Instrumental Variables and Local Average Treatment Effects

Francis J. DiTraglia

University of Oxford

Treatment Effects: The Basics

Where have we been and where are we headed?

Lecture #1

- If D is randomly assigned, it is straightforward to learn the ATE.
- If D is not randomly assigned, selection bias / confounding imply that $\mathbb{E}[Y|D=1] - \mathbb{E}[Y|D=0]$ usually doesn't tell us what we want to know.

Lecture #2

- Even if D wasn't randomly assigned, perhaps there's no selection bias *after we adjust* for observed variables X. This is called selection on observables.
- Avoiding **bad controls** requires a causal model; DAGs help us reason about these.

Lectures #3–5

When the selection-on-observables approach fails, is there anything else we could try?

Recall from Lecture #2

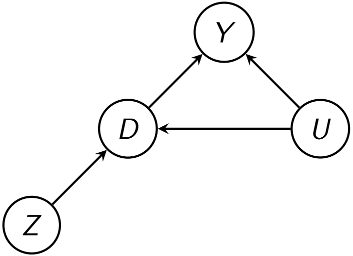
Back-door Path

- A path between treatment and outcome starting with edge pointing *into* treatment.
- Back-door paths are *non-causal*: only edges pointing *out* from treatment represent causal effects.

Back-door Criterion

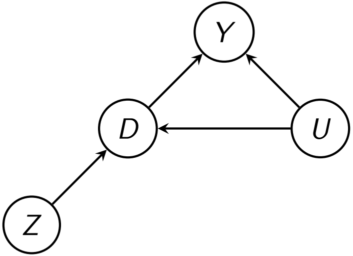
1. List all the paths that connect treatment and outcome.
2. Check which of them *open*. A path is *open* unless it contains a collider.
3. Check which of them are *back-door paths*: contain an arrow pointing at D.
4. If there are no open back-door paths, you're done. If not, look for nodes you can condition on to **block** remaining open back-door paths without opening new ones.

Exercise: If D and Z are binary, which statements are true?



1. $\mathbb{E}[Y|D=1] - \mathbb{E}[Y|D=0] = D \to Y$ causal effect
2. $\mathbb{E}[D|Z=1] - \mathbb{E}[D|Z=0] = Z \to D$ causal effect
3. $\mathbb{E}[Y|Z=1] - \mathbb{E}[Y|Z=0] = Z \to Y$ causal effect
4. We can learn the $D \to Y$ effect by conditioning on U.
5. We can learn the $D \to Y$ effect by conditioning on Z.

Solution: 1 and 5 are False, 2–4 are True[1]

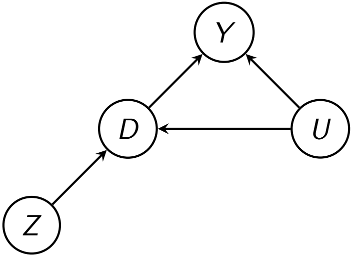


1. $\mathbb{E}[Y|D=1] - \mathbb{E}[Y|D=0] \neq D \to Y$ causal effect
2. $\mathbb{E}[D|Z=1] - \mathbb{E}[D|Z=0] = Z \to D$ causal effect
3. $\mathbb{E}[Y|Z=1] - \mathbb{E}[Y|Z=0] = Z \to Y$ causal effect
4. We can learn the $D \to Y$ effect by conditioning on U.
5. We can't learn the $D \to Y$ effect by conditioning on Z.

- Conditioning on U blocks the backdoor path $D \leftarrow U \to Y$.
- No open backdoor paths between Z and D or between Z and Y.
- Conditioning on Z does not block the backdoor path $D \leftarrow U \to Y$.

[1] Conditioning on Z is a *disastrous idea*: see my blog post "A Good Instrument is a Bad Control".

In this DAG, Z is a so-called "Instrumental Variable"



Setting

- Want to learn the $D \to Y$ causal effect
- U represents unobserved causes of D and Y.
- Can't use selection on observables.

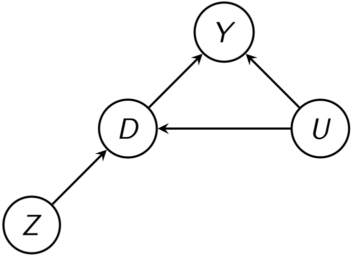
Relevance

Z and D are adjacent: Z causes D.

Exogeneity / Exclusion

Z and U are *not adjacent* and Z and Y are *not* adjacent.

Example: Effectiveness of Charter Schools



Research Question

Does attending a charter school increase math scores?

Unobserved Counfounders

U could include "ability", "grit", family background, etc.

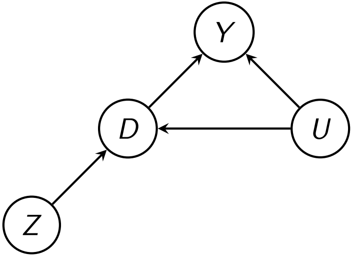
What are we looking for?

Observed variable Z that causes charter school attendance but is unrelated to U and has no direct effect on math scores.

Clever Idea

When oversubscribed, some charter schools use a lottery to choose which students are admitted. Let $Z = 1$ if a student **wins the lottery**.

Instrumental Variable Intuition



From our Warm-up Exercise:

- $\mathbb{E}[D|Z=1] - \mathbb{E}[D|Z=0] = Z \to D$ causal effect
- $\mathbb{E}[Y|Z=1] - \mathbb{E}[Y|Z=0] = Z \to Y$ causal effect

- Z only affects Y *through* its causal effect on D, which in turn affects Y.
- Therefore: $(Z \to Y \text{ effect}) = (Z \to D \text{ effect}) \times (D \to Y \text{ effect})$.

$$(D \to Y \text{ effect}) = \frac{(Z \to Y \text{ effect})}{(Z \to D \text{ effect})} = \frac{\mathbb{E}[Y|Z=1] - \mathbb{E}[Y|Z=0]}{\mathbb{E}[D|Z=1] - \mathbb{E}[D|Z=0]}$$

The "Textbook" Linear, Homogeneous Effects Model

- Linear causal model with homogeneous treatment effects: $Y \leftarrow \alpha + \beta D + U$
- Model says that changing D has the same effect for *everyone*: increasing D by one unit increases Y by β units
- D doesn't have to be binary; if it is we can make a link with potential outcomes:

$$
\begin{aligned}
D = 0 \implies Y &= \alpha + U & \implies Y_0 &= \alpha + U \\
D = 1 \implies Y &= (\alpha + \beta) + U & \implies Y_1 &= (\alpha + \beta) + U
\end{aligned}
$$

- Therefore, if D is binary, $\beta = Y_1 - Y_0$, a constant that is the same for everyone.
- Linearity isn't an extra assumption if D is binary
- Since $\beta = Y_1 - Y_0$ is constant, it equals $\mathbb{E}(Y_1 - Y_0) \equiv$ ATE.
- The next few slides assume you know a bit about linear regression.

Recall: Linear Regression and Exogeneity

Exogeneity

In the causal model ($Y \leftarrow \alpha + \beta D + U$) we say that D is **exogenous** if $\mathrm{Cov}(D, U) = 0$.

Population Linear Regression

The slope coefficient from a regression of Y on D is $\beta_{\mathrm{OLS}} \equiv \dfrac{\mathrm{Cov}(D, Y)}{\mathrm{Var}(D)}$.

Properties of Covariance

$$\begin{aligned} \mathrm{Cov}(X, W) &= \mathrm{Cov}(W, X) & \mathrm{Cov}(aX + b, W) &= a\mathrm{Cov}(X, W) \\ \mathrm{Cov}(X, X) &= \mathrm{Var}(X) & \mathrm{Cov}(X, W + V) &= \mathrm{Cov}(X, W) + \mathrm{Cov}(X, V) \end{aligned}$$

Therefore

$$\beta_{\mathrm{OLS}} \equiv \frac{\mathrm{Cov}(D, Y)}{\mathrm{Var}(D)} = \frac{\mathrm{Cov}(D, \alpha + \beta D + U)}{\mathrm{Var}(D)} = \frac{\beta \mathrm{Cov}(D, D) + \mathrm{Cov}(D, U)}{\mathrm{Var}(D)} = \beta + \frac{\mathrm{Cov}(D, U)}{Var(D)}$$

If D is Exogenous and Binary, Linear Regression Gives the ATE

$$Y \leftarrow \alpha + \beta D + U, \quad \beta = Y_1 - Y_0 = \mathbb{E}(Y_1 - Y_0) \equiv \text{ATE}$$

$$\text{Cov}(D, U) = 0 \implies \beta_{\text{OLS}} \equiv \frac{\text{Cov}(D, Y)}{\text{Var}(D)} = \beta + \frac{\text{Cov}(D, U)}{\text{Var}(D)} = \beta = \text{ATE}$$

Wait a second...

How does this relate to $\text{ATE} = \mathbb{E}[Y|D = 1] - \mathbb{E}[Y|D = 0]$ from Lecture #1?

Recall: The Fundamental Decomposition

$$\underbrace{\mathbb{E}(Y|D=1)-\mathbb{E}(Y|D=0)}_{\text{Observed Difference of Means}} = \underbrace{\mathbb{E}(Y_1 - Y_0|D=1)}_{\text{TOT}} + \underbrace{[\mathbb{E}(Y_0|D=1)-\mathbb{E}(Y_0|D=0)]}_{\text{Selection Bias}}$$

Homogeneous Effects Model

- $(Y \leftarrow \alpha + \beta D + U)$ equivalent to $Y_0 = \alpha + U$ and $Y_1 = (\alpha + \beta) + U$
- $\text{TOT} \equiv \mathbb{E}(Y_1 - Y_0|D=1) = \mathbb{E}(\beta|D=1) = \beta = \text{ATE}$
- $\mathbb{E}(Y_0|D=1) - \mathbb{E}(Y_0|D=0) = \mathbb{E}(\alpha + U|D=1) - \mathbb{E}(\alpha + U|D=0)$
- Hence, the fundamental decomposition becomes

$$\mathbb{E}(Y|D=1)-\mathbb{E}(Y|D=0) = \beta + [\mathbb{E}(U|D=1) - \mathbb{E}(U|D=0)]$$

Does this agree with the regression expression from above?

Recall: Properties of $\mathbb{E}(W|X=x) \equiv \sum_{\text{all } x} w \cdot \mathbb{P}(W=w|X=x)$

Linearity

$$\mathbb{E}(cW|X=x) = c\mathbb{E}(W|X=x)$$

$$\mathbb{E}(W+Z|X=x) = \mathbb{E}(W|X=x) + \mathbb{E}(Z|X=x)$$

Iterated Expectations

$$\mathbb{E}(W) = \mathbb{E}_X[\mathbb{E}(W|X)] \equiv \sum_{\text{all } x} \mathbb{E}(W|X=x)\mathbb{P}(X=x)$$

$$\mathbb{E}(W|Z=z) = \mathbb{E}_{(X|Z=z)}[\mathbb{E}(W|X,Z=z)] \equiv \sum_{\text{all } x} \mathbb{E}(W|X=x,Z=z)\mathbb{P}(X=x|Z=z)$$

One more property: “Taking Out What is Known”

Mathematics

$$\mathbb{E}[f(X) \cdot W|X] = f(X) \cdot \mathbb{E}[W|X] \quad \text{for any (measurable) function } f \text{ of } X.$$

Intuition

- $\mathbb{E}[W|X]$ is the expectation of W if we *pretend* that we know the realization of X.
- The realization of X is simply a constant; so is the realization of $f(X)$.
- We can pull constants in front of an expectation.

$\text{Cov}(W,X)/\text{Var}(X) = \mathbb{E}(W|X=1) - \mathbb{E}(W|X=0)$ for binary X.

Step 1

Recall that $\mathbb{E}(X) = \mathbb{P}(X=1) \equiv p$ and $\text{Var}(X) = p(1-p)$ if X is binary.

Step 2

Recall that $\text{Cov}(W,X) = \mathbb{E}(WX) - \mathbb{E}(W)\mathbb{E}(X)$ so we only need $\mathbb{E}(WX)$ and $\mathbb{E}(W)$.

Step 3

Iterated Expectations: $\mathbb{E}(W) = \mathbb{E}_X\left[\mathbb{E}(W|X)\right] = \mathbb{E}(W|X=1)p + \mathbb{E}(W|X=0)(1-p)$.

Step 4 – Exercise: Show that $\mathbb{E}(WX) = \mathbb{E}(W|X=1)p$.

$\text{Cov}(W,X)/\text{Var}(X) = \mathbb{E}(W|X=1) - \mathbb{E}(W|X=0)$ for binary X.

Step 1

Recall that $\mathbb{E}(X) = \mathbb{P}(X=1) \equiv p$ and $\text{Var}(X) = p(1-p)$ if X is binary.

Step 2

Recall that $\text{Cov}(W,X) = \mathbb{E}(WX) - \mathbb{E}(W)\mathbb{E}(X)$ so we only need $\mathbb{E}(WX)$ and $\mathbb{E}(W)$.

Step 3

Iterated Expectations: $\mathbb{E}(W) = \mathbb{E}_X\left[\mathbb{E}(W|X)\right] = \mathbb{E}(W|X=1)p + \mathbb{E}(W|X=0)(1-p)$.

Step 4 – Exercise: Show that $\mathbb{E}(WX) = \mathbb{E}(W|X=1)p$.

Iterated Expectations and Taking Out What is Known

$$
\begin{aligned}
\mathbb{E}(WX) &= \mathbb{E}_X[X\mathbb{E}(W|X)] = 0 \times \mathbb{E}(W|X=0)(1-p) + 1 \times \mathbb{E}(W|X=1)p \\
&= \mathbb{E}(W|X=1)p
\end{aligned}
$$

$\text{Cov}(W, X)/\text{Var}(X) = \mathbb{E}(W|X=1) - \mathbb{E}(W|X=0)$ for binary X.

Previous Slide

- Step 1: $\mathbb{E}(X) = p$ and $\text{Var}(X) = p(1-p)$
- Step 2: $\text{Cov}(W, X) = \mathbb{E}(WX) - \mathbb{E}(W)\mathbb{E}(X)$
- Step 3: $\mathbb{E}(W) = \mathbb{E}(W|X=1)p + \mathbb{E}(W|X=0)(1-p)$
- Step 4: $\mathbb{E}(WX) = \mathbb{E}(W|X=1)p$

Putting the Pieces Together

$$
\begin{aligned}
\text{Cov}(W, X) &= \mathbb{E}(WX) - \mathbb{E}(W)\mathbb{E}(X) \\
&= \mathbb{E}(W|X=1)p - [\mathbb{E}(W|X=1)p + \mathbb{E}(W|X=0)(1-p)]p \\
&= \mathbb{E}(W|X=1)p(1-p) - \mathbb{E}(W|X=0)(1-p)p \\
&= [\mathbb{E}(W|X=1) - \mathbb{E}(W|X=0)]\text{Var}(X)
\end{aligned}
$$

This also makes sense if you think of regression and dummy variables...

So yes: everything works out as it should!

Fundamental Decomposition

$$\underbrace{\mathbb{E}(Y|D=1) - \mathbb{E}(Y|D=0)}_{\text{Observed Difference of Means}} = \underbrace{\mathbb{E}(Y_1 - Y_0|D=1)}_{\text{TOT}} + \underbrace{[\mathbb{E}(Y_0|D=1) - \mathbb{E}(Y_0|D=0)]}_{\text{Selection Bias}}$$

Homogeneous Effects Model

$$\mathbb{E}(Y|D=1) - \mathbb{E}(Y|D=0) = \beta + [\mathbb{E}(U|D=1) - \mathbb{E}(U|D=0)]$$

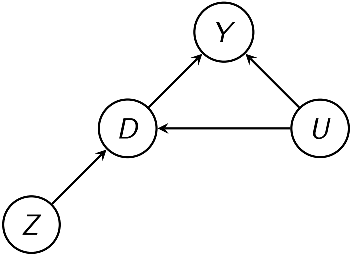
Regression Version

$$\beta_{\text{OLS}} \equiv \frac{\text{Cov}(D,Y)}{\text{Var}(D)} = \beta + \frac{\text{Cov}(D,U)}{\text{Var}(D)}$$

Previous Slide

If X is binary then $\dfrac{\text{Cov}(W,X)}{\text{Var}(X)} = \mathbb{E}(W|X=1) - \mathbb{E}(W|X=0)$.

The “Textbook” Instrumental Variables Model



Linear, Homogeneous Model

$Y \leftarrow \alpha + \beta D + U$ (notice: doesn't include Z!)

Endogenous Treament

The treatment D is **endogenous** if $\text{Cov}(D, U) \neq 0$.

Instrument Relevance

Z is **relevant** if $\text{Cov}(Z, D) \neq 0$, i.e. $Z \rightarrow D$.

Instrument Exogeneity / Exclusion

Z is **exogenous** if $\text{Cov}(Z, U) = 0$; i.e. $Z \not\leftrightarrow U$ and $Z \not\leftrightarrow Y$.

Valid Instrument

Z is a **valid instrument** if it is relevant and exogenous.

The "Textbook" Instrumental Variables Model

Linear, Homogeneous Model

$Y \leftarrow \alpha + \beta D + U$

Valid Instrument

Z is **relevant** and **exogenous**: $\text{Cov}(Z, D) \neq 0$ and $\text{Cov}(Z, U) = 0$

Exercise – Show that $\text{Cov}(Z, Y)/\text{Cov}(Z, D) = \beta$.

The "Textbook" Instrumental Variables Model

Linear, Homogeneous Model

$Y \leftarrow \alpha + \beta D + U$

Valid Instrument

Z is **relevant** and **exogenous**: $\text{Cov}(Z, D) \neq 0$ and $\text{Cov}(Z, U) = 0$

Exercise – Show that $\text{Cov}(Z, Y)/\text{Cov}(Z, D) = \beta$.

$$\beta_{\text{IV}} \equiv \frac{\text{Cov}(Z, Y)}{\text{Cov}(Z, D)} = \frac{\text{Cov}(Z, \alpha + \beta D + U)}{\text{Cov}(Z, D)} = \frac{\beta \text{Cov}(Z, D) + \text{Cov}(Z, U)}{\text{Cov}(Z, D)} = \beta = \text{ATE}$$

Notice

When Z is binary this coincides with our idea from earlier in the lecture:

$$\beta_{\text{IV}} \equiv \frac{\text{Cov}(Z, Y)}{\text{Cov}(Z, D)} = \frac{\text{Cov}(Z, Y)/\text{Var}(Z)}{\text{Cov}(Z, D)/\text{Var}(Z)} = \frac{\mathbb{E}(Y|Z = 1) - \mathbb{E}(Y|Z = 0)}{\mathbb{E}(D|Z = 1) - \mathbb{E}(D|Z = 0)}$$

What's the role of instrument relevance?

Exercise: Why do we need $\text{Cov}(Z, D) \neq 0$?

What's the role of instrument relevance?

Exercise: Why do we need $\text{Cov}(Z, D) \neq 0$?

- Math answer: appears in *denominator* of the IV expression; can't divide by zero!
- Causal inference answer: $\text{Cov}(Z, D)$ means Z has no causal effect on D.

Exercise: can we test either of the IV assumptions?

What's the role of instrument relevance?

Exercise: Why do we need $\text{Cov}(Z, D) \neq 0$?

- Math answer: appears in *denominator* of the IV expression; can't divide by zero!
- Causal inference answer: $\text{Cov}(Z, D)$ means Z has no causal effect on D.

Exercise: can we test either of the IV assumptions?

- $\text{Cov}(Z, D)$ is something we can calculate from data, so we can test it.
- $\text{Cov}(Z, U)$ depends on U, something we don't observe. That's not *quite* the end of the story though: register for *Beyond the Basics* if you want to learn more!

Simulation Example

```
set.seed(1234)
n <- 5000
u <- rnorm(n)
z <- rbinom(n, size = 1, prob = 0.5)
cov(z, u) # exogenous instrument
```

```
## [1] -0.0005708841
```

```
d <- rbinom(n, size = 1, prob = plogis(2 * z - u - 1))
cov(d, u) # endogenous treatment
```

```
## [1] -0.1871822
```

```
cov(d, z) # relevant instrument
```

```
## [1] 0.09425341
```

Simulation Example

```r
alpha <- 0
beta <- 1
y <- alpha + beta * d + u

cov(d, y) / var(d) # OLS
```

```
## [1] 0.2513902
```

```r
cov(d, u) / var(d) # This plus OLS should be approximately beta
```

```
## [1] -0.7486098
```

```r
cov(z, y) / cov(z, d) # IV
```

```
## [1] 0.9939431
```

Simulation Example

```r
cov(z, y) / cov(z, d) # IV
```

```
## [1] 0.9939431
```

```r
numerator <- mean(y[z == 1]) - mean(y[z == 0])
denominator <- mean(d[z == 1]) - mean(d[z == 0])

numerator / denominator # Should be identical to IV
```

```
## [1] 0.9939431
```

But treatment effects are heterogeneous!

The Rest of the Lecture

- "Textbook" IV solves selection bias but assumes homogeneous effects.
- Does β_{IV} have any meaning if treatment effects vary?

Crucial Question

Who gets treated and why?

Easiest Way to Understand

Experiments with **non-compliance**: the treatment that is *assigned* may not be the one that is *received*

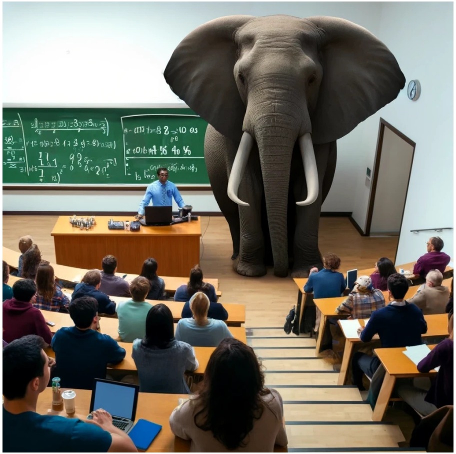


Figure 1: The Elephant in the Room.

Example: Pawn Lending in Mexico City[2]

Pawn Lending

- Valuable object (**pawn**) as collateral; receive loan for 70% of its appraised value.
- Regain your pawn by repaying loan plus interest by the deadline, otherwise lose it.

Status Quo Contract

- Single payment due at the end of three months; no reminders.
- Over 40% of borrowers default, losing their pawn and any payments made.
- Strictly worse off than if they'd sold their pawn for 100% of its appraised value!

New "Commitment" Contract

Monthly payments, small penalties for late payment & reminders. Fewer defaults?

[2] See "The Controlled Choice Design and Privated Paternalism in Pawnshop Borrowing" for more.

Example: Pawn Lending in Mexico City

Randomized Controlled Trial

- $Z = 0 \implies$ status quo contract
- $Z = 1 \implies$ *choice* of contracts

One-sided Non-compliance

- Everyone with $Z = 0$ receives the status quo contract
- People with $Z = 1$ can opt-in to the new "commitment" contract.

Research Question

What is the causal effect of *receiving* the new contract.

Figure 2: Commitment Choice.

Compliers: People who only take the treatment when offered.

One-sided Non-compliance

Z is randomly assigned; $Z = 0 \implies D = 0$; $Z = 1 \implies$ free to choose D.

First Stage: $Z \to D$

- Effect of treatment **offer** on treatment **receipt**; probably varies across people!
- One-sided Non-compliance $\implies$ two possible $Z \to D$ effects
  - **Effect is zero**: $D = 0$ regardless of Z. (cf. "doomed" from disease example)
  - **Effect is one**: switch from $D = 0$ to $D = 1$. (cf. "cured" from disease example)

Complier

- Someone who only takes treatment when offered: $Z \to D$ effect is **one**
- Pawn Example: someone who would choose the commitment contract, if offered.
- It's likely that compliers have **systematically different treatment effects!**

IV with Heterogeneous Treatment Effects: One-sided Non-compliance

Let $C = 1$ if complier, zero otherwise. Then:

$$D = C \cdot Z \quad \implies \quad Y = Y_0 + D(Y_1 - Y_0) = Y_0 + C \cdot Z(Y_1 - Y_0).$$

Assumption: $Z \perp\!\!\!\perp (C, Y_0, Y_1)$

$$\mathbb{E}(Y|Z = 1) = \mathbb{E}[Y_0 + C \cdot (Y_1 - Y_0)|Z = 1] = \mathbb{E}(Y_0) + \mathbb{E}[C \cdot (Y_1 - Y_0)]$$

$$\mathbb{E}(Y|Z = 0) = \mathbb{E}(Y_0|Z = 0) = \mathbb{E}(Y_0)$$

Intent to Treat: $(Z \to Y)$

$$\begin{aligned} \text{ITT} &\equiv \mathbb{E}(Y|Z = 1) - \mathbb{E}(Y|Z = 0) = \mathbb{E}[C \cdot (Y_1 - Y_0)] \\ &= \mathbb{E}_C [C \cdot \mathbb{E}(Y_1 - Y_0|C)] = \mathbb{E}(Y_1 - Y_0|C = 1)\mathbb{P}(C = 1) \end{aligned}$$

IV with Heterogeneous Treatment Effects: One-sided Non-compliance

Previous Slide

$\boxed{D = C \cdot Z}$, $\boxed{\text{Assumption: } Z \perp\!\!\!\perp (C, Y_0, Y_1)}$, and $\boxed{\text{ITT} = \mathbb{E}(Y_1 - Y_0 | C = 1)\mathbb{P}(C = 1)}$

First Stage: $(Z \to D)$

$$\text{FS} \equiv \mathbb{E}(D|Z = 1) - \mathbb{E}(D|Z = 0) = \mathbb{E}(C|Z = 1) - 0 = \mathbb{E}(C) = \mathbb{P}(C = 1)$$

Result: $\text{IV} = \mathbb{E}(Y_1 - Y_0 | C = 1)$

- Under 1-sided non-compliance & heterogeneous treatment effects, IV equals the average causal effect **for compliers**.
- Since we divide by $\mathbb{P}(C = 1)$, need this to be positive.

One-sided Non-compliance: The Compliers are The Treated

Previous Slide

$\text{IV} = \text{ITT}/\text{FS} = \mathbb{E}(Y_1 - Y_0 | C = 1)$.

Two Observations

- Conditioning on $(Z = 1, C = 1)$ is *equivalent* to conditioning on $D = 1$.
- Properties[3] of conditional independence: $Z \perp\!\!\!\perp (Y_0, Y_1, C) \implies Z \perp\!\!\!\perp (Y_1 - Y_0) | C$.

Punchline

Under 1-sided non-compliance and heterogeneous treatment effects, IV equals TOT!

$$\begin{aligned} \text{TOT} &\equiv \mathbb{E}(Y_1 - Y_0 | D = 1) = \mathbb{E}(Y_1 - Y_0 | Z = 1, C = 1) \\ &= \mathbb{E}(Y_1 - Y_0 | C = 1) = \text{IV} \end{aligned}$$

[3] Specifically: "Weak Union" and "Decomposition". See https://expl.ai/LXPVDDN and Chapter 2.

Example: Pawn Lending in Mexico City

- Only 11% choose commitment.
- TOT for default is negative: commitment lowers default *for the sort of person who chooses it*
- Low take-up leads to relatively imprecise estimates.

Figure 3: He probably didn't choose commitment.

Example: The 1944 British Education Act[4]

The minimum school-leaving age in Britain increased from 14 to 15 in 1947. Within two years of this policy change, the portion of 14-year-olds leaving school fell from 57% to less than 10%.

The finding that some adults reported finishing school at age 14, even after the school-leaving age had been changed, may reflect measurement error, noncompliance, or delayed enforcement.

What is the causal effect of staying in school until 15 on wage?

Figure 4: What a difference a year makes!

[4] Quotes from Oreopoulos (2006).

Example: The 1944 British Education Act

Simplified Version

After policy change all **must** be treated; before some **choose** to be treated.

$Z = 0$

- Turned 14 just before policy change.
- Can choose $D = 0$ or $D = 1$

$Z = 1$

- Turned 14 just after policy change.
- Forced to have $D = 1$

Always-Taker

No $Z \to D$ causal effect: would say in school until age 15 regardless.

Figure 5: What a difference a year makes!

Two Kinds of One-sided Non-compliance

Pawn Lending Example

- $Z = 0 \implies D = 0$ but $Z = 1 \implies$ can choose $D = 0$ or 1.
- Someone who chooses $D = 1$ when $Z = 1$ is called a **complier**.
- Assumptions: $Z \perp\!\!\!\perp (C, Y_0, Y_1)$ and there are at least **some** compliers.
- IV gives average causal effect for **compliers**; equivalent to TOT

British Education Example

- $Z = 1 \implies D = 1$ but $Z = 0 \implies$ can choose $D = 0$ or 1.
- Someone who chooses $D = 1$ when $Z = 0$ is called an **always-taker**
- Assumptions: $Z \perp\!\!\!\perp (A, Y_0, Y_1)$ and **not everyone** is an always-taker.
- IV gives the average causal effect for people who **are not always-takers**.
- Equivalent to the **treatment on the untreated:** $\mathbb{E}(Y_1 - Y_0 | D = 0)$.

Example: KIPP Academy Lynn[5]

The nation's largest network of charter schools is the Knowledge is Power Program (KIPP).

KIPP schools target low income and minority students and . . . feature a long school day and year, selective teacher hiring, strict behavior norms, and encourage a strong student work ethic.

Descriptive accounts of KIPP suggest positive achievement effects, but critics argue that the apparent KIPP advantage reflects differences between students who attend traditional public schools and students that choose to attend KIPP.

Figure 6: Terrifying artist's rendition of a Charter School Lottery.

[5] Angrist et al (2010) and Angrist et al (2012)

Example: KIPP Academy Lynn[6]

KIPP Lynn . . . is the only charter school in Lynn Masschusetts, a low income city north of Boston.

Statewide regulations require Massachusetts charter schools to use a lottery when oversubscribed.

The 2005-2008 admissions lotteries are used here to develop a quasi-experimental research design. These randomized lotteries allow us to estimate the causal effect of KIPP Lynn on achievement, solving the problem of selection bias that plagues most studies of school effectiveness.

[6] Angrist et al (2010) and Angrist et al (2012)

Example: KIPP Academy Lynn[7]

Lottery

$Z = 1$ if *offered* place at KIPP Lynn.

Two-sided Noncompliance

- $Z = 0 \not\Rightarrow D = 0$; $Z = 1 \not\Rightarrow D = 1$
- 25% of lottery winners **didn't** attend KIPP.
- 3.5% of lottery losers **did** attend KIPP.

Research Question

What is the causal effect of attending KIPP Lynn ($D = 1$) on math test scores Y?

[7] Slightly simpler version of this example as presented in *Mastering 'Metrics*.

Two-sided Non-compliance and Potential Treatments

Potential Treatments (D_0, D_1)

- D_0 is a person's D if $Z = 0$
- D_1 is a person's D if $Z = 1$
- Observe $D = (1 - Z)D_0 + ZD_1$
- Compare to the disease example!

Type	D_0	D_1	$(D_1 - D_0)$
Never-taker (N)	0	0	0
Always-taker (A)	1	1	0
Complier (C)	0	1	1
Defier (D)	1	0	-1

Table 1: The four "compliance types" and their respective causal effects of Z on D.

KIPP Example

- **Never-takers** would **not attend** KIPP regardless of the lottery outcome.
- **Always-takers** would **attend** KIPP regardless of the lottery outcome.
- **Compliers** would attend KIPP if they won the lottery, but not if they lost.
- **Defiers** would only attend KIPP if they **lost** the lottery, just to spite you!

Assumption 1: No Defiers

What's the problem?

If treatment effects vary, need to compare average values of Y_1 and Y_0 for *same group of people* to learn a causal effect.

Type	D_0	D_1	$D(Z)$
Never-taker (N)	0	0	0
Always-taker (A)	1	1	1
Complier (C)	0	1	Z
Defier (D)	1	0	$1-Z$

Table 2: The four "compliance types" and their treatment take-up rules.

With Defiers

- Can't tell if someone with $(Z=1, D=0)$ is a never-taker or defier.
- Can't tell if someone with $(Z=0, D=1)$ is an always-taker or defier.
- Notice: there were **automatically** no defiers in the one-sided examples!

Assumption 1: No Defiers

Without Defiers

- $(Z = 1, D = 0) \implies$ never-taker.
- $(Z = 0, D = 1) \implies$ always-taker.

Notation

- $A = 1$ if always-taker, zero otherwise
- $C = 1$ if complier, zero otherwise

Implication

No Defiers implies that $D = A + C \cdot Z$ and hence

$$Y = Y_0 + D(Y_1 - Y_0) = Y_0 + (A + C \cdot Z)(Y_1 - Y_0)$$

Type	D_0	D_1	$D(Z)$
Never-taker (N)	0	0	0
Always-taker (A)	1	1	1
Complier (C)	0	1	Z

Table 3: The *three* "compliance types" if we assume no defiers.

Assumption 2: $Z \perp\!\!\!\perp (Y_0, Y_1, C, A)$

Previous Slide

No Defiers Assumption $\Rightarrow$ $\boxed{D = A + C \cdot Z}$ hence $\boxed{Y = Y_0 + (A + C \cdot Z)(Y_1 - Y_0)}$.

Using Assumption 2

$$\mathbb{E}(Y|Z=1) = \mathbb{E}[Y_0 + (A + C)(Y_1 - Y_0)|Z=1] = \mathbb{E}[Y_0 + (A + C)(Y_1 - Y_0)]$$

$$\mathbb{E}(Y|Z=0) = \mathbb{E}[Y_0 + A(Y_1 - Y_0)|Z=0] = \mathbb{E}[Y_0 + A(Y_1 - Y_0)]$$

$$\text{ITT} \equiv \mathbb{E}(Y|Z=1) - \mathbb{E}(Y|Z=0) = \mathbb{E}[C(Y_1 - Y_0)] = \mathbb{E}(Y_1 - Y_0|C=1)\mathbb{P}(C=1)$$

$$\text{FS} \equiv \mathbb{E}(D|Z=1) - \mathbb{E}(D|Z=0) = \mathbb{E}(C + A|Z=1) - \mathbb{E}(A|Z=0) = \mathbb{E}(C)$$

Therefore: $\text{IV} = \mathbb{E}(Y_1 - Y_0|C=1)$.

This is often called the Local Average Treatment Effect (LATE)

Example: KIPP Academy Lynn

- The local average treatment effect of attending KIPP Academy Lynn for one year is approximately half a standard deviation of math test scores.
- This is quite a sizable effect, but remember that it is *not* the ATE!
- We might wonder how the effect for compliers differs from that for the population at large.

Discussion of IV with Heterogeneous Treatment Effects

- If treatment effects are heterogeneous, IV does **not** give us the ATE:
  - One-sided non-compliance: TOT or TUT
  - Two-sided non-compliance: LATE
- Who are the compliers? Better LATE than nothing?
- Different instruments for the same treatment can yield different causal effects, since different people would choose to comply.
- Three assumptions:
  1. Relevance: $\mathbb{E}(D|Z=1) \neq \mathbb{E}(D|Z=0)$ is testable.
  2. No defiers (only in needed in 2-sided case)
  3. Exclusion/Exogeneity: $Z \perp\!\!\!\perp (Y_0, Y_1, C, A)$ is not.
- Crucial question is whether Z could have a causal effect **of its own**.

Much more to say about IV! Why not sign up for *Beyond the Basics* in September?

Appendix

Derivations for *The Other Kind* of One-sided Non-compliance

Intent-to-treat: $Z \rightarrow Y$

$\text{ITT} = \mathbb{E}(Y|Z=1) - \mathbb{E}(Y|Z=0)$

Treatment Take-up

- $A = 1$ if always-taker
- $D = Z + A \cdot (1 - Z)$

Outcome

$Y = Y_0 + D(Y_1 - Y_0)$

Combining

$Y = Y_0 + [Z + A \cdot (1 - Z)](Y_1 - Y_0)$

Assumption

$Z \perp\!\!\!\perp (A, Y_0, Y_1)$

$$\begin{aligned}
\mathbb{E}(Y|Z=1) &= \mathbb{E}[Y_0 + (Y_1 - Y_0)|Z=1] \\
&= \mathbb{E}(Y_1)
\end{aligned}$$

$$\begin{aligned}
\mathbb{E}(Y|Z=0) &= \mathbb{E}[Y_0 + A \cdot (Y_1 - Y_0)|Z=0] \\
&= \mathbb{E}(Y_0) + \mathbb{E}[A \cdot (Y_1 - Y_0)]
\end{aligned}$$

$$\begin{aligned}
\text{ITT} &= \mathbb{E}(Y_1) - \mathbb{E}(Y_0) - \mathbb{E}[A \cdot (Y_1 - Y_0)] \\
&= \mathbb{E}[(1 - A)(Y_1 - Y_0)] \\
&= \mathbb{E}_A\left[(1 - A) \cdot \mathbb{E}(Y_1 - Y_0|A)\right] \\
&= \mathbb{E}(Y_1 - Y_0|A=0)\mathbb{P}(A=0)
\end{aligned}$$

Derivations for *The Other Kind* of One-sided Non-compliance

Intent-to-treat: $Z \to Y$

$\text{ITT} = \mathbb{E}(Y_1 - Y_0|A = 0)\mathbb{P}(A = 0)$

First-Stage: $Z \to D$

$$\begin{aligned} \text{FS} &\equiv \mathbb{E}(D|Z = 1) - \mathbb{E}(D|Z = 0) \\ &= 1 - \mathbb{E}(D|Z = 0) \end{aligned}$$

Treatment Take-up

- $A = 1$ if always-taker
- $D = Z + A \cdot (1 - Z)$

Assumption

$Z \perp\!\!\!\perp (A, Y_0, Y_1)$

$$\begin{aligned} \text{FS} &= 1 - \mathbb{E}(A|Z = 0) \\ &= 1 - \mathbb{E}(A) \\ &= 1 - \mathbb{P}(A = 1) = \mathbb{P}(A = 0) \end{aligned}$$

$$\begin{aligned} \text{IV} \equiv \frac{\text{ITT}}{\text{FS}} &= \frac{\mathbb{E}(Y_1 - Y_0|A = 1)\mathbb{P}(A = 0)}{\mathbb{P}(A = 0)} \\ &= \mathbb{E}(Y_1 - Y_0|A = 0) \end{aligned}$$

In this case IV equals the ATE for people who only take the treatment *when forced to do so.*

Derivations for the *The Other Kind* of One-sided Non-compliance

Previous Slide

$\text{IV} = \text{ITT}/\text{FS} = \mathbb{E}(Y_1 - Y_0 | A = 0)$.

Two Observations

- Conditioning on $(Z = 0, A = 0)$ is *equivalent* to conditioning on $D = 0$.
- Properties[8] of conditional independence: $Z \perp\!\!\!\perp (Y_0, Y_1, A) \implies Z \perp\!\!\!\perp (Y_1 - Y_0) | A$.

Punchline

Under this form of 1-sided non-compliance, IV is the *treated on the untreated* effect:

$$\begin{aligned} \text{TUT} &\equiv \mathbb{E}(Y_1 - Y_0 | D = 0) = \mathbb{E}(Y_1 - Y_0 | Z = 0, A = 0) \\ &= \mathbb{E}(Y_1 - Y_0 | A = 0) = \text{IV} \end{aligned}$$

[8] Specifically: "Weak Union" and "Decomposition". See https://expl.ai/LXPVDDN and Chapter 2.