Advanced Econometrics I

Jürgen Meinecke

Lecture 9 of 12

Research School of Economics, Australian National University

Roadmap

- Instrumental Variables Estimation
  - Potential Outcomes Framework, Treatment Effects
  - Treatment Heterogeneity: What is IV Estimating?
  - Example

Some new terminology

You are interested in the effect of a *treatment* X_i a person receives on an outcome Y_i

To keep things simple, the treatment is binary: $X_i \in \{0, 1\}$

The outcome is a function of the treatment: $Y_i(X_i)$

Only two *potential outcomes* per person: $Y_i(0)$ and $Y_i(1)$

The individual treatment effect (ITE) is: $Y_i(1) - Y_i(0)$,
that is: the difference in potential outcomes

This is the effect of the treatment on person i

Problem:
The individual treatment effect is never observed
(bc only one of the potential outcomes is observed per person)

How can we solve the problem of the missing counterfactual?

One idea would be to find an otherwise identical person $j \neq i$ who did not receive the treatment

For that person, the observed factual would be $Y_j(0)$

The individual treatment effect would be $Y_i(1) - Y_j(0)$

We do not observe if $Y_i(p) = Y_j(p)$ for $p \in \{0, 1\}$

What else can be done?

We are moving goal posts:
Instead of studying the *individual treatment effect* we look at the *average treatment effect*

The ATE is given by $\mathrm{E}\left(Y_i(1) - Y_i(0)\right)$

It is the effect on the average person in the population

While it would be great to know about the ITE, learning about the ATE also is immensely important

People who like to run regression want to know:
Can I estimate the ATE using OLS?

How does ATE relate to our regression β?

The potential outcomes framework is closely related to the study of *randomized controlled trials (RCT)*

RCTs have their roots in medical literature

A typical example is that of a medication that is randomly offered to some part of a sample and a placebo treatment to the other part

What are examples of treatments in economics?

- job training
- changes of legislation
- reducing class size in primary school
- sending out fake CVs to employers

The point here is: you have found a credible way to assign treatment *randomly*, that is X_i can be viewed as random

The potential outcomes framework can be mapped into a regression model

In the data you observe (X_i, Y_i)

Still, to keep things simple, binary treatment: $X_i \in \{0, 1\}$

Observed outcome is given by

$$\begin{aligned} Y_i &:= Y_i(1) \cdot X_i + Y_i(0) \cdot (1 - X_i) \\ &= (Y_i(1) - Y_i(0))X_i + Y_i(0) \\ &= \underbrace{\mathrm{E}\left(Y_i(0)\right)}_{\beta_0} + \underbrace{\left(Y_i(1) - Y_i(0)\right)}_{\beta_{1i}} X_i + \underbrace{\left(Y_i(0) - \mathrm{E}\left(Y_i(0)\right)\right)}_{\tilde{u}_i} \\ &= \beta_0 + \beta_{1i} X_i + \tilde{u}_i \end{aligned}$$

The last line looks like a regression

Careful though:

$$Y_i = \beta_0 + \beta_{1i} X_i + \tilde{u}_i$$

The slope coefficient is *indidvidual specific* (it has an i-subscript)

The coefficient β_{1i} is the individual treatment effect

You wouldn't use OLS here unless you think that β_{1i} is constant

Even then you would still need $E(X_i \tilde{u}_i) = 0$, which will be implied by random treatment (as we show soon)

Let's turn $Y_i = \beta_0 + \beta_{1i} X_i + \tilde{u}_i$ into a regression model in which the slope coefficient does not have an i-subscript

$$
\begin{aligned}
Y_i &= \mathrm{E}(Y_i(0)) + (Y_i(1) - Y_i(0))\, X_i + (Y_i(0) - \mathrm{E}(Y_i(0))) \\
&= \mathrm{E}(Y_i(0)) + (Y_i(1) - Y_i(0))\, X_i + (Y_i(0) - \mathrm{E}(Y_i(0))) \\
&\qquad + \mathrm{E}\left(Y_i(1) - Y_i(0)\right) \cdot X_i - \mathrm{E}(Y_i(1) - Y_i(0)) \cdot X_i \\
&= \mathrm{E}(Y_i(0)) + \mathrm{E}(Y_i(1) - Y_i(0)) \cdot X_i \\
&\qquad + \Big((Y_i(0) - \mathrm{E}(Y_i(0))) \\
&\qquad + \quad (Y_i(1) - Y_i(0))\, X_i - \mathrm{E}(Y_i(1) - Y_i(0)) \cdot X_i \Big) \\
&= \beta_0 + \beta_1 \cdot X_i + u_i,
\end{aligned}
$$

where $\beta_0 := \mathrm{E}(Y_i(0))$ and $\beta_1 := \mathrm{E}(Y_i(1) - Y_i(0))$ and everything in big parentheses is u_i

Notice that $\beta_1 := \mathrm{E}\left(Y_i(1) - Y_i(0)\right)$ is equal to the ATE

Let's check if $\mathsf{E}(X_i u_i) = 0$

$$
\begin{aligned}
\mathsf{E}(u_i|X_i) &= \mathsf{E}\Big((Y_i(0) - \mathsf{E}(Y_i(0))) \\
&\qquad + (Y_i(1) - Y_i(0))\, X_i - \mathsf{E}(Y_i(1) - Y_i(0)) \cdot X_i | X_i \Big) \\
&= \mathsf{E}(Y_i(0)|X_i) - \mathsf{E}(\mathsf{E}(Y_i(0))|X_i) \\
&\qquad + \mathsf{E}(Y_i(1)|X_i) \cdot X_i - \mathsf{E}(Y_i(0)|X_i) \cdot X_i \\
&\qquad - \mathsf{E}(\mathsf{E}(Y_i(1))|X_i) \cdot X_i + \mathsf{E}(\mathsf{E}(Y_i(0))|X_i) \cdot X_i \\
&= \mathsf{E}(Y_i(0)|X_i) - \mathsf{E}(Y_i(0)) \\
&\qquad + \mathsf{E}(Y_i(1)|X_i) \cdot X_i - \mathsf{E}(Y_i(0)|X_i) \cdot X_i \\
&\qquad - \mathsf{E}(Y_i(1)) \cdot X_i + \mathsf{E}(Y_i(0)) \cdot X_i \\
&= \mathsf{E}(Y_i(0)) - \mathsf{E}(Y_i(0)) \\
&\qquad + \mathsf{E}(Y_i(1)) \cdot X_i - \mathsf{E}(Y_i(0)) \cdot X_i \\
&\qquad - \mathsf{E}(Y_i(1)) \cdot X_i + \mathsf{E}(Y_i(0)) \cdot X_i \\
&= 0
\end{aligned}
$$

Yep!

The first equality follows by definition

The second equality follows by breaking up all individual terms

The third equality follows by the law of total probability

The fourth equality follows if X_i is assigned randomly

The treatment effect literature typically writes: $(Y_i(1), Y_i(0)) \perp\!\!\!\perp X_i$

Theorem (OLS in Randomized Controlled Trial)

Suppose you have available data (X_i, Y_i) from a randomized controlled trial. In particular, X_i is a randomly assigned treatment dummy variable. Then the OLS estimator of β_1 in the model $Y_i = \beta_0 + \beta_1 X_i + u_i$ is a consistent estimator of the average treatment effect $E(Y_i(1) - Y_i(0))$.

Roadmap

Instrumental Variables Estimation

- Potential Outcomes Framework, Treatment Effects
- Treatment Heterogeneity: What is IV Estimating?
- Example

What if you cannot effectively randomize treatment?

Earlier I said that job training is a treatment

In practice, there's no way you can randomly assign job training as a treatment and expect full compliance

If you cannot effectively randomize treatment then X_i may not be independent of potential outcomes

There's a clever work around:
randomize *eligibility* for treatment instead

Let Z_i be eligibility for treatment, with $Z_i \in \{0, 1\}$

It is not a coincidence that we are using the letter Z_i here:
eligibility will play the role of an instrumental variable

We conjecture that $\hat{\beta}^{\text{IV}}$ could be a good estimator in this setting

Is this true? Does $\hat{\beta}^{\text{IV}}$ estimate the ATE?

Like before, let's study a model in which treatment effect is heterogeneous

$$Y_i = \beta_{0i} + \beta_{1i} X_i + u_i \qquad \text{(equation of interest)}$$
$$X_i = \pi_{0i} + \pi_{1i} Z_i + v_i \qquad \text{(first stage)}$$

where

$$\beta_{1i} = Y_i(1) - Y_i(0)$$
$$\pi_{1i} = X_i(1) - X_i(0),$$

with $Y_i(p) = Y_i(X_i = p)$ and $X_i(p) = X_i(Z_i = p)$ for $p \in \{0, 1\}$

Using a little bit of math, it can be shown that

$$\hat{\beta}^{\text{IV}} = \frac{\mathsf{E}\left(\beta_{1i} \cdot \pi_{1i}\right)}{\mathsf{E}\left(\pi_{1i}\right)} + \mathsf{o}_p(1) \neq \mathsf{E}\left(\beta_{1i}\right) + \mathsf{o}_p(1)$$

(you will show this in assignment 9)

Two results here:

- IV estimator does not converge to the ATE (bad news?)
- Instead it converges to $\mathrm{E}\left(\beta_{1i} \cdot \pi_{1i}\right) / \mathrm{E}\left(\pi_{1i}\right)$ (looks complicated)

Let's take a closer look at the probability limit

For no apparent reason, let's call it LATE

$$\text{LATE} := \frac{\mathrm{E}\left(\beta_{1i} \cdot \pi_{1i}\right)}{\mathrm{E}\left(\pi_{1i}\right)}$$

What is LATE and how does it relate to ATE?

Here is a useful way to contrast them:

$$\text{ATE} = \mathrm{E}\left(\beta_{1i}\right)$$

$$\text{LATE} = \frac{\mathrm{E}\left(\beta_{1i} \cdot \pi_{1i}\right)}{\mathrm{E}\left(\pi_{1i}\right)} = \mathrm{E}\left(\beta_{1i} \cdot \frac{\pi_{1i}}{\mathrm{E}\left(\pi_{1i}\right)}\right)$$

- interpret $\frac{\pi_{1i}}{\mathrm{E}(\pi_{1i})}$ as weights
- then the rhs is equal to the expected value of β_{1i} adjusted for these weights
- in other words: the rhs is the weighted average of β_{1i}
- ideally, we would not want any weights in there (because we are after the ATE, which is the simple average)
- some intuition for the weights: when π_{1i} is large relative to $\mathrm{E}(\pi_{1i})$ then the weight is large; therefore people with large π_{1i} influence the IV estimator more (their Z_i have a strong impact on X_i)

Putting things together: $\hat{\beta}^{\text{IV}}$ estimates the causal effect for those individuals for whom Z_i is most influential
(those with large π_{1i})

LATE is the acronym for *local average treatment effect*

The LATE can be understood as the ATE for the subpopulation whose treatment X_i is most heavily influenced by the instrument Z_i

LATE is an ATE only for this peculiar ("local") subpopulation; it is not equal to the ATE in the population

Actually, we can relate ATE and LATE:

Notice that $\text{Cov}(\beta_{1i}, \pi_{1i}) = \text{E}(\beta_{1i}, \pi_{1i}) - \text{E}(\beta_{1i}) \cdot \text{E}(\pi_{1i})$

It follows

$$\begin{aligned}
\text{LATE} &:= \frac{\text{E}(\beta_{1i} \cdot \pi_{1i})}{\text{E}(\pi_{1i})} \\
&= \frac{\text{E}(\beta_{1i})\text{E}(\pi_{1i}) + \text{Cov}(\beta_{1i}, \pi_{1i})}{\text{E}(\pi_{1i})} \\
&= \text{E}(\beta_{1i}) + \frac{\text{Cov}(\beta_{1i}, \pi_{1i})}{\text{E}(\pi_{1i})} \\
&= \text{ATE} + \frac{\text{Cov}(\beta_{1i}, \pi_{1i})}{\text{E}(\pi_{1i})}
\end{aligned}$$

In words: LATE equals ATE plus “some stuff”

From previous slide

$$\text{LATE} = \text{ATE} + \frac{\text{Cov}(\beta_{1i}, \pi_{1i})}{\text{E}(\pi_{1i})}$$

But what exactly is "some stuff"?

It is the covariance between the two individual-specific parameters β_{1i} and π_{1i}

If the treatment effect β_{1i} tends to be large for individuals for whom the effect of the instrument π_{1i} is also large, then $\text{Cov}(\beta_{1i}, \pi_{1i}) > 0$ and therefore LATE $>$ ATE
(supposing $\text{E}(\pi_{1i}) > 0$)

On the other hand, if the treatment effect β_{1i} tends to be small for individuals for whom the effect of the instrument π_{1i} is also large, then $\text{Cov}(\beta_{1i}, \pi_{1i}) < 0$ and therefore LATE $<$ ATE

When does IV estimate the ATE?

- If $\beta_{1i} = \beta_1$ (no heterogeneity in equation of interest)
- If $\pi_{1i} = \pi_1$ (no heterogeneity in first stage equation)
- If β_{1i} and π_{1i} vary but are independently distributed

But these three are unrealistic

In general, $\hat{\beta}^{\text{IV}}$ does not estimate ATE

Whether this is important depends on the application

Define four exhaustive and mutually exclusive types based on their treatment response wrt to a particular value of $Z_i \in \{0, 1\}$

$$\begin{cases} \text{always taker} & X_i(0) = 1 \text{ and } X_i(1) = 1 \\ \text{complier} & X_i(0) = 0 \text{ and } X_i(1) = 1 \\ \text{defier} & X_i(0) = 1 \text{ and } X_i(1) = 0 \\ \text{never taker} & X_i(0) = 0 \text{ and } X_i(1) = 0 \end{cases}$$

This results in the following values of π_{1i} for these types:

$$\begin{aligned} \pi_{1i} &= X_i(1) - X_i(0) \\ &= \begin{cases} 0 & \text{always taker} \\ 1 & \text{complier} \\ -1 & \text{defier} \\ 0 & \text{never taker} \end{cases} \end{aligned}$$

Let's say the proportions of these four types are $\tau_{AT}, \tau_C, \tau_D, \tau_{NT}$, adding up to one

Furthermore, for simplicity claim that $\tau_D = 0$ (no defiers)

Then

$$\begin{aligned} \mathsf{E}(\pi_{1i}) &= \tau_{AT}\mathsf{E}(\pi_{1i}|AT) + \tau_C\mathsf{E}(\pi_{1i}|C) + \tau_{NT}\mathsf{E}(\pi_{1i}|NT) \\ &= \tau_C\mathsf{E}(\pi_{1i}|C) \\ &= \tau_C \end{aligned}$$

Likewise

$$\mathsf{E}\left(\beta_{1i} \cdot \pi_{1i}\right) = \tau_C\mathsf{E}\left(\beta_{1i}|C\right)$$

Therefore

$$\text{LATE} = \frac{\mathsf{E}\left(\beta_{1i} \cdot \pi_{1i}\right)}{\mathsf{E}\left(\pi_{1i}\right)} = \mathsf{E}\left(\beta_{1i}|C\right) \neq \mathsf{E}\left(\beta_{1i}\right)$$

So

$$\begin{aligned} \text{LATE} &= \text{E}\left(\beta_{1i} | C\right) \\ &= \text{E}\left(Y_i(1) - Y_i(0) | C\right) \end{aligned}$$

This is important because it says that LATE is the ATE for the subpopulation of compliers

The four types AT, NT, D, and C differ in how their outcomes respond to a treatment

We would not expect a *homogenous* treatment effect, that is, each of these four types would have the same treatment effect

LATE is the treatment effect for one particular type, the compliers

IV estimation successfuly estimates that local treatment effect

When Z_i is binary, there's a special form for the IV estimator

$$\hat{\beta}^{\text{IV}} = \frac{s_{ZY}}{s_{XZ}} = \frac{\hat{\text{E}}(Y_i|Z_i = 1) - \hat{\text{E}}(Y_i|Z_i = 0)}{\hat{\text{E}}(X_i|Z_i = 1) - \hat{\text{E}}(X_i|Z_i = 0)}$$

It is customary to write

$$\hat{\beta}^{\text{IV}} = \frac{\bar{Y}_1 - \bar{Y}_0}{\bar{X}_1 - \bar{X}_0}$$

with

$$\bar{Y}_1 := \frac{\sum_{i=1}^N Z_i Y_i}{\sum_{i=1}^N Z_i} \qquad \bar{Y}_0 := \frac{\sum_{i=1}^N (1 - Z_i) Y_i}{\sum_{i=1}^N (1 - Z_i)}$$

$$\bar{X}_1 := \frac{\sum_{i=1}^N Z_i X_i}{\sum_{i=1}^N Z_i} \qquad \bar{X}_0 := \frac{\sum_{i=1}^N (1 - Z_i) X_i}{\sum_{i=1}^N (1 - Z_i)}$$

This represenation of $\hat{\beta}^{\text{IV}}$ is called the *Wald estimator*

More generally, the Wald estimator is really any estimator that compares averages in grouped data as portrayed here

Roadmap

Instrumental Variables Estimation

- Potential Outcomes Framework, Treatment Effects
- Treatment Heterogeneity: What is IV Estimating?
- Example

Let me present an application taken from Angrist and Pischke, "*Mostly Harmless Econometrics*", (2008)

The actual underlying paper is Bloom et al., "*The Benefits and Costs of JTPA Title II-A Programs: Key Findings from the National Job Training Partnership Act Study*", (1997)

Note:
I will be presenting a much simplified version of the paper

Background for understanding the paper

Their research question:
Is job training beneficial for economically disadvantaged adults?

People were randomly *made eligible* for job training

This is an example where treatment was not randomly assigned but instead the *eligibility* for treatment was

Key variables:

- X_i: treatment dummy equal 1 if received job training
- Z_i: dummy equal 1 if offered job training (randomly assigned)
- Y_i: total earnings in the 30-months period after random assignment

Typical example of *one-sided compliance*:

$$Z_i = 0 \Rightarrow X_i = 0$$

$$Z_i = 1 \Rightarrow X_i \in \{0, 1\}$$

A person in the control group cannot access treatment

You might expect that $X_i = Z_i$

But many people refuse the offer of treatment (takes effort!)

In the job training example

$$\widehat{\Pr}(X_i = 1 | Z_i = 1) = \hat{\mathrm{E}}(X_i | Z_i = 1) \approx 0.6$$

$$\widehat{\Pr}(X_i = 1 | Z_i = 0) = \hat{\mathrm{E}}(X_i | Z_i = 1) \approx 0.02$$

More or less confirms one-sided compliance

The estimation results...

	OLS	ITT		LATE
	$\hat{\mathrm{E}}(Y_i \mid X_i = 1)$ $-\hat{\mathrm{E}}(Y_i \mid X_i = 0)$	$\hat{\mathrm{E}}(Y_i \mid Z_i = 1)$ $-\hat{\mathrm{E}}(Y_i \mid Z_i = 0)$	$\hat{\mathrm{E}}(X_i \mid Z_i = 1)$ $-\hat{\mathrm{E}}(X_i \mid Z_i = 0)$	
Men	\$3,970	\$1,117	0.61	\$1,825
Women	\$2,133	\$1,243	0.64	\$1,942

ITT: *intention-to-treat* effect; the effect you would have calculated under full compliance

Here it gives you a sort of lower bound

But compliance was only around 60% therefore ITT underestimates LATE

LATE is the treatment effect for *compliers*:
the subpopulation who are willing to take the treatment if offered