Instrumental Variables

EC 607, Set 9

Edward Rubin
Spring 2021

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.*,

$$(\mathrm{Y}_{0i}, \mathrm{Y}_{1i}) \perp\!\!\!\perp \mathrm{D}_i | \mathrm{X}_i$$

for **observable** variables X_i.

Selection on observables and/or unobservables

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

$$(\mathrm{Y}_{0i},\, \mathrm{Y}_{1i}) \perp\!\!\!\perp \mathrm{D}_i | \mathrm{X}_i$$

for **observable** variables X_i.

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

1. $(\mathrm{Y}_{0i},\, \mathrm{Y}_{1i}) \perp \mathrm{Z}_i$
2. $\mathrm{Cov}(\mathrm{Z}_i,\, \mathrm{D}_i) \neq 0$

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 $\mathbf{Y}_{0i}$ and $\mathbf{Y}_{1i}$).

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 $\mathbf{Y}_{0i}$ and $\mathbf{Y}_{1i}$).

(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 $\mathbf{Y}_{0i}$ and $\mathbf{Y}_{1i}$).

(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) $\mathbf{X}_i$.

Selection on observables and/or unobservables

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

(We want to avoid selection bias.)

- **Selection-on-observables designs** assume that we can control for all *bad variation* (selection) in $\mathbf{D}_i$ through a known (observed) $\mathbf{X}_i$.
- **Selection-on-unobservables designs** assume that we can extract **part of** the *good variation* in $\mathbf{D}_i$ (generally using some $\mathbf{Z}_i$) and then use this *good* part of $\mathbf{D}_i$ to estimate the effect of $\mathbf{D}_i$ on $\mathbf{Y}_i$.

Selection on observables and/or unobservables

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

(We want to avoid selection bias.)

- **Selection-on-observables designs** assume that we can control for all *bad variation* (selection) in $\mathbf{D}_i$ through a known (observed) $\mathbf{X}_i$.

- **Selection-on-unobservables designs** assume that we can extract **part of** the *good variation* in $\mathbf{D}_i$ (generally using some $\mathbf{Z}_i$) and then use this *good* part of $\mathbf{D}_i$ to estimate the effect of $\mathbf{D}_i$ on $\mathbf{Y}_i$. We throw away the rest of $\mathbf{D}_i$ (it includes *bad variation*).

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.*

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 $\mathbf{D}_i$.

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.

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 $\mathbf{Y}_i$.

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.

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)

$$\mathbf{Y}_i = \beta_0 + \beta_1 \mathbf{D}_i + \varepsilon_i \tag{1}$$

To guarantee consistent OLS estimates for β_1, want $\mathbf{Cov}(\mathbf{D}_i,\, \varepsilon_i) = \mathbf{0}$.
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.

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)

$$\mathbf{Y}_i = \beta_0 + \beta_1 \mathbf{D}_i + \varepsilon_i \tag{1}$$

To guarantee consistent OLS estimates for β_1, want $\mathbf{Cov}(\mathbf{D}_i,\, \varepsilon_i) = \mathbf{0}$.
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

$$\text{Y}_i = \beta_0 + \beta_1 \text{D}_i + \varepsilon_i \tag{1}$$

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

1. $\text{Cov}(\text{Z}_i,\, \text{D}_i) \neq 0$

Instrumental variables

Definition

For our model

$$\mathbf{Y}_i = \beta_0 + \beta_1 \mathbf{D}_i + \varepsilon_i \tag{1}$$

A valid **instrument** is a variable $\mathbf{Z}_i$ such that

1. $\mathbf{Cov}(\mathbf{Z}_i, \mathbf{D}_i) \neq 0$
   our instrument correlates with treatment

Definition

For our model

$$\mathbf{Y}_i = \beta_0 + \beta_1 \mathbf{D}_i + \varepsilon_i \tag{1}$$

A valid **instrument** is a variable $\mathbf{Z}_i$ such that

1. $\mathbf{Cov}(\mathbf{Z}_i, \mathbf{D}_i) \neq \mathbf{0}$
   our instrument correlates with treatment (so we can keep part of $\mathbf{D}_i$)

Instrumental variables

Definition

For our model

$$\mathbf{Y}_i = \beta_0 + \beta_1 \mathbf{D}_i + \varepsilon_i \tag{1}$$

A valid **instrument** is a variable $\mathbf{Z}_i$ such that

1. $\mathbf{Cov}(\mathbf{Z}_i,\, \mathbf{D}_i) \neq \mathbf{0}$
   our instrument correlates with treatment (so we can keep part of $\mathbf{D}_i$)

2. $\mathbf{Cov}(\mathbf{Z}_i,\, \varepsilon_i) = \mathbf{0}$

Instrumental variables

Definition

For our model

$$\mathbf{Y}_i = \beta_0 + \beta_1 \mathbf{D}_i + \varepsilon_i \tag{1}$$

A valid **instrument** is a variable $\mathbf{Z}_i$ such that

1. $\mathbf{Cov}(\mathbf{Z}_i,\ \mathbf{D}_i) \neq \mathbf{0}$
   our instrument correlates with treatment (so we can keep part of $\mathbf{D}_i$)

2. $\mathbf{Cov}(\mathbf{Z}_i,\ \varepsilon_i) = \mathbf{0}$
   our instrument is uncorrelated with other (non-$\mathbf{D}_i$) determinants of $\mathbf{Y}_i$

Instrumental variables

Definition

For our model

$$\mathbf{Y}_i = \beta_0 + \beta_1 \mathbf{D}_i + \varepsilon_i \tag{1}$$

A valid **instrument** is a variable $\mathbf{Z}_i$ such that

1. $\mathbf{Cov}(\mathbf{Z}_i,\, \mathbf{D}_i) \neq 0$
   our instrument correlates with treatment (so we can keep part of $\mathbf{D}_i$)

2. $\mathbf{Cov}(\mathbf{Z}_i,\, \varepsilon_i) = 0$
   our instrument is uncorrelated with other (non-$\mathbf{D}_i$) determinants of $\mathbf{Y}_i$,
   i.e., $\mathbf{Z}_i$ is excludable from equation (1).

Instrumental variables

Definition

For our model

$$\text{Y}_i = \beta_0 + \beta_1 \text{D}_i + \varepsilon_i \tag{1}$$

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

1. $\text{Cov}(\text{Z}_i,\, \text{D}_i) \neq 0$
   our instrument correlates with treatment (so we can keep part of D_i)

2. $\text{Cov}(\text{Z}_i,\, \varepsilon_i) = 0$
   our instrument is uncorrelated with other (non-D_i) determinants of Y_i, *i.e.*, Z_i is excludable from equation (1). **(exclusion restriction)**

Example

Back to the returns to a college degree,

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

OLS is likely biased.

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.

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.

The IV estimator

The IV estimator for our model

$$\mathrm{Y}_i = \beta_0 + \beta_1 \mathrm{D}_i + \varepsilon_i \tag{1}$$

with (valid) instrument Z_i is

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

Instrument variables

The IV estimator

The IV estimator for our model

$$\mathrm{Y}_i = \beta_0 + \beta_1 \mathrm{D}_i + \varepsilon_i \tag{1}$$

with (valid) instrument Z_i is

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

If you have no covariates, then

$$\hat{\beta}_{\mathrm{IV}} = \frac{\mathrm{Cov}(\mathrm{Z}_i,\ \mathrm{Y}_i)}{\mathrm{Cov}(\mathrm{Z}_i,\ \mathrm{D}_i)}$$

The IV estimator

The IV estimator for our model

$$\mathrm{Y}_i = \beta_0 + \beta_1 \mathrm{D}_i + \varepsilon_i \tag{1}$$

with (valid) instrument Z_i is

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

If you have additional (exogenous) covariates $\mathbf{X}_i$, then

$$\mathbf{Z} = \begin{bmatrix} \mathrm{Z}_i & \mathbf{X}_i \end{bmatrix}$$

$$\mathbf{D} = \begin{bmatrix} \mathrm{D}_i & \mathbf{X}_i \end{bmatrix}$$

Proof: Consistency

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

$$\mathbf{Y}_i = \beta_0 + \beta_1 \mathbf{X}_i + \varepsilon_i \tag{1}$$

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

Proof: Consistency

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

$$\mathrm{Y}_i = \beta_0 + \beta_1 \mathrm{X}_i + \varepsilon_i \tag{1}$$

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

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

Instrumental variables

Proof: Consistency

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

$$\text{Y}_i = \beta_0 + \beta_1 \text{X}_i + \varepsilon_i \tag{1}$$

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

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

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

Proof: Consistency

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

$$\mathrm{Y}_i = \beta_0 + \beta_1 \mathrm{X}_i + \varepsilon_i \tag{1}$$

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

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

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

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

Instrumental variables

Proof: Consistency

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

$$\mathbf{Y}_i = \beta_0 + \beta_1 \mathbf{X}_i + \varepsilon_i \tag{1}$$

$$\begin{aligned}
&\operatorname{plim}\left(\hat{\beta}_{IV}\right) \\
&\quad = \operatorname{plim}\left(\left(\mathrm{Z}'\mathrm{D}\right)^{-1}\left(\mathrm{Z}'\mathrm{Y}\right)\right) \\
&\quad = \operatorname{plim}\left(\left(\mathrm{Z}'\mathrm{D}\right)^{-1}\left(\mathrm{Z}'\mathrm{D}\beta + \mathrm{Z}'\varepsilon\right)\right) \\
&\quad = \operatorname{plim}\left(\left(\mathrm{Z}'\mathrm{D}\right)^{-1}\left(\mathrm{Z}'\mathrm{D}\right)\beta\right) + \operatorname{plim}\left(\frac{1}{N}\mathrm{Z}'\mathrm{D}\right)^{-1}\operatorname{plim}\left(\frac{1}{N}\mathrm{Z}'\varepsilon\right) \\
&\quad = \beta \quad ✔
\end{aligned}$$

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 Z_i on our endogenous variable D_i and (predetermined) covariates X_i. Save $\widehat{\mathrm{D}}_i$.

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

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 Z_i on our endogenous variable $\mathbf{D}_i$ and (predetermined) covariates $\mathbf{X}_i$. Save $\widehat{\mathrm{D}}_i$.

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

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

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

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.

$$\hat{\beta}_{2SLS} = \left(\mathrm{D}'\mathrm{P_Z}\mathrm{D}\right)^{-1}\left(\mathrm{D}'\mathrm{P_Z}\mathrm{Y}\right)$$

$$\mathrm{P_Z} = \mathrm{Z}\left(\mathrm{Z}'\mathrm{Z}\right)^{-1}\mathrm{Z}'$$

where $\mathbf{D}$ is a matrix of our treatment and predetermined covariates ($\mathbf{X}_i$) and 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. $\text{D}_i = \gamma_1 \text{Z}_i + \gamma_2 \text{X}_i + u_i$
2. $\text{Y}_i = \beta_1 \widehat{\text{D}}_i + \beta_2 \text{X}_i + \varepsilon_i$

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. $\text{D}_i = \gamma_1 \text{Z}_i + \gamma_2 \text{X}_i + u_i$
2. $\text{Y}_i = \beta_1 \widehat{\text{D}}_i + \beta_2 \text{X}_i + \varepsilon_i$

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).

$$\text{Y}_i = \pi_1 \text{Z}_i + \pi_2 \text{X}_i + u_i$$

The reduced form

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

1. $\mathbf{D}_i = \gamma_1 \mathrm{Z}_i + \gamma_2 \mathbf{X}_i + u_i$
2. $\mathbf{Y}_i = \beta_1 \widehat{\mathrm{D}}_i + \beta_2 \mathbf{X}_i + \varepsilon_i$

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

The **reduced form** regresses the outcome $\mathbf{Y}_i$ (LHS of the second stage) on our instrument Z_i and covariates $\mathbf{X}_i$ (RHS of the first stage).

$$\mathbf{Y}_i = \pi_1 \mathrm{Z}_i + \pi_2 \mathbf{X}_i + u_i$$

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* (D_i).

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.

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 $\mathbf{Cov}(\mathbf{Z}_i,\, \varepsilon_i) = \mathbf{0}$.

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 $\mathbf{Cov}(\mathbf{Z}_i,\, \varepsilon_i) = \mathbf{0}$.
- 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 $\mathbf{Cov}(\mathbf{Z}_i,\, \varepsilon_i) = \mathbf{0}$.
- Offers insights into your estimates

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

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.

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

The reduced form, intuition

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

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

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.

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

What's the interpretation/intuition?

Back to our example: $\widehat{\beta}_1 =$ 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.

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

What's the interpretation/intuition?

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

$\widehat{\pi}_1$ 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.

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

What's the interpretation/intuition?

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

$\widehat{\pi}_1$ 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.

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

What's the interpretation/intuition?

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

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

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

The reduced form, intuition

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

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

What's the interpretation/intuition?

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

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

$\widehat{\gamma}_1$ 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.

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

What's the interpretation/intuition?

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

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

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

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.*, $\widehat{\gamma}_1 = 0.50$.†

† 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.*, $\widehat{\gamma}_1 = 0.50$.†

Our reduced-form estimate of $\widehat{\pi}_1 = \$5{,}000$ 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.*, $\widehat{\gamma}_1 = 0.50$.[†]

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

However, half of the winners did not graduate, so $\widehat{\pi}_1$ "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.*, $\widehat{\gamma}_1 = 0.50$.†

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

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

Thus, we want to double $\widehat{\pi}_1$, *i.e.*, divide by $\widehat{\gamma}_1$: $\widehat{\pi}_1/\widehat{\gamma}_1$ = \$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? $\left(\widehat{\beta}_1^{\text{IV}} = \dfrac{\widehat{\pi}_1}{\widehat{\gamma}_1}\right)$

Two-stage least squares

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

Derivation

Two-stage least squares

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

Derivation

$$\widehat{\beta}_1^{\text{IV}} = \left(\mathrm{Z}'\mathrm{D}\right)^{-1}\left(\mathrm{Z}'\mathrm{Y}\right)$$

Two-stage least squares

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

Derivation

$$
\begin{aligned}
\widehat{\beta}_1^{\text{IV}} &= \left(\text{Z}'\text{D}\right)^{-1}\left(\text{Z}'\text{Y}\right) \\
&= \left(\widetilde{\text{Z}}'\widetilde{\text{D}}\right)^{-1}\left(\widetilde{\text{Z}}'\text{Y}\right) \qquad \text{applying FWL to reduce } \mathbf{D} \text{ and } \mathbf{Z} \text{ to vectors.}
\end{aligned}
$$

Two-stage least squares

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

Derivation

$$
\begin{aligned}
\widehat{\beta}_1^{\text{IV}} &= \left(\mathrm{Z}'\mathrm{D}\right)^{-1}\left(\mathrm{Z}'\mathrm{Y}\right) \\
&= \left(\widetilde{\mathrm{Z}}'\widetilde{\mathrm{D}}\right)^{-1}\left(\widetilde{\mathrm{Z}}'\mathrm{Y}\right) \qquad \text{applying FWL to reduce } \mathbf{D} \text{ and } \mathbf{Z} \text{ to vectors.} \\
&= \frac{\mathrm{Cov}\left(\widetilde{\mathrm{Z}}_i,\ \mathrm{Y}_i\right)}{\mathrm{Cov}\left(\widetilde{\mathrm{Z}}_i,\ \widetilde{\mathrm{D}}_i\right)}
\end{aligned}
$$

Two-stage least squares

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

Derivation

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

Two-stage least squares

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

Derivation

$$\widehat{\beta}_1^{\text{IV}} = \left(\text{Z}'\text{D}\right)^{-1}\left(\text{Z}'\text{Y}\right)$$

$$= \left(\widetilde{\text{Z}}'\widetilde{\text{D}}\right)^{-1}\left(\widetilde{\text{Z}}'\text{Y}\right) \qquad \text{applying FWL to reduce } \mathbf{D} \text{ and } \mathbf{Z} \text{ to vectors.}$$

$$= \frac{\text{Cov}\left(\widetilde{\text{Z}}_i,\ \text{Y}_i\right)}{\text{Cov}\left(\widetilde{\text{Z}}_i,\ \widetilde{\text{D}}_i\right)} = \frac{\text{Cov}\left(\widetilde{\text{Z}}_i,\ \text{Y}_i\right)/\ \text{Var}\left(\widetilde{\text{Z}}_i\right)}{\text{Cov}\left(\widetilde{\text{Z}}_i,\ \widetilde{\text{D}}_i\right)/\ \text{Var}\left(\widetilde{\text{Z}}_i\right)}$$

$$= \frac{\widehat{\pi}_1}{\widehat{\gamma}_1} \quad ✔$$

Let's push a bit deeper into IV's 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 $\mathbf{Y}_i$

† 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 $\mathbf{Y}_i$

$$\mathbf{Y}_i = \beta_0 + \beta_1 \mathbf{D}_i + \varepsilon_i$$

$\mathbf{D}_i$ indicates whether $\boldsymbol{i}$ *takes* the treatment (medication). ε_i captures all other factors that affect $\mathbf{Y}_i$.

† Credit/thanks go to Michael Anderson for this example—and much of these notes.

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 $\mathbf{Y}_i$

$$\mathbf{Y}_i = \beta_0 + \beta_1 \mathbf{D}_i + \varepsilon_i$$

$\mathbf{D}_i$ indicates whether i *takes* the treatment (medication). ε_i captures all other factors that affect $\mathbf{Y}_i$. Or in potential-outcomes framework:

$$\begin{aligned} \mathbf{Y}_i &= \mathbf{Y}_{1i}\mathbf{D}_i + \mathbf{Y}_{0i}(1 - \mathbf{D}_i) \\ \mathbf{Y}_{0i} &= \beta_0 + \varepsilon_i \\ \mathbf{Y}_{1i} &= \mathbf{Y}_{0i} + \beta_1 \end{aligned}$$

† 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.

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 fights against the estimated effect.

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 fights 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 fights 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): $\widehat{\beta}_1^{\text{ITT}} = \overline{\text{Y}}_{\text{Trt}} - \overline{\text{Y}}_{\text{Ctrl}}$

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 fights 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): $\widehat{\beta}_1^{\text{ITT}} = \overline{\text{Y}}_{\text{Trt}} - \overline{\text{Y}}_{\text{Ctrl}}$

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 fights 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): $\widehat{\beta}_1^{\text{ITT}} = \overline{\text{Y}}_{\text{Trt}} - \overline{\text{Y}}_{\text{Ctrl}}$

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 fights 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): $\widehat{\beta}_1^{\text{ITT}} = \overline{\text{Y}}_{\text{Trt}} - \overline{\text{Y}}_{\text{Ctrl}}$

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 $\mathbf{D}_i$ with intention to treat $\mathbf{Z}_i$.

The IV solution

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

The IV solution

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

1. $\mathbf{Cov}(\mathbf{Z}_i,\, \varepsilon_i) = \mathbf{0}$ as $\mathbf{Z}_i$ was randomly assigned (exclusion restriction).

IV: Mechanics and intuition

The IV solution

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

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

The IV solution

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

1. $\mathbf{Cov}(\mathbf{Z}_i,\, \varepsilon_i) = \mathbf{0}$ as $\mathbf{Z}_i$ was randomly assigned (exclusion restriction).
2. $\mathbf{Cov}(\mathbf{Z}_i,\, \mathbf{D}_i) \neq \mathbf{0}$ 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.

Noncompliance

Noncompliant individuals do not abide by their treatment assignment.

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

$$\text{D}_i = \gamma_1 \text{Z}_i + u_i$$

Noncompliance, continued

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

$$\mathrm{D}_i = \gamma_1 \mathrm{Z}_i + u_i$$

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

Noncompliance, continued

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

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

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

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

$$\mathbf{Y}_i = \pi_1 \mathbf{Z}_i + v_i$$

IV: Mechanics and intuition

Noncompliance, continued

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

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

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

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

$$\mathbf{Y}_i = \pi_1 \mathbf{Z}_i + v_i$$

which we know IV rescales using the first stage

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

Noncompliance, continued

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

Noncompliance, continued

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

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

Noncompliance, continued

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

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

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

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

Noncompliance, continued

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

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

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

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

IV: Mechanics and intuition

Noncompliance, continued

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

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

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

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

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

Noncompliance, continued

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

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

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

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

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

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

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

IV: Mechanics and intuition

Takeaways

Main points

1. IV **rescales** the causal effect of $\mathbf{Z}_i$ on $\mathbf{Y}_i$ by the causal effect of $\mathbf{Z}_i$ on $\mathbf{D}_i$.

IV: Mechanics and intuition

Takeaways

Main points

1. IV **rescales** the causal effect of $\mathbf{Z}_i$ on $\mathbf{Y}_i$ by the causal effect of $\mathbf{Z}_i$ on $\mathbf{D}_i$.

2. IV **does not** compare treated compliers to untreated compliers.

IV: Mechanics and intuition

Takeaways

Main points

1. IV **rescales** the causal effect of $\mathbf{Z}_i$ on $\mathbf{Y}_i$ by the causal effect of $\mathbf{Z}_i$ on $\mathbf{D}_i$.

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

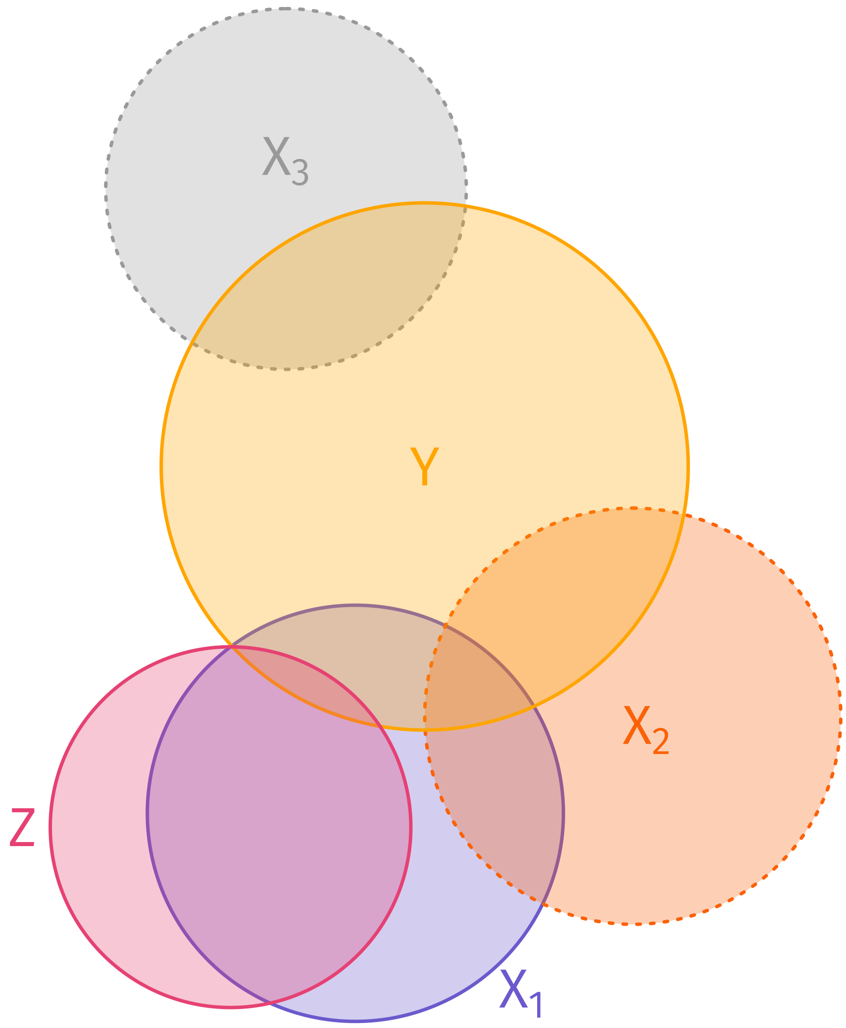


Figure 2

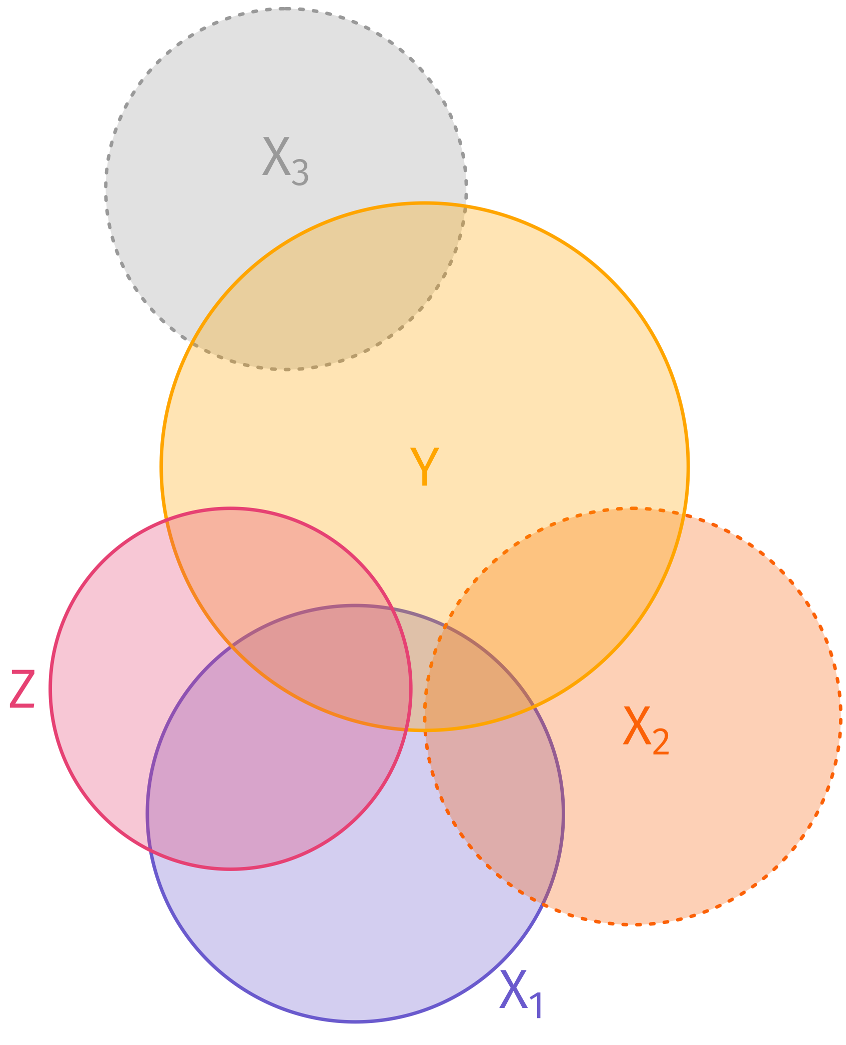


Figure 3

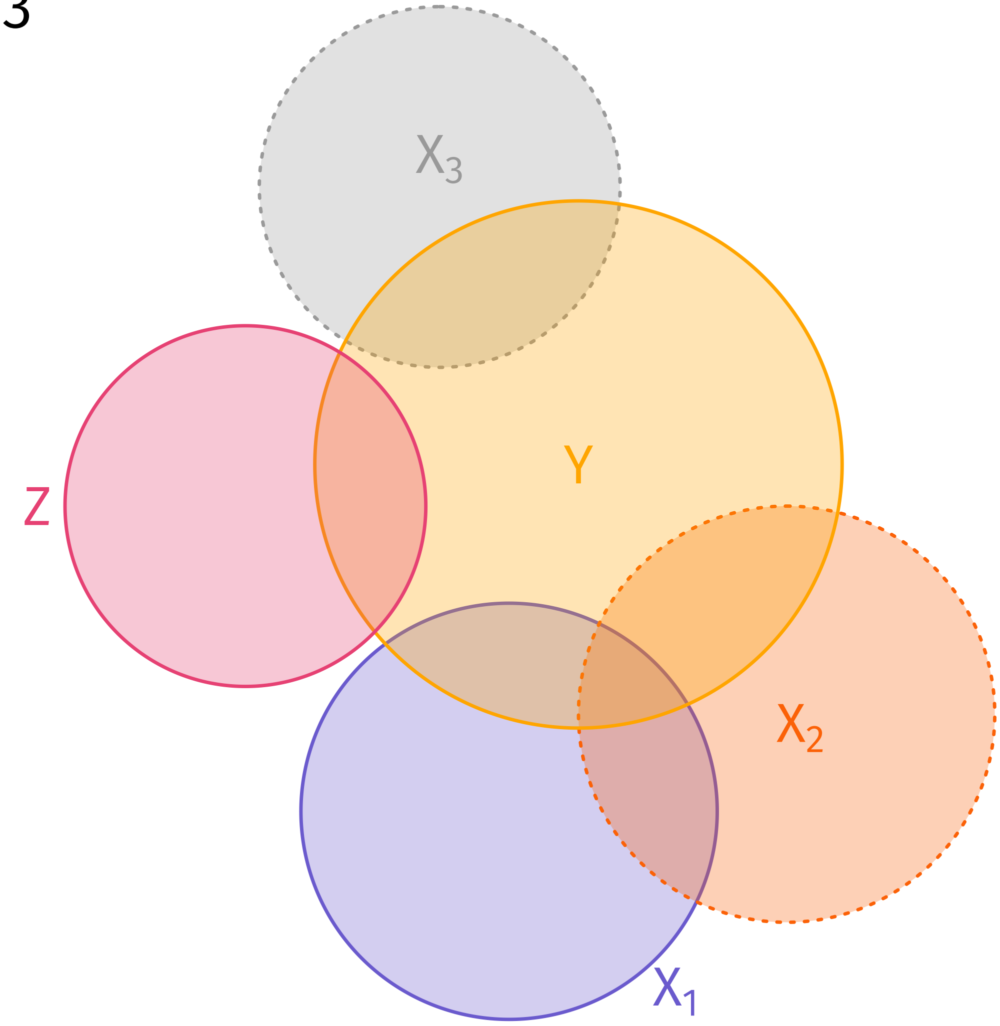


Figure 4

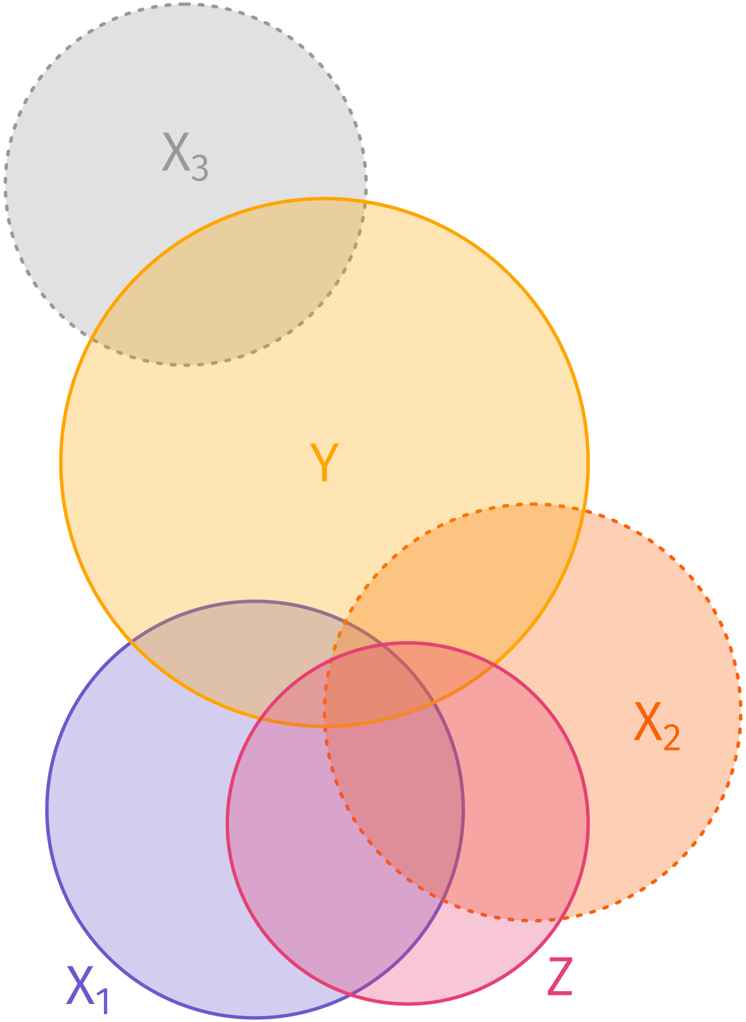
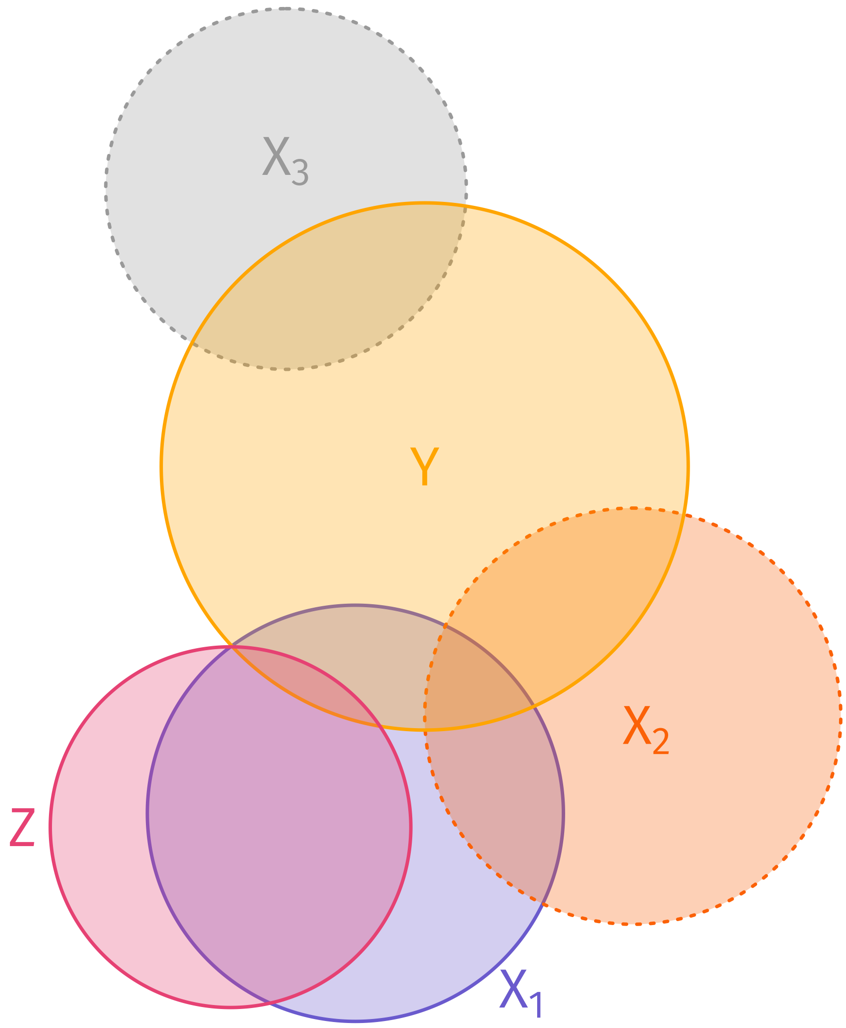


Figure 1



Can you draw the DAGs?

IV + heterogeneity

Recap

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

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.

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**.

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.

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**.

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

1. **Compliers** $\mathbf{D}_i = 1$ iff $\mathbf{Z}_i = 1$.
2. **Always-takers** $\mathbf{D}_i = 1 \ \forall \mathbf{Z}_i$.
3. **Never-takers** $\mathbf{D}_i = 0 \ \forall \mathbf{Z}_i$.
4. **Defiers** $\mathbf{D}_i = 1$ iff $\mathbf{Z}_i = 0$.

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.

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

1. Compliers $\mathbf{D}_i = 1$ iff $\mathbf{Z}_i = 1$. Only take pills **when treated**.
2. Always-takers $\mathbf{D}_i = 1\ \forall \mathbf{Z}_i$. **Always** take pills.
3. Never-takers $\mathbf{D}_i = 0\ \forall \mathbf{Z}_i$. **Never** take pills.
4. Defiers $\mathbf{D}_i = 1$ iff $\mathbf{Z}_i = 0$. 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*.

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*.

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

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*.

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 $\left(\beta_{1,j} = 0\right)$.

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 $\left(\beta_{1,j} = 0\right)$.

Suppose individuals know their response to blood-pressure medication.

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

The LATE: Medical-trial example

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

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 $\left(\beta_{1,j} = 0\right)$.

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 $\beta_{1,j} = 0$.

The LATE: Medical-trial example

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

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 $\left(\widehat{\beta}_1^{\text{IV}} = 0\right)$.

IV + heterogeneity

The LATE

Q So is IV actually inconsistent?

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.

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.

Monotonicity

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

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

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

Monotonicity

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

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

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

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

Monotonicity

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

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 $\Pr(\text{defier}) = 1/3$.

Monotonicity

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

In this case, the IV estimand is

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

which is not bound between τ_c and τ_d.

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

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

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

Multiple instruments

In practice

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

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

while our second stage is still

$$\mathrm{Y}_i = \beta_0 + \beta_1 \widehat{\mathrm{D}}_i + \beta_2 \mathrm{X}_i + v_i$$

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[†]

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

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

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)

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

`estimatr`

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

Today: `esitmatr` and `iv_robust()`.

2SLS and R

`estimatr`

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

Today: `esitmatr` and `iv_robust()`.

Specifically, we give `iv_robust()` the relationship that we want separted from the instrument by `|`

2SLS and R

`estimatr`

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

Today: `esitmatr` and `iv_robust()`.

Specifically, we give `iv_robust()` the relationship that we want separted from the instrument by `|`, *e.g.*,

```
# Estimate 2SLS
iv_robust(Y ~ D | Z, data = sample_df, se_type = "classical") %>%
  tidy() %>% select(1:5)
```

```
#>          term estimate std.error statistic      p.value
#> 1 (Intercept) 5.786204 2.9744230  1.945320 0.0546020456
#> 2           D 1.107801 0.3043264  3.640173 0.0004372703
```

Now in two stages!

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

```r
# First stage
stage1 = lm_robust(D ~ Z, data = sample_df, se_type = "classical")
# First-stage results
stage1 %>% tidy() %>% select(1:5)
```

```
#>          term  estimate std.error statistic      p.value
#> 1 (Intercept) 8.8226148 0.3169568 27.835389 2.486413e-48
#> 2           Z 0.3257347 0.1031506  3.157857 2.112927e-03
```

Second stage

We just need to add $\widehat{\mathrm{D}}_i$ to our dataset.

```r
# Add fitted (first-stage) values to data
sample_df %<>% mutate(D_hat = stage1$fitted.values)
# Second stage
stage2 = lm_robust(Y ~ D_hat, data = sample_df, se_type = "classical")
# Second-stage results
stage2 %>% tidy() %>% select(1:5)
```

```
#>          term estimate std.error statistic    p.value
#> 1 (Intercept) 5.786204 5.4132099  1.068904 0.28773854
#> 2       D_hat 1.107801 0.5538496  2.000184 0.04824759
```

Standard errors

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

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 $\mathbf{D}_i$ via $\mathbf{Z}_i$.
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

Instrumental variables

Two-stage least squares