frac{1}{N}\mathbf{Z}'\varepsilon\right)$$

Instrumental variables

Proof: Consistency

With a valid instrument Z_i , $\hat{\beta}_{IV}$ is a consistent estimator for β_1 in

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i \quad (1)$$

$$\text{plim}(\hat{\beta}_{IV})$$

$$= \text{plim}\left((Z'D)^{-1} (Z'Y)\right)$$

$$= \text{plim}\left((Z'D)^{-1} (Z'D\beta + Z'\varepsilon)\right)$$

$$= \text{plim}\left((Z'D)^{-1} (Z'D) \beta\right) + \text{plim}\left(\frac{1}{N} Z'D\right)^{-1} \text{plim}\left(\frac{1}{N} Z'\varepsilon\right)$$

$$= \beta \quad \checkmark$$

Two-stage least squares

Two-stage least squares

Setup

You'll commonly see IV implemented as a two-stage process known as **two-stage least squares** (2SLS).

Two-stage least squares

Setup

You'll commonly see IV implemented as a two-stage process known as **two-stage least squares** (2SLS).

First stage Estimate the effect of the instrument \mathbf{Z}_i on our endogenous variable \mathbf{D}_i and (predetermined) covariates \mathbf{X}_i . Save $\hat{\mathbf{D}}_i$.

$$\mathbf{D}_i = \gamma_1 \mathbf{Z}_i + \gamma_2 \mathbf{X}_i + u_i$$

Two-stage least squares

Setup

You'll commonly see IV implemented as a two-stage process known as **two-stage least squares** (2SLS).

First stage Estimate the effect of the instrument \mathbf{Z}_i on our endogenous variable \mathbf{D}_i and (predetermined) covariates \mathbf{X}_i . Save $\hat{\mathbf{D}}_i$.

$$\mathbf{D}_i = \gamma_1 \mathbf{Z}_i + \gamma_2 \mathbf{X}_i + u_i$$

Second stage Estimate the model we wanted—but only using the variation in \mathbf{D}_i that correlates with \mathbf{Z}_i , i.e., $\hat{\mathbf{D}}_i$.

$$\mathbf{Y}_i = \beta_1 \hat{\mathbf{D}}_i + \beta_2 \mathbf{X}_i + \varepsilon_i$$

Note The controls \mathbf{X}_i must match in the first and second stages.

Two-stage least squares

IV estimation

This two-step procedure, with a valid instrument, produces an estimator $\hat{\beta}_1$ that is consistent for β_1 .

$$\hat{\beta}_{2SLS} = (\mathbf{D}'\mathbf{P}_Z\mathbf{D})^{-1} (\mathbf{D}'\mathbf{P}_Z\mathbf{Y})$$

$$\mathbf{P}_Z = \mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'$$

where \mathbf{D} is a matrix of our treatment and predetermined covariates (\mathbf{X}_i) and \mathbf{Z} is a matrix of our instrument and our predetermined covariates.

Two-stage least squares

IV estimation

Important notes

- The controls (\mathbf{X}_i) must match in the first and second stages.
- *Related:* Nonlinear first stages can mess things up.
- If you have exactly **one instrument** and exactly **one endogenous variable**, then 2SLS and IV are identical.
- Your second-stage standard errors are not correct.

Two-stage least squares

The reduced form

In addition to the regressions within the two stages of 2SLS

1. $D_i = \gamma_1 Z_i + \gamma_2 X_i + u_i$

2. $Y_i = \beta_1 \hat{D}_i + \beta_2 X_i + \varepsilon_i$

there is a third important and related regression: the reduced form.

Two-stage least squares

The reduced form

In addition to the regressions within the two stages of 2SLS

1. $D_i = \gamma_1 Z_i + \gamma_2 X_i + u_i$
2. $Y_i = \beta_1 \hat{D}_i + \beta_2 X_i + \varepsilon_i$

there is a third important and related regression: the reduced form.

The **reduced form** regresses the outcome Y_i (LHS of the second stage) on our instrument Z_i and covariates X_i (RHS of the first stage).

$$Y_i = \pi_1 Z_i + \pi_2 X_i + u_i$$

Two-stage least squares

The reduced form

In addition to the regressions within the two stages of 2SLS

1. $D_i = \gamma_1 Z_i + \gamma_2 X_i + u_i$
2. $Y_i = \beta_1 \hat{D}_i + \beta_2 X_i + \varepsilon_i$

there is a third important and related regression: the reduced form.

The **reduced form** regresses the outcome Y_i (LHS of the second stage) on our instrument Z_i and covariates X_i (RHS of the first stage).

$$Y_i = \pi_1 Z_i + \pi_2 X_i + u_i$$

Thus, the reduced form provides a consistent estimate of the causal effect of our instrument on the outcome.

Two-stage least squares

The reduced form, continued

While the reduced form estimates the causal effect of the instrument on our outcome, we're often actually interested in the effect of *treatment* (\mathbf{D}_i).

Two-stage least squares

The reduced form, continued

While the reduced form estimates the causal effect of the instrument on our outcome, we're often actually interested in the effect of *treatment* (\mathbf{D}_i).

That said, the reduced form is still incredibly helpful/important:

- Clarifies your source of identifying variation.

Two-stage least squares

The reduced form, continued

While the reduced form estimates the causal effect of the instrument on our outcome, we're often actually interested in the effect of *treatment* (D_i).

That said, the reduced form is still incredibly helpful/important:

- Clarifies your source of identifying variation.
- Does not suffer from *weak instruments* problems.

Two-stage least squares

The reduced form, continued

While the reduced form estimates the causal effect of the instrument on our outcome, we're often actually interested in the effect of *treatment* (\mathbf{D}_i).

That said, the reduced form is still incredibly helpful/important:

- Clarifies your source of identifying variation.
- Does not suffer from *weak instruments* problems.
- Only requires $\text{Cov}(\mathbf{Z}_i, \varepsilon_i) = 0$.

Two-stage least squares

The reduced form, continued

While the reduced form estimates the causal effect of the instrument on our outcome, we're often actually interested in the effect of *treatment* (\mathbf{D}_i).

That said, the reduced form is still incredibly helpful/important:

- Clarifies your source of identifying variation.
- Does not suffer from *weak instruments* problems.
- Only requires $\text{Cov}(\mathbf{Z}_i, \varepsilon_i) = \mathbf{0}$.
- Offers insights into your estimates

Two-stage least squares

The reduced form, continued

While the reduced form estimates the causal effect of the instrument on our outcome, we're often actually interested in the effect of *treatment* (\mathbf{D}_i).

That said, the reduced form is still incredibly helpful/important:

- Clarifies your source of identifying variation.
- Does not suffer from *weak instruments* problems.
- Only requires $\text{Cov}(\mathbf{Z}_i, \varepsilon_i) = 0$.
- Offers insights into your estimates

$$\hat{\beta}_1^{2SLS} = \frac{\hat{\pi}_1}{\hat{\gamma}_1}$$

when you have exactly one instrument.

Two-stage least squares

The reduced form, intuition

This expression for the 2SLS (and IV) estimator can be very helpful.

$$\hat{\beta}_1^{2SLS} = \frac{\hat{\pi}_1}{\hat{\gamma}_1} = \frac{\text{Reduced-form estimate}}{\text{First-stage estimate}}$$

Two-stage least squares

The reduced form, intuition

This expression for the 2SLS (and IV) estimator can be very helpful.

$$\hat{\beta}_1^{2SLS} = \frac{\hat{\pi}_1}{\hat{\gamma}_1} = \frac{\text{Reduced-form estimate}}{\text{First-stage estimate}}$$

What's the interpretation/intuition?

Two-stage least squares

The reduced form, intuition

This expression for the 2SLS (and IV) estimator can be very helpful.

$$\hat{\beta}_1^{2SLS} = \frac{\hat{\pi}_1}{\hat{\gamma}_1} = \frac{\text{Reduced-form estimate}}{\text{First-stage estimate}}$$

What's the interpretation/intuition?

Back to our example: $\hat{\beta}_1$ = est. effect of college graduation on income.

Two-stage least squares

The reduced form, intuition

This expression for the 2SLS (and IV) estimator can be very helpful.

$$\hat{\beta}_1^{2SLS} = \frac{\hat{\pi}_1}{\hat{\gamma}_1} = \frac{\text{Reduced-form estimate}}{\text{First-stage estimate}}$$

What's the interpretation/intuition?

Back to our example: $\hat{\beta}_1$ = est. effect of college graduation on income.

$\hat{\pi}_1$ gives the estimated causal effect of the scholarship lottery on income

Two-stage least squares

The reduced form, intuition

This expression for the 2SLS (and IV) estimator can be very helpful.

$$\hat{\beta}_1^{2SLS} = \frac{\hat{\pi}_1}{\hat{\gamma}_1} = \frac{\text{Reduced-form estimate}}{\text{First-stage estimate}}$$

What's the interpretation/intuition?

Back to our example: $\hat{\beta}_1$ = est. effect of college graduation on income.

$\hat{\pi}_1$ gives the estimated causal effect of the scholarship lottery on income, but what share of lottery winners graduate? We need to rescale if $< 100\%$.

Two-stage least squares

The reduced form, intuition

This expression for the 2SLS (and IV) estimator can be very helpful.

$$\hat{\beta}_1^{2SLS} = \frac{\hat{\pi}_1}{\hat{\gamma}_1} = \frac{\text{Reduced-form estimate}}{\text{First-stage estimate}}$$

What's the interpretation/intuition?

Back to our example: $\hat{\beta}_1$ = est. effect of college graduation on income.

$\hat{\pi}_1$ gives the estimated causal effect of the scholarship lottery on income, but what share of lottery winners graduate? We need to rescale if $< 100\%$.

$\hat{\gamma}_1$ estimates the effect of winning the scholarship lottery on graduation

Two-stage least squares

The reduced form, intuition

This expression for the 2SLS (and IV) estimator can be very helpful.

$$\hat{\beta}_1^{2SLS} = \frac{\hat{\pi}_1}{\hat{\gamma}_1} = \frac{\text{Reduced-form estimate}}{\text{First-stage estimate}}$$

What's the interpretation/intuition?

Back to our example: $\hat{\beta}_1$ = est. effect of college graduation on income.

$\hat{\pi}_1$ gives the estimated causal effect of the scholarship lottery on income, but what share of lottery winners graduate? We need to rescale if $< 100\%$.

$\hat{\gamma}_1$ estimates the effect of winning the scholarship lottery on graduation—the share of winners who graduated due to winning.

Two-stage least squares

The reduced form, intuition

This expression for the 2SLS (and IV) estimator can be very helpful.

$$\hat{\beta}_1^{2SLS} = \frac{\hat{\pi}_1}{\hat{\gamma}_1} = \frac{\text{Reduced-form estimate}}{\text{First-stage estimate}}$$

What's the interpretation/intuition?

Back to our example: $\hat{\beta}_1$ = est. effect of college graduation on income.

$\hat{\pi}_1$ gives the estimated causal effect of the scholarship lottery on income, but what share of lottery winners graduate? We need to rescale if $< 100\%$.

$\hat{\gamma}_1$ estimates the effect of winning the scholarship lottery on graduation—the share of winners who graduated due to winning. We can scale with $\hat{\gamma}_1$!

Two-stage least squares

The reduced form, example

To see why this scaling makes sense, imagine that 50% of lottery winners graduate from college due to the lottery, *i.e.*, $\hat{\gamma}_1 = 0.50$.[†]

[†] Imagine none of the applicants would have graduated otherwise

Two-stage least squares

The reduced form, example

To see why this scaling makes sense, imagine that 50% of lottery winners graduate from college due to the lottery, *i.e.*, $\hat{\gamma}_1 = 0.50$.[†]

Our reduced-form estimate of $\hat{\pi}_1 = \$5,000$ says that lottery winners make \$5,000 more than the control group, on average.

[†] Imagine none of the applicants would have graduated otherwise

Two-stage least squares

The reduced form, example

To see why this scaling makes sense, imagine that 50% of lottery winners graduate from college due to the lottery, *i.e.*, $\hat{\gamma}_1 = 0.50$.[†]

Our reduced-form estimate of $\hat{\pi}_1 = \$5,000$ says that lottery winners make \$5,000 more than the control group, on average.

However, half of the winners did not graduate, so $\hat{\pi}_1$ "underestimates" the effect of college graduation by combining graduates by nongraduates.

[†] Imagine none of the applicants would have graduated otherwise

Two-stage least squares

The reduced form, example

To see why this scaling makes sense, imagine that 50% of lottery winners graduate from college due to the lottery, *i.e.*, $\hat{\gamma}_1 = 0.50$.[†]

Our reduced-form estimate of $\hat{\pi}_1 = \$5,000$ says that lottery winners make \$5,000 more than the control group, on average.

However, half of the winners did not graduate, so $\hat{\pi}_1$ "underestimates" the effect of college graduation by combining graduates by nongraduates.

Thus, we want to double $\hat{\pi}_1$, *i.e.*, divide by $\hat{\gamma}_1$: $\hat{\pi}_1 / \hat{\gamma}_1 = \$5,000 / 0.5 = \$10,000$.

[†] Imagine none of the applicants would have graduated otherwise

Two-stage least squares

Q How do we get this magical expression? $\left(\hat{\beta}_1^{\text{IV}} = \frac{\hat{\pi}_1}{\hat{\gamma}_1} \right)$

Two-stage least squares

Q How do we get this magical expression? $\left(\hat{\beta}_1^{\text{IV}} = \frac{\hat{\pi}_1}{\hat{\gamma}_1} \right)$

Derivation

Two-stage least squares

Q How do we get this magical expression? $\left(\hat{\beta}_1^{\text{IV}} = \frac{\hat{\pi}_1}{\hat{\gamma}_1} \right)$

Derivation

$$\hat{\beta}_1^{\text{IV}} = (\mathbf{Z}'\mathbf{D})^{-1} (\mathbf{Z}'\mathbf{Y})$$

Two-stage least squares

Q How do we get this magical expression? $\left(\hat{\beta}_1^{\text{IV}} = \frac{\hat{\pi}_1}{\hat{\gamma}_1} \right)$

Derivation

$$\begin{aligned} \hat{\beta}_1^{\text{IV}} &= (\mathbf{Z}'\mathbf{D})^{-1} (\mathbf{Z}'\mathbf{Y}) \\ &= (\tilde{\mathbf{Z}}'\tilde{\mathbf{D}})^{-1} (\tilde{\mathbf{Z}}'\mathbf{Y}) \quad \text{applying FWL to reduce } \mathbf{D} \text{ and } \mathbf{Z} \text{ to vectors.} \end{aligned}$$

Two-stage least squares

Q How do we get this magical expression? $\left(\hat{\beta}_1^{\text{IV}} = \frac{\hat{\pi}_1}{\hat{\gamma}_1} \right)$

Derivation

$$\begin{aligned}\hat{\beta}_1^{\text{IV}} &= (\mathbf{Z}'\mathbf{D})^{-1} (\mathbf{Z}'\mathbf{Y}) \\ &= (\tilde{\mathbf{Z}}'\tilde{\mathbf{D}})^{-1} (\tilde{\mathbf{Z}}'\mathbf{Y}) \quad \text{applying FWL to reduce } \mathbf{D} \text{ and } \mathbf{Z} \text{ to vectors.} \\ &= \frac{\text{Cov}(\tilde{\mathbf{Z}}_i, \mathbf{Y}_i)}{\text{Cov}(\tilde{\mathbf{Z}}_i, \tilde{\mathbf{D}}_i)}\end{aligned}$$

Two-stage least squares

Q How do we get this magical expression? $\left(\hat{\beta}_1^{\text{IV}} = \frac{\hat{\pi}_1}{\hat{\gamma}_1} \right)$

Derivation

$$\begin{aligned}\hat{\beta}_1^{\text{IV}} &= (\mathbf{Z}'\mathbf{D})^{-1} (\mathbf{Z}'\mathbf{Y}) \\ &= (\tilde{\mathbf{Z}}'\tilde{\mathbf{D}})^{-1} (\tilde{\mathbf{Z}}'\mathbf{Y}) \quad \text{applying FWL to reduce } \mathbf{D} \text{ and } \mathbf{Z} \text{ to vectors.} \\ &= \frac{\text{Cov}(\tilde{\mathbf{Z}}_i, \mathbf{Y}_i)}{\text{Cov}(\tilde{\mathbf{Z}}_i, \tilde{\mathbf{D}}_i)} = \frac{\text{Cov}(\tilde{\mathbf{Z}}_i, \mathbf{Y}_i) / \text{Var}(\tilde{\mathbf{Z}}_i)}{\text{Cov}(\tilde{\mathbf{Z}}_i, \tilde{\mathbf{D}}_i) / \text{Var}(\tilde{\mathbf{Z}}_i)}\end{aligned}$$

Two-stage least squares

Q How do we get this magical expression? $\left(\hat{\beta}_1^{\text{IV}} = \frac{\hat{\pi}_1}{\hat{\gamma}_1} \right)$

Derivation

$$\begin{aligned}\hat{\beta}_1^{\text{IV}} &= (\mathbf{Z}'\mathbf{D})^{-1} (\mathbf{Z}'\mathbf{Y}) \\ &= (\tilde{\mathbf{Z}}'\tilde{\mathbf{D}})^{-1} (\tilde{\mathbf{Z}}'\mathbf{Y}) \quad \text{applying FWL to reduce } \mathbf{D} \text{ and } \mathbf{Z} \text{ to vectors.} \\ &= \frac{\text{Cov}(\tilde{\mathbf{Z}}_i, \mathbf{Y}_i)}{\text{Cov}(\tilde{\mathbf{Z}}_i, \tilde{\mathbf{D}}_i)} = \frac{\text{Cov}(\tilde{\mathbf{Z}}_i, \mathbf{Y}_i) / \text{Var}(\tilde{\mathbf{Z}}_i)}{\text{Cov}(\tilde{\mathbf{Z}}_i, \tilde{\mathbf{D}}_i) / \text{Var}(\tilde{\mathbf{Z}}_i)} \\ &= \frac{\hat{\pi}_1}{\hat{\gamma}_1} \quad \checkmark\end{aligned}$$

Let's push a bit deeper into IV's mechanics and intuition.

IV: Mechanics and intuition

Setup

In this section, we'll use medical trials as a working example.[†]

[†] Credit/thanks go to [Michael Anderson](#) for this example—and much of these notes.

IV: Mechanics and intuition

Setup

In this section, we'll use medical trials as a working example.[†]

We are interested in the regression model for the effect of some treatment (e.g., blood-pressure medication) on medical outcome Y_i

[†] Credit/thanks go to [Michael Anderson](#) for this example—and much of these notes.

IV: Mechanics and intuition

Setup

In this section, we'll use medical trials as a working example.[†]

We are interested in the regression model for the effect of some treatment (e.g., blood-pressure medication) on medical outcome Y_i

$$Y_i = \beta_0 + \beta_1 D_i + \varepsilon_i$$

D_i indicates whether i takes the treatment (medication). ε_i captures all other factors that affect Y_i .

[†] Credit/thanks go to [Michael Anderson](#) for this example—and much of these notes.

IV: Mechanics and intuition

Setup

In this section, we'll use medical trials as a working example.[†]

We are interested in the regression model for the effect of some treatment (e.g., blood-pressure medication) on medical outcome Y_i

$$Y_i = \beta_0 + \beta_1 D_i + \varepsilon_i$$

D_i indicates whether i takes the treatment (medication). ε_i captures all other factors that affect Y_i . Or in potential-outcomes framework:

$$\begin{aligned} Y_i &= Y_{1i} D_i + Y_{0i} (1 - D_i) \\ Y_{0i} &= \beta_0 + \varepsilon_i \\ Y_{1i} &= Y_{0i} + \beta_1 \end{aligned}$$

[†] Credit/thanks go to [Michael Anderson](#) for this example—and much of these notes.

IV: Mechanics and intuition

Research design

Goal **Estimate the effect of blood-pressure medication** on blood pressure.

IV: Mechanics and intuition

Research design

Goal **Estimate the effect of blood-pressure medication** on blood pressure.

Challenge **Selection bias:** Even if treatment reduces blood pressure, selection bias will fight against the estimated effect.

IV: Mechanics and intuition

Research design

Goal **Estimate the effect of blood-pressure medication** on blood pressure.

Challenge **Selection bias:** Even if treatment reduces blood pressure, selection bias will fight against the estimated effect.

Solution **Randomized medical trial:** Ask randomly chosen individuals in treatment group to take the pill. Controls get placebo (or nothing).

IV: Mechanics and intuition

Research design

Goal **Estimate the effect of blood-pressure medication** on blood pressure.

Challenge **Selection bias:** Even if treatment reduces blood pressure, selection bias will fight against the estimated effect.

Solution **Randomized medical trial:** Ask randomly chosen individuals in treatment group to take the pill. Controls get placebo (or nothing).

Analysis 1 **Intention to treat (ITT):** $\hat{\beta}_1^{\text{ITT}} = \bar{Y}_{\text{Trt}} - \bar{Y}_{\text{Ctrl}}$

IV: Mechanics and intuition

Research design

Goal **Estimate the effect of blood-pressure medication** on blood pressure.

Challenge **Selection bias:** Even if treatment reduces blood pressure, selection bias will fight against the estimated effect.

Solution **Randomized medical trial:** Ask randomly chosen individuals in treatment group to take the pill. Controls get placebo (or nothing).

Analysis 1 **Intention to treat (ITT):** $\hat{\beta}_1^{\text{ITT}} = \bar{Y}_{\text{Trt}} - \bar{Y}_{\text{Ctrl}}$

ITT problem **Bias from noncompliance:** People don't always follow rules. E.g., treated folks who don't take pills; control folks who take pills.

IV: Mechanics and intuition

Research design

Goal **Estimate the effect of blood-pressure medication** on blood pressure.

Challenge **Selection bias:** Even if treatment reduces blood pressure, selection bias will fight against the estimated effect.

Solution **Randomized medical trial:** Ask randomly chosen individuals in treatment group to take the pill. Controls get placebo (or nothing).

Analysis 1 **Intention to treat (ITT):** $\hat{\beta}_1^{\text{ITT}} = \bar{Y}_{\text{Trt}} - \bar{Y}_{\text{Ctrl}}$

ITT problem **Bias from noncompliance:** People don't always follow rules. E.g., treated folks who don't take pills; control folks who take pills.

Analysis 2 **IV!**

IV: Mechanics and intuition

Research design

Goal **Estimate the effect of blood-pressure medication** on blood pressure.

Challenge **Selection bias:** Even if treatment reduces blood pressure, selection bias will fight against the estimated effect.

Solution **Randomized medical trial:** Ask randomly chosen individuals in treatment group to take the pill. Controls get placebo (or nothing).

Analysis 1 **Intention to treat (ITT):** $\hat{\beta}_1^{\text{ITT}} = \bar{Y}_{\text{Trt}} - \bar{Y}_{\text{Ctrl}}$

ITT problem **Bias from noncompliance:** People don't always follow rules. E.g., treated folks who don't take pills; control folks who take pills.

Analysis 2 **IV!** Instrument medication D_i with intention to treat Z_i .

IV: Mechanics and intuition

The IV solution

First question: Is Z_i a valid instrument for D_i ?

IV: Mechanics and intuition

The IV solution

First question: Is Z_i a valid instrument for D_i ?

1. $\text{Cov}(Z_i, \varepsilon_i) = 0$ as Z_i was randomly assigned (exclusion restriction).

IV: Mechanics and intuition

The IV solution

First question: Is Z_i a valid instrument for D_i ?

1. $\text{Cov}(Z_i, \varepsilon_i) = 0$ as Z_i was randomly assigned (exclusion restriction).
2. $\text{Cov}(Z_i, D_i) \neq 0$ if assignment to treatment changes the likelihood you take the pills (first stage).

IV: Mechanics and intuition

The IV solution

First question: Is \mathbf{Z}_i a valid instrument for \mathbf{D}_i ?

1. $\text{Cov}(\mathbf{Z}_i, \varepsilon_i) = 0$ as \mathbf{Z}_i was randomly assigned (exclusion restriction).
2. $\text{Cov}(\mathbf{Z}_i, \mathbf{D}_i) \neq 0$ if assignment to treatment changes the likelihood you take the pills (first stage).

$\therefore \mathbf{Z}_i$ is a valid instrument for \mathbf{D}_i and IV consistently estimates β_1 .

IV: Mechanics and intuition

Noncompliance

Noncompliant individuals do not abide by their treatment assignment.

IV: Mechanics and intuition

Noncompliance

Noncompliant individuals do not abide by their treatment assignment.

Let's see how IV "solves" this problems.

IV: Mechanics and intuition

Noncompliance

Noncompliant individuals do not abide by their treatment assignment.

Let's see how IV "solves" this problems.

First, assume noncompliance only affects treated individuals—*i.e.*, treated folks sometimes don't take their pills; control folks never take pills.

IV: Mechanics and intuition

Noncompliance, continued

The **first stage** recovers the share of treatment individuals who take the pill

$$D_i = \gamma_1 Z_i + u_i$$

IV: Mechanics and intuition

Noncompliance, continued

The **first stage** recovers the share of treatment individuals who take the pill

$$D_i = \gamma_1 Z_i + u_i$$

i.e., if 50% of treated individuals take the medication, $\hat{\gamma}_1 = 0.50$.

IV: Mechanics and intuition

Noncompliance, continued

The **first stage** recovers the share of treatment individuals who take the pill

$$D_i = \gamma_1 Z_i + u_i$$

i.e., if 50% of treated individuals take the medication, $\hat{\gamma}_1 = 0.50$.

The **reduced form** estimates the *ITT*

$$Y_i = \pi_1 Z_i + v_i$$

IV: Mechanics and intuition

Noncompliance, continued

The **first stage** recovers the share of treatment individuals who take the pill

$$D_i = \gamma_1 Z_i + u_i$$

i.e., if 50% of treated individuals take the medication, $\hat{\gamma}_1 = 0.50$.

The **reduced form** estimates the *ITT*

$$Y_i = \pi_1 Z_i + v_i$$

which we know IV rescales using the first stage

$$\hat{\beta}_1^{\text{IV}} = \frac{\hat{\pi}_1}{\hat{\gamma}_1} = \frac{\hat{\pi}_1}{0.50} = 2 \times \hat{\pi}_1$$

IV: Mechanics and intuition

Noncompliance, continued

IV solves the noncompliance issue by rescaling by the rate of compliance.

IV: Mechanics and intuition

Noncompliance, continued

IV solves the noncompliance issue by rescaling by the rate of compliance.

If everyone perfectly complies, then $\hat{\gamma}_1 = 1$ and $\hat{\beta}_1^{\text{IV}} = \hat{\pi}_1 / 1 = \hat{\beta}_1^{\text{ITT}}$.

IV: Mechanics and intuition

Noncompliance, continued

IV solves the noncompliance issue by rescaling by the rate of compliance.

If everyone perfectly complies, then $\hat{\gamma}_1 = 1$ and $\hat{\beta}_1^{\text{IV}} = \hat{\pi}_1/1 = \hat{\beta}_1^{\text{ITT}}$.

Further example $N_{\text{Trt}} = 10$; trt. compliance = 50%; ctrl. compliance = 100%.

$$\bar{Y}_{\text{Trt}} = \frac{5(\beta_0 + \beta_1) + 5(\beta_0)}{10} = \beta_0 + \frac{\beta_1}{2}$$

IV: Mechanics and intuition

Noncompliance, continued

IV solves the noncompliance issue by rescaling by the rate of compliance.

If everyone perfectly complies, then $\hat{\gamma}_1 = 1$ and $\hat{\beta}_1^{\text{IV}} = \hat{\pi}_1/1 = \hat{\beta}_1^{\text{ITT}}$.

Further example $N_{\text{Trt}} = 10$; trt. compliance = 50%; ctrl. compliance = 100%.

$$\bar{Y}_{\text{Trt}} = \frac{5(\beta_0 + \beta_1) + 5(\beta_0)}{10} = \beta_0 + \frac{\beta_1}{2} \text{ and } \bar{Y}_{\text{Ctrl}} = \beta_0.$$

IV: Mechanics and intuition

Noncompliance, continued

IV solves the noncompliance issue by rescaling by the rate of compliance.

If everyone perfectly complies, then $\hat{\gamma}_1 = 1$ and $\hat{\beta}_1^{\text{IV}} = \hat{\pi}_1/1 = \hat{\beta}_1^{\text{ITT}}$.

Further example $N_{\text{Trt}} = 10$; trt. compliance = 50%; ctrl. compliance = 100%.

$$\bar{Y}_{\text{Trt}} = \frac{5(\beta_0 + \beta_1) + 5(\beta_0)}{10} = \beta_0 + \frac{\beta_1}{2} \text{ and } \bar{Y}_{\text{Ctrl}} = \beta_0.$$

So our reduced-form estimate (the ITT) is $\hat{\gamma}_1 = \frac{\beta_1}{2}$ (half the true effect).

IV: Mechanics and intuition

Noncompliance, continued

IV solves the noncompliance issue by rescaling by the rate of compliance.

If everyone perfectly complies, then $\hat{\gamma}_1 = 1$ and $\hat{\beta}_1^{\text{IV}} = \hat{\pi}_1/1 = \hat{\beta}_1^{\text{ITT}}$.

Further example $N_{\text{Trt}} = 10$; trt. compliance = 50%; ctrl. compliance = 100%.

$$\bar{Y}_{\text{Trt}} = \frac{5(\beta_0 + \beta_1) + 5(\beta_0)}{10} = \beta_0 + \frac{\beta_1}{2} \text{ and } \bar{Y}_{\text{Ctrl}} = \beta_0.$$

So our reduced-form estimate (the ITT) is $\hat{\gamma}_1 = \frac{\beta_1}{2}$ (half the true effect).

IV consistently estimates β_1 via rescaling the ITT by the rate of compliance

$$\hat{\beta}_1^{\text{IV}} = \frac{\pi}{\gamma} = \frac{\beta_1/2}{1/2} = \beta_1$$

IV: Mechanics and intuition

Takeaways

Main points

1. IV **rescales** the causal effect of Z_i on Y_i by the causal effect of Z_i on D_i .

IV: Mechanics and intuition

Takeaways

Main points

1. IV **rescales** the causal effect of Z_i on Y_i by the causal effect of Z_i on D_i .
2. IV **does not** compare treated compliers to untreated compliers.

IV: Mechanics and intuition

Takeaways

Main points

1. IV **rescales** the causal effect of Z_i on Y_i by the causal effect of Z_i on D_i .
2. IV **does not** compare treated compliers to untreated compliers.
Such a comparison/estimator would re-introduce selection bias.

Thus far, we assumed homogeneous treatment effects.

Q What happens **when treatment effects are heterogeneous**?

A Let's recall what our instruments are doing (with Venn diagrams!).

Credit **Glen Waddell** introduced me to IV via Venn.

Figure 1

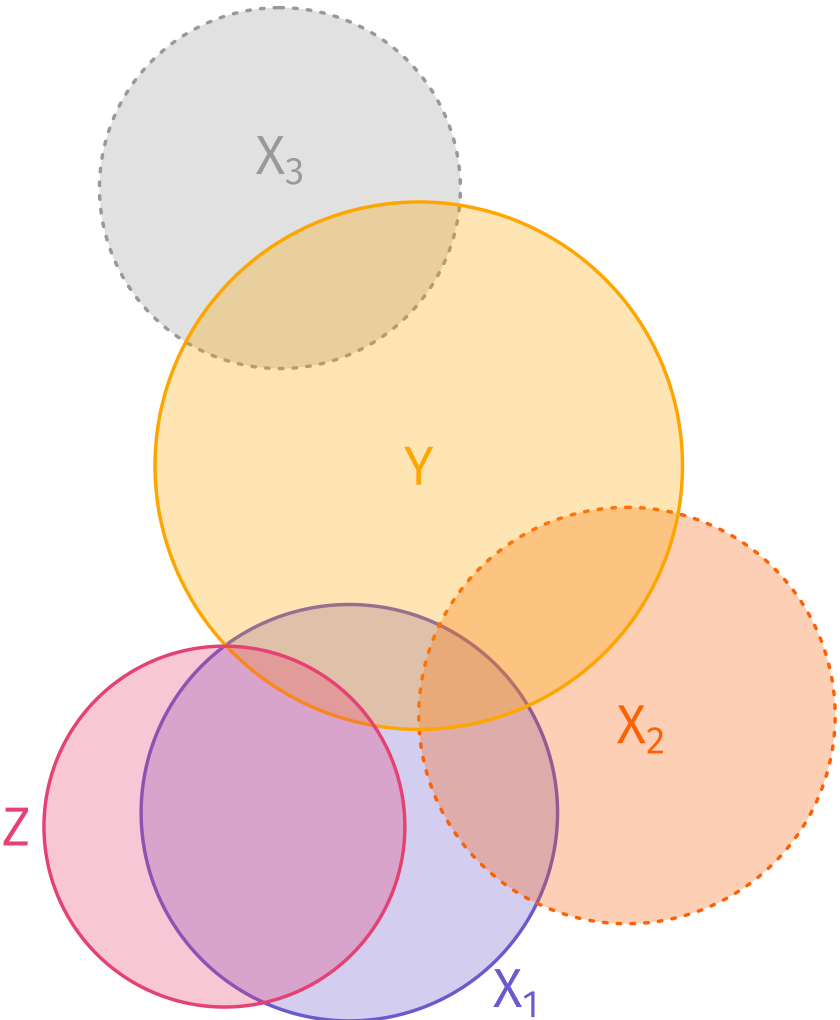


Figure 2

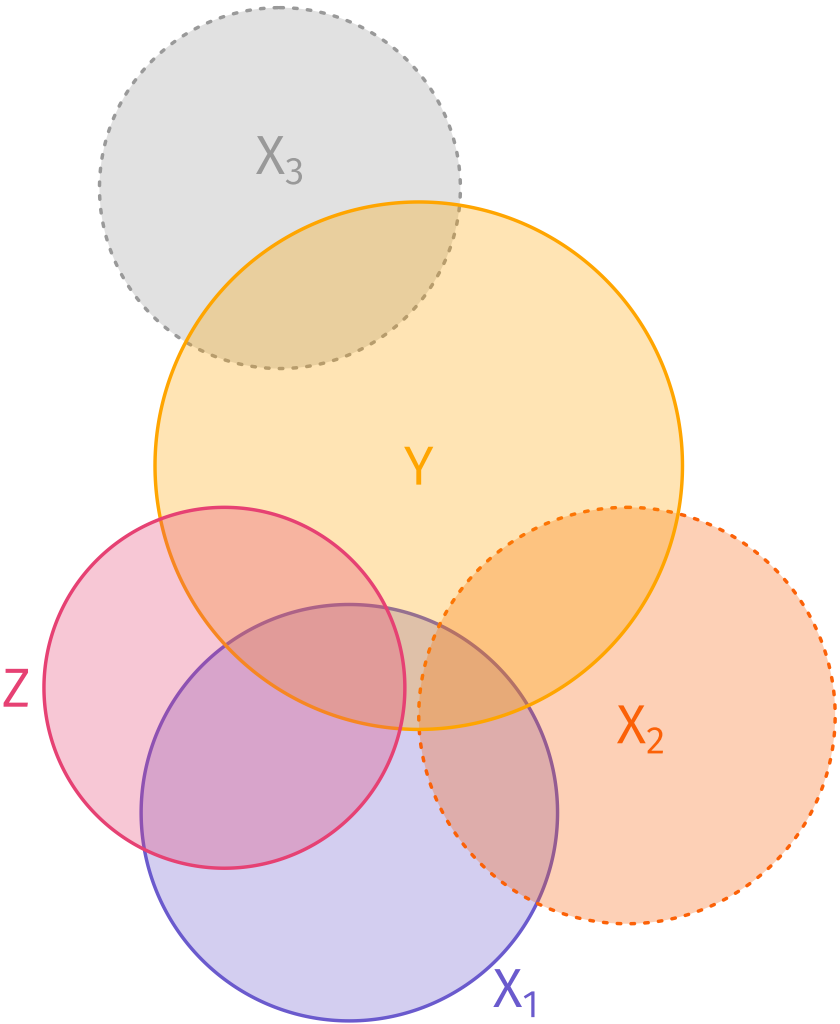


Figure 3

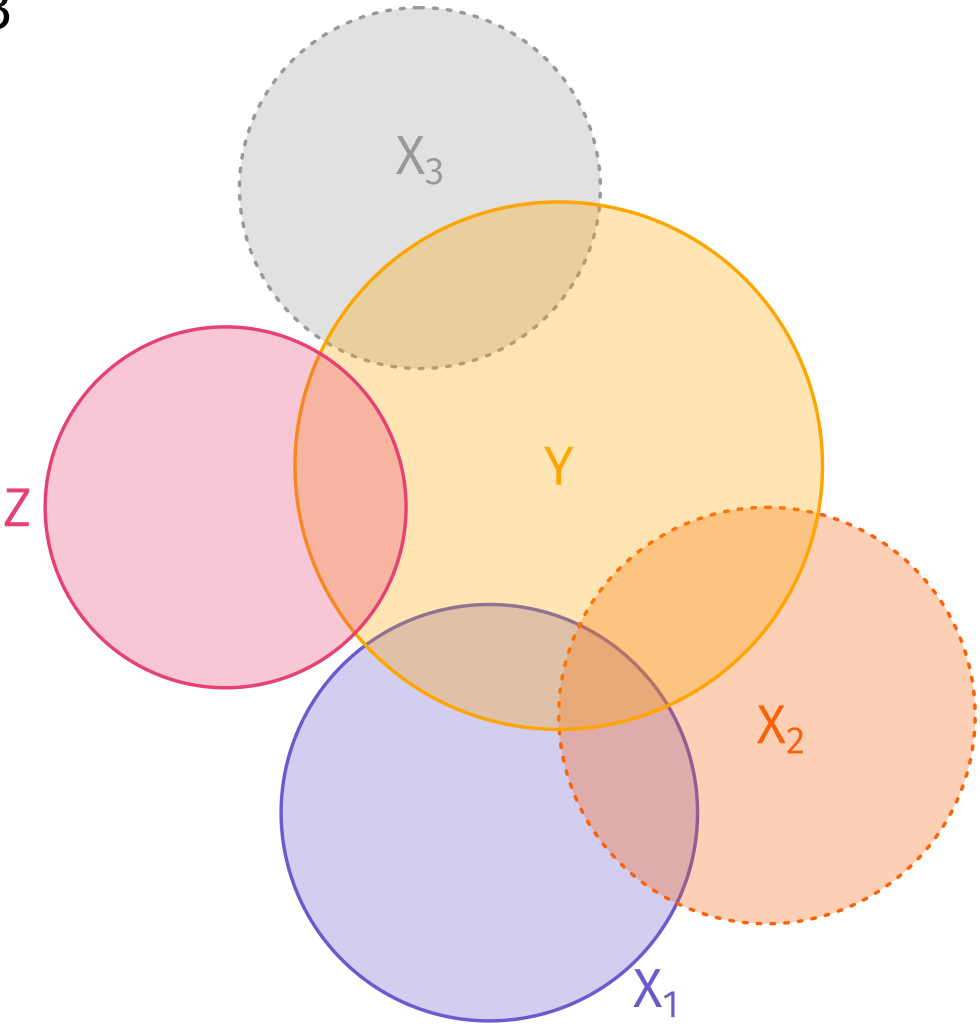


Figure 4

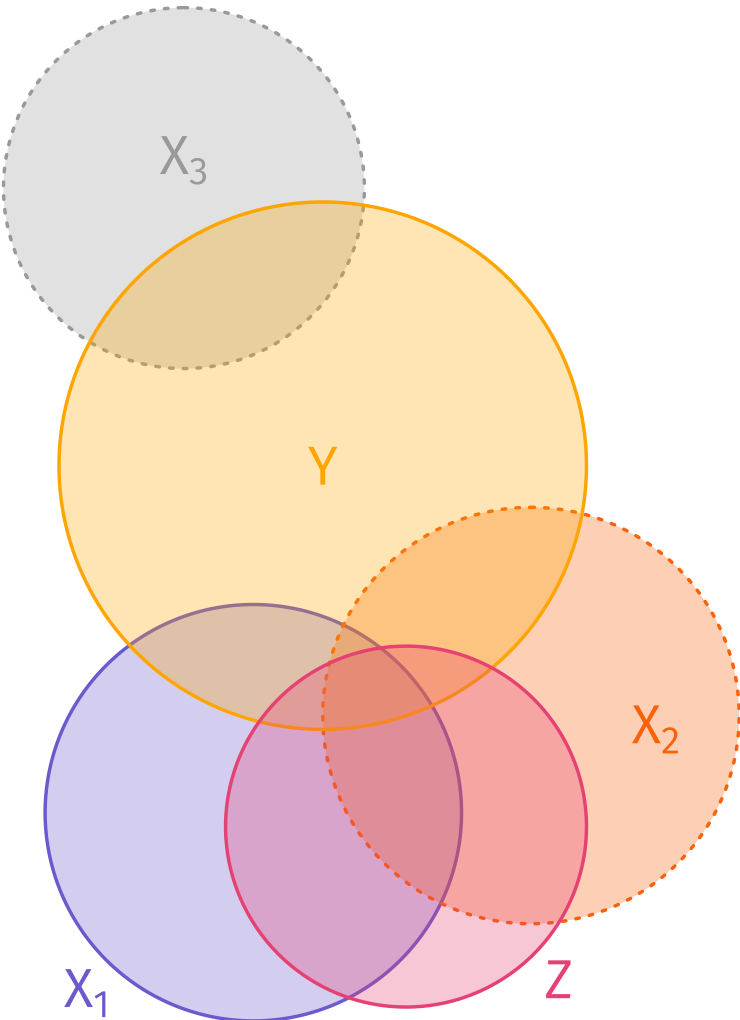
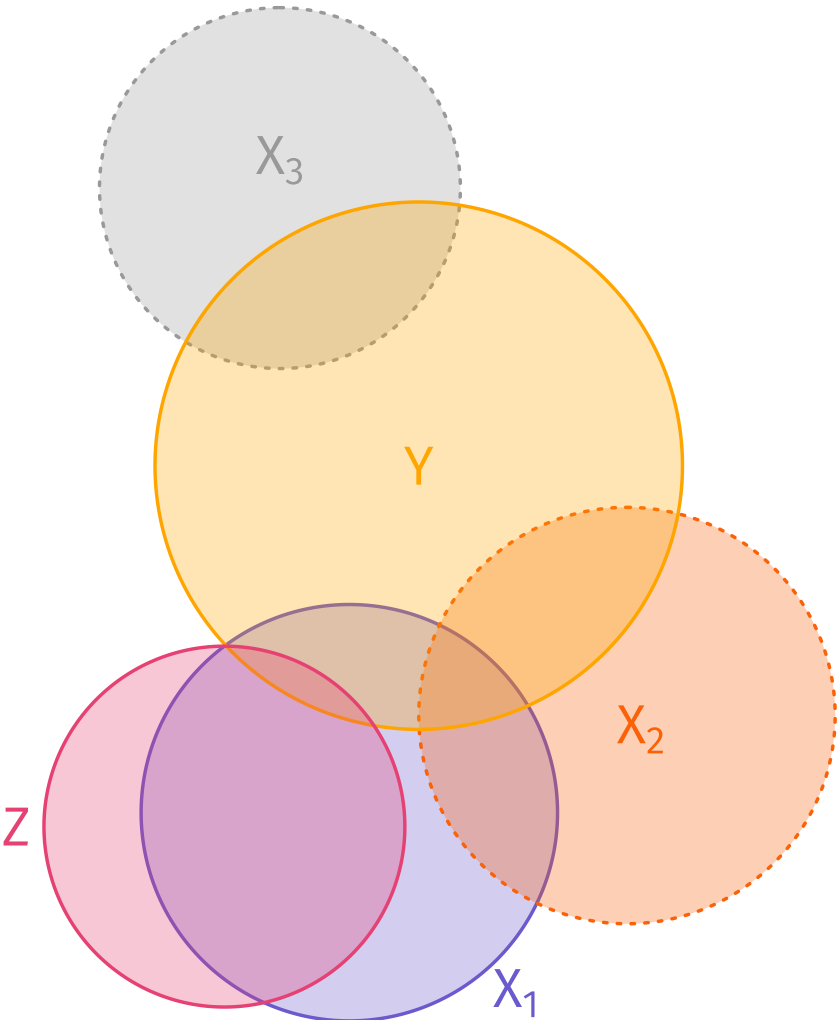


Figure 1



Can you draw the DAGs?

IV + heterogeneity

Recap

Throughout the course, we've discussed two concepts of treatment effects.

IV + heterogeneity

Recap

Throughout the course, we've discussed two concepts of treatment effects.

1. **Average treatment effect (ATE)** The average treatment effect for an individual randomly drawn from our sample.

IV + heterogeneity

Recap

Throughout the course, we've discussed two concepts of treatment effects.

1. **Average treatment effect (ATE)** The average treatment effect for an individual randomly drawn from our sample.
2. **Treatment on the treated (TOT)** The average treatment effect for a ***treated*** individual randomly drawn from our sample.

IV + heterogeneity

Recap

Throughout the course, we've discussed two concepts of treatment effects.

1. **Average treatment effect (ATE)** The average treatment effect for an individual randomly drawn from our sample.
2. **Treatment on the treated (TOT)** The average treatment effect for a ***treated*** individual randomly drawn from our sample.

When we assume homogeneous/constant treatment effects, $ATE = TOT$.

IV + heterogeneity

Recap

Throughout the course, we've discussed two concepts of treatment effects.

1. **Average treatment effect (ATE)** The average treatment effect for an individual randomly drawn from our sample.
2. **Treatment on the treated (TOT)** The average treatment effect for a ***treated*** individual randomly drawn from our sample.

When we assume homogeneous/constant treatment effects, $ATE = TOT$.

Q If treatment effects vary, then what do IV and 2SLS estimate?

IV + heterogeneity

Recap

Throughout the course, we've discussed two concepts of treatment effects.

1. **Average treatment effect (ATE)** The average treatment effect for an individual randomly drawn from our sample.
2. **Treatment on the treated (TOT)** The average treatment effect for a ***treated*** individual randomly drawn from our sample.

When we assume homogeneous/constant treatment effects, $ATE = TOT$.

Q If treatment effects vary, then what do IV and 2SLS estimate?

A Not ATE.

IV + heterogeneity

Recap

Throughout the course, we've discussed two concepts of treatment effects.

1. **Average treatment effect (ATE)** The average treatment effect for an individual randomly drawn from our sample.
2. **Treatment on the treated (TOT)** The average treatment effect for a ***treated*** individual randomly drawn from our sample.

When we assume homogeneous/constant treatment effects, $ATE = TOT$.

Q If treatment effects vary, then what do IV and 2SLS estimate?

A Not ATE. And not TOT.

IV + heterogeneity

Recap

Throughout the course, we've discussed two concepts of treatment effects.

1. **Average treatment effect (ATE)** The average treatment effect for an individual randomly drawn from our sample.
2. **Treatment on the treated (TOT)** The average treatment effect for a ***treated*** individual randomly drawn from our sample.

When we assume homogeneous/constant treatment effects, $ATE = TOT$.

Q If treatment effects vary, then what do IV and 2SLS estimate?

A Not ATE. And not TOT. They estimate the LATE.[†]

[†] See Angrist, Imbens, and Rubin (1996).

IV + heterogeneity

The LATE

IV generally estimates the **LATE**—the **Local Average Treatment Effect**.

IV + heterogeneity

The LATE

IV generally estimates the **LATE**—the **Local Average Treatment Effect**.

Recall IV "works" by isolating variation in \mathbf{D}_i induced by our instrument \mathbf{Z}_i .

IV + heterogeneity

The LATE

IV generally estimates the **LATE**—the **Local Average Treatment Effect**.

Recall IV "works" by isolating variation in \mathbf{D}_i induced by our instrument \mathbf{Z}_i .

In other words: IV focuses on the individuals whose \mathbf{D}_i changes due to \mathbf{Z}_i .

Angrist, Imbens, and Rubin (1996) call these folks **compliers**.

IV + heterogeneity

The LATE

IV generally estimates the **LATE**—the **Local Average Treatment Effect**.

Recall IV "works" by isolating variation in D_i induced by our instrument Z_i .

In other words: IV focuses on the individuals whose D_i changes due to Z_i .

Angrist, Imbens, and Rubin (1996) call these folks **compliers**.

However, *compliers* are only one of four possible groups.

1. **Compliers** $D_i = 1$ iff $Z_i = 1$.
2. **Always-takers** $D_i = 1 \forall Z_i$.
3. **Never-takers** $D_i = 0 \forall Z_i$.
4. **Defiers** $D_i = 1$ iff $Z_i = 0$.

IV + heterogeneity

The LATE

IV generally estimates the **LATE**—the **Local Average Treatment Effect**.

Recall IV "works" by isolating variation in D_i induced by our instrument Z_i .

In other words: IV focuses on the individuals whose D_i changes due to Z_i .

Angrist, Imbens, and Rubin (1996) call these folks **compliers**.

However, *compliers* are only one of four possible groups.

1. **Compliers** $D_i = 1$ iff $Z_i = 1$.

Only take pills **when treated**.

2. **Always-takers** $D_i = 1 \forall Z_i$.

Always take pills.

3. **Never-takers** $D_i = 0 \forall Z_i$.

Never take pills.

4. **Defiers** $D_i = 1$ iff $Z_i = 0$.

Only take pills **when untreated**.

IV + heterogeneity

The LATE

Because IV only uses variation in \mathbf{D}_i that correlates with \mathbf{Z}_i , IV mechanically drops *always-takers* and *never-takers*.

IV + heterogeneity

The LATE

Because IV only uses variation in \mathbf{D}_i that correlates with \mathbf{Z}_i , IV mechanically drops *always-takers* and *never-takers*.

Most IV derivations/applications assume away the existence of *defiers*.

IV + heterogeneity

The LATE

Because IV only uses variation in D_i that correlates with Z_i , IV mechanically drops *always-takers* and *never-takers*.

Most IV derivations/applications assume away the existence of *defiers*.

Thus, IV estimates a treatment effect **using only compliers**.

IV + heterogeneity

The LATE

Because IV only uses variation in D_i that correlates with Z_i , IV mechanically drops *always-takers* and *never-takers*.

Most IV derivations/applications assume away the existence of *defiers*.

Thus, IV estimates a treatment effect **using only compliers**.

Hence the "local" in *local average treatment effect*.

IV + heterogeneity

The LATE: Medical-trial example

Imagine treatment works for some ($\beta_{1,i} < 0$) and not for others ($\beta_{1,j} = 0$).

Suppose individuals know their response to blood-pressure medication.

IV + heterogeneity

The LATE: Medical-trial example

Imagine treatment works for some ($\beta_{1,i} < 0$) and not for others ($\beta_{1,j} = 0$).

Suppose individuals know their response to blood-pressure medication.

- $\beta_{1,i} < 0$ individuals always take the pill.

IV + heterogeneity

The LATE: Medical-trial example

Imagine treatment works for some ($\beta_{1,i} < 0$) and not for others ($\beta_{1,j} = 0$).

Suppose individuals know their response to blood-pressure medication.

- $\beta_{1,i} < 0$ individuals always take the pill.
- $\beta_{1,j} = 0$ individuals only take the pill when treated.

IV + heterogeneity

The LATE: Medical-trial example

Imagine treatment works for some ($\beta_{1,i} < 0$) and not for others ($\beta_{1,j} = 0$).

Suppose individuals know their response to blood-pressure medication.

- $\beta_{1,i} < 0$ individuals always take the pill.
- $\beta_{1,j} = 0$ individuals only take the pill when treated.

Then our compliers will be individuals for whom $\beta_{1,j} = 0$.

IV + heterogeneity

The LATE: Medical-trial example

Imagine treatment works for some ($\beta_{1,i} < 0$) and not for others ($\beta_{1,j} = 0$).

Suppose individuals know their response to blood-pressure medication.

- $\beta_{1,i} < 0$ individuals always take the pill.
- $\beta_{1,j} = 0$ individuals only take the pill when treated.

Then our compliers will be individuals for whom $\beta_{1,j} = 0$.

Thus, IV's LATE will indicate no treatment effect ($\widehat{\beta}_1^{\text{IV}} = 0$).

IV + heterogeneity

The LATE

Q So is IV actually inconsistent?

IV + heterogeneity

The LATE

Q So is IV actually inconsistent?

A It depends what you are trying to estimate (and how you interpret it).

IV doesn't estimate the ATE or TOT, so it would be inconsistent for them.[†]

[†] Just as the TOT is not consistent for the ATE.

IV + heterogeneity

The LATE

Q So is IV actually inconsistent?

A It depends what you are trying to estimate (and how you interpret it).

IV doesn't estimate the ATE or TOT, so it would be inconsistent for them.[†]

IV estimates the *local* average treatment effect.

[†] Just as the TOT is not consistent for the ATE.

IV + heterogeneity

The LATE

Q So is IV actually inconsistent?

A It depends what you are trying to estimate (and how you interpret it).

IV doesn't estimate the ATE or TOT, so it would be inconsistent for them.[†]

IV estimates the *local* average treatment effect.

Takeaway Because IV identifies off of compliers, it estimates an average treatment effect for these individuals (who *comply* with the instrument).

[†] Just as the TOT is not consistent for the ATE.

IV + heterogeneity

The LATE

Q So is IV actually inconsistent?

A It depends what you are trying to estimate (and how you interpret it).

IV doesn't estimate the ATE or TOT, so it would be inconsistent for them.[†]

IV estimates the *local* average treatment effect.

Takeaway Because IV identifies off of compliers, it estimates an average treatment effect for these individuals (who *comply* with the instrument).

*Takeaway*₂ Different instruments have different LATEs.

[†] Just as the TOT is not consistent for the ATE.

IV + heterogeneity

Monotonicity

We've already written down the two classical IV/2SLS assumptions

- *First stage*: $\text{Cov}(\mathbf{Z}_i, \mathbf{D}_i) > 0$
- *Exclusion restriction*: $\text{Cov}(\mathbf{Z}_i, \varepsilon_i) = 0$

but we need a third assumption to get ensure IV's complier-based LATE interpretation.

IV + heterogeneity

Monotonicity

We've already written down the two classical IV/2SLS assumptions

- *First stage*: $\text{Cov}(\mathbf{Z}_i, \mathbf{D}_i) > 0$
- *Exclusion restriction*: $\text{Cov}(\mathbf{Z}_i, \varepsilon_i) = 0$

but we need a third assumption to get ensure IV's complier-based LATE interpretation.

- **Monotonicity (Uniformity)**: $\mathbf{D}_i(z) \geq \mathbf{D}_i(z')$ or $\mathbf{D}_i(z) \leq \mathbf{D}_i(z') \quad \forall i$
Heckman: *Uniformity of responses across persons.*
Imbens and Angrist (1994): Instrument has monotone effect on \mathbf{D}_i .

IV + heterogeneity

Monotonicity

If "defiers" exist, then monotonicity/uniformity is violated.

IV + heterogeneity

Monotonicity

If "defiers" exist, then monotonicity/uniformity is violated.

In this case, the IV estimand is

$$\frac{\tau_c \Pr(\text{complier}) - \tau_d \Pr(\text{defier})}{\Pr(\text{complier}) - \Pr(\text{defier})}$$

which is not bound between τ_c and τ_d .

IV + heterogeneity

Monotonicity

If "defiers" exist, then monotonicity/uniformity is violated.

In this case, the IV estimand is

$$\frac{\tau_c \Pr(\text{complier}) - \tau_d \Pr(\text{defier})}{\Pr(\text{complier}) - \Pr(\text{defier})}$$

which is not bound between τ_c and τ_d .

Example $\tau_c = 1$ and $\tau_d = 2$. $\Pr(\text{complier}) = 2/3$ and $\Pr(\text{defier}) = 1/3$.

IV + heterogeneity

Monotonicity

If "defiers" exist, then monotonicity/uniformity is violated.

In this case, the IV estimand is

$$\frac{\tau_c \Pr(\text{complier}) - \tau_d \Pr(\text{defier})}{\Pr(\text{complier}) - \Pr(\text{defier})}$$

which is not bound between τ_c and τ_d .

Example $\tau_c = 1$ and $\tau_d = 2$. $\Pr(\text{complier}) = 2/3$ and $\Pr(\text{defier}) = 1/3$.

Then the "LATE" is 0.[†]

[†] Some people would instead say that there is no LATE when you violate monotonicity.

Until now, we've focused on using a single instrument.

The 2SLS estimator accomodates multiple instruments.[†]

[†] Whether you can find multiple valid instruments is another question.

Multiple instruments

Multiple instruments

Motivation

Q Why include multiple instruments?

Multiple instruments

Motivation

Q Why include multiple instruments?

A Multiple instruments can capture more variation in \mathbf{D}_i (efficiency).

Multiple instruments

Motivation

Q Why include multiple instruments?

A Multiple instruments can capture more variation in \mathbf{D}_i (efficiency).

Using terminology from the *system-of-equations* literature,

- one instrument for one endogenous variable: **just identified**
- multiple instruments for one endogenous variable: **over identified**

Multiple instruments

In practice

With (valid) instruments Z_{1i} and Z_{2i} , or first stage becomes

$$D_i = \gamma_0 + \gamma_1 Z_{1i} + \gamma_2 Z_{2i} + \gamma_3 X_i + u_i$$

Multiple instruments

In practice

With (valid) instruments \mathbf{Z}_{1i} and \mathbf{Z}_{2i} , or first stage becomes

$$\mathbf{D}_i = \gamma_0 + \gamma_1 \mathbf{Z}_{1i} + \gamma_2 \mathbf{Z}_{2i} + \gamma_3 \mathbf{X}_i + u_i$$

while our second stage is still

$$\mathbf{Y}_i = \beta_0 + \beta_1 \hat{\mathbf{D}}_i + \beta_2 \mathbf{X}_i + v_i$$

Multiple instruments

Example: Quarter of birth

Back to our quest to estimate the returns to education.

Multiple instruments

Example: Quarter of birth

Back to our quest to estimate the returns to education.

Angrist and Krueger (1991) proposed *quarter of birth* as a set of instruments for years of schooling.

Multiple instruments

Example: Quarter of birth

Back to our quest to estimate the returns to education.

Angrist and Krueger (1991) proposed *quarter of birth* as a set of instruments for years of schooling.

Accordingly, their first stage looks something like[†]

$$\begin{aligned}\text{Schooling}_i = & \gamma_0 + \gamma_1 \mathbb{I}(\text{Born Q1})_i + \gamma_2 \mathbb{I}(\text{Born Q2})_i \\ & + \gamma_3 \mathbb{I}(\text{Born Q3})_i + \gamma_4 \mathbb{I}(\text{Born Q4})_i \\ & + \gamma_5 \mathbf{X}_i + u_i\end{aligned}$$

[†] We need to drop one of the quarter-of-birth indicators to avoid perfect collinearity.

Multiple instruments

Example: Quarter of birth

Q Is quarter of birth a valid instrument?

Multiple instruments

Example: Quarter of birth

Q Is quarter of birth a valid instrument?

Q1 Why would quarter of birth affect schooling? (*First stage*)

Multiple instruments

Example: Quarter of birth

Q Is quarter of birth a valid instrument?

Q1 Why would quarter of birth affect schooling? (*First stage*)

A1 Students cannot drop out of school until a certain age, and quarter of birth affects your age at the time you begin school.

Multiple instruments

Example: Quarter of birth

Q Is quarter of birth a valid instrument?

Q1 Why would quarter of birth affect schooling? (*First stage*)

A1 Students cannot drop out of school until a certain age, and quarter of birth affects your age at the time you begin school.

Example Some states require students to stay in school until they are 16.

- Students who start school at age **6** drop out after **10** years of schooling.
- Students who start school at age **5** drop out after **11** years of schooling.

Multiple instruments

Example: Quarter of birth

If students must begin school in calendar year in which they turn 6

- December birthdates: begin school at 5.75; drop out with 10.25 yrs.
- January birthdates: begin school at 6.75; drop out with 9.25 yrs.

Multiple instruments

Example: Quarter of birth

If students must begin school in calendar year in which they turn 6

- December birthdates: begin school at 5.75; drop out with 10.25 yrs.
- January birthdates: begin school at 6.75; drop out with 9.25 yrs.

For some group, quarter of birth may affect the number of years in school.

Multiple instruments

Example: Quarter of birth

It turns out that the first stage is also pretty weak in this setting.

Weak instruments can cause several problems for 2SLS/IV:

Multiple instruments

Example: Quarter of birth

It turns out that the first stage is also pretty weak in this setting.

Weak instruments can cause several problems for 2SLS/IV:

1. Our estimator is a ratio of the reduced form and the first stage, so a weak first stage can blow up reduced-form estimates (amplifying reduced-form noise/bias).

Multiple instruments

Example: Quarter of birth

It turns out that the first stage is also pretty weak in this setting.

Weak instruments can cause several problems for 2SLS/IV:

1. Our estimator is a ratio of the reduced form and the first stage, so a weak first stage can blow up reduced-form estimates (amplifying reduced-form noise/bias).
2. Many weak instruments lead to a finite-sample issue in which 2SLS is biased toward OLS—our first stage is essentially overfitting.

Multiple instruments

Example: Quarter of birth

It turns out that the first stage is also pretty weak in this setting.

Weak instruments can cause several problems for 2SLS/IV:

1. Our estimator is a ratio of the reduced form and the first stage, so a weak first stage can blow up reduced-form estimates (amplifying reduced-form noise/bias).
2. Many weak instruments lead to a finite-sample issue in which 2SLS is biased toward OLS—our first stage is essentially overfitting.

What about our other requirements for a valid instrument?

Multiple instruments

Example: Quarter of birth

Q2 Is quarter of birth uncorrelated with ε_i (*excludable*)?

Multiple instruments

Example: Quarter of birth

Q2 Is quarter of birth uncorrelated with ε_i (*excludable*)?

A2 While quarter of birth may be fairly arbitrary for some families, other families might time births.

If these birth timers differ from other couples along other dimensions (*e.g.*, income or education), then quarter of birth may correlate with ε_i .

Multiple instruments

Example: Quarter of birth

Q3 Is the effect monotone?

Multiple instruments

Example: Quarter of birth

Q3 Is the effect monotone?

A3 Some[†] argue that monotonicity may be violated in this setting.

[†] E.g., Aliprantis (2012)

Multiple instruments

Example: Quarter of birth

Q3 Is the effect monotone?

A3 Some[†] argue that monotonicity may be violated in this setting.

Consider December births.

[†] *E.g.*, Aliprantis (2012)

Multiple instruments

Example: Quarter of birth

Q3 Is the effect monotone?

A3 Some[†] argue that monotonicity may be violated in this setting.

Consider December births.

- Original idea: December birthdates will start school at age 5.7, inducing more years of education before 16.

[†] E.g., Aliprantis (2012)

Multiple instruments

Example: Quarter of birth

Q3 Is the effect monotone?

A3 Some[†] argue that monotonicity may be violated in this setting.

Consider December births.

- Original idea: December birthdates will start school at age 5.7, inducing more years of education before 16.
- *Redshirting* idea: Parents hold back December kids so they can be older (*i.e.*, 6.7), inducing fewer years of education before 16.

[†] *E.g.*, Aliprantis (2012)

2SLS and R

feols

You can implement 2SLS/IV in many ways in R.

Today: `fixest` and `feols()`

There are others, e.g., `estimatr::iv_robust()` and `lfe::felm()`

2SLS and R

feols

You can implement 2SLS/IV in many ways in R.

Today: `fixest` and `feols()`

There are others, e.g., `estimatr::iv_robust()` and `lfe::felm()`

Specifically, `feols()` wants the exogenous "part" of the equation, a `I`, and the 'link' between the endogenous regressors and the instruments

2SLS and R

feols

You can implement 2SLS/IV in many ways in R.

Today: `fixest` and `feols()`

There are others, e.g., `estimatr::iv_robust()` and `lfe::felm()`

Specifically, `feols()` wants the exogenous "part" of the equation, a `|`, and the 'link' between the endogenous regressors and the instruments, e.g.,

```
# Estimate 2SLS  
feols(Y ~ 1 | D ~ Z, data = sample_df) %>% tidy()
```

```
#> # A tibble: 2 × 5  
#>   term          estimate std.error statistic  p.value  
#>   <chr>         <dbl>    <dbl>    <dbl>    <dbl>  
#> 1 (Intercept)     5.79     2.97     1.95 0.0546  
#> 2 fit_D           1.11     0.304    3.64 0.000437
```


2SLS and R

Now in two stages!

Of course, we can estimate 2SLS in two stages.

```
# First stage  
stage1 = feols(D ~ Z, data = sample_df)  
# First-stage results  
stage1 %>% tidy()
```

```
#> # A tibble: 2 × 5  
#>   term          estimate std.error statistic  p.value  
#>   <chr>          <dbl>     <dbl>     <dbl>   <dbl>  
#> 1 (Intercept)    8.82      0.317     27.8  2.49e-48  
#> 2 Z              0.326     0.103      3.16  2.11e- 3
```

2SLS and \mathbb{R}

Second stage

We just need to add \hat{D}_i to our dataset.

```
# Add fitted (first-stage) values to data
sample_df %<>% mutate(D_hat = stage1$fitted.values)
# Second stage
stage2 = feols(Y ~ D_hat, data = sample_df)
# Second-stage results
stage2 %>% tidy()
```

```
#> # A tibble: 2 × 5
#>   term          estimate std.error statistic p.value
#>   <chr>          <dbl>    <dbl>    <dbl>    <dbl>
#> 1 (Intercept)     5.79     5.41     1.07    0.288
#> 2 D_hat           1.11     0.554    2.00    0.0482
```

2SLS and \mathbb{R}

Standard errors

However, recall that our second-stage standard errors are not correct.

2SLS and R

Standard errors

However, recall that our second-stage standard errors are not correct.

Second-stage results

Term	Est.	S.E.	t stat.	p-Value
Int	5.786	5.413	1.07	0.2877
D hat	1.108	0.554	2.00	0.0482

2SLS and R

Standard errors

However, recall that our second-stage standard errors are not correct.

Second-stage results

Term	Est.	S.E.	t stat.	p-Value
Int	5.786	5.413	1.07	0.2877
D hat	1.108	0.554	2.00	0.0482

2SLS results

Term	Est.	S.E.	t stat.	p-Value
Int	5.786	2.974	1.95	0.0546
D	1.108	0.304	3.64	0.0004

IV and 2SLS

Conclusions

1. IV/2SLS focus on **isolating some "good" variation** in D_i via Z_i .
2. Important **requirements**: strong first stage, excludability, monotonicity.
3. IV and 2SLS **rescale the reduced form** with the first stage.
4. Estimates are **LATE from compliers**.
5. Different instruments can produce **different LATEs**.
6. A **weak first stage** can lead to problems.

Table of contents

Admin

1. Schedule

Instrumental variables

1. Research designs
2. Introduction
3. Definition
4. DAG
5. Example
6. IV estimator

Two-stage least squares

1. Setup
2. The reduced form
 - Defined
 - Intuition
 - Example
 - Derivation
3. Intuition and mechanics
 - Noncompliance
 - Rescaling
4. Heterogeneous treatment effects
 - Venn diagram
 - LATE
 - Example
 - Monotonicity
5. Multiple instruments
 - Example
6. 2SLS and \mathbb{R}
7. Conclusions