Conducting Empirical Research

Lecture 5

Jonas Grunau

2024-02-05

Feedback on the exercises

For the **modelling part**:

- Have you managed to run a model within the Mimi framework by now?

For the **empirical part**:

- How far have you made it through the coding exercise?
- Do you need more help?

Exercise: Finding a research question

- Take 20 minutes to discuss with your neighbour what you came up with

Person A:

- What is your research question? Why it is important?
- Briefly describe what empirical approach you would take

Person B:

- Critically challenge the assumptions underlying the proposed methodology
  - Are they likely to hold? Under which conditions?

5 minutes break

Let's now get to the core of causal identification…

- How do we get from **observed data** to the **counterfactual**?
  - All we have at the start is observed data (at best) …
  - … and the counterfactual is the opposite of observed - unobservable *by definition*. Can only be estimated!

→ Empirical research that aims to recover causal effects hinges upon **estimating a (plausible) counterfactual**.

Getting from the observed data to a counterfactual…

Rubin causal model of potential outcomes

$$\underbrace{Y_i}_{\text{Observed outcome}} = D_i Y_i^1 + (1 - D_i) Y_i^0$$

- Y_i^1 (Y_i^0) is the **potential outcome** when unit i is treated (untreated)
- D_i equals 1 if a unit i is treated (zero otherwise)

Unit-specific treatment effect:

$$\delta_i = Y_i^1 - Y_i^0$$

What kind of treatment effect are we interested in?

Average treatment effect (ATE):

$$\begin{aligned} ATE &= E[\delta_i] \\ &= E[Y_i^1 - Y_i^0] \\ &= E[Y_i^1] - E[Y_i^0] \end{aligned}$$

Average treatment effect on the treated (ATT):

$$\begin{aligned} ATT &= E\big[\delta_i \mid D_i = 1\big] \\ &= E\big[Y_i^1 - Y_i^0 \mid D_i = 1\big] \\ &= E\big[Y_i^1 \mid D_i = 1\big] - E\big[Y_i^0 \mid D_i = 1\big] \end{aligned}$$

Average treatment effect on the untreated (ATU):

$$\begin{aligned} ATU &= E\big[\delta_i \mid D_i = 0\big] \\ &= E\big[Y_i^1 - Y_i^0 \mid D_i = 0\big] \\ &= E\big[Y_i^1 \mid D_i = 0\big] - E\big[Y_i^0 \mid D_i = 0 \end{aligned}$$

Example: Should patients receive surgery X?

- Treatment: Surgery X for sick patients
- Control: No surgery
- Y: Impact of surgery X on remaining life span

Unit	**?**	δ
1	.	.
2	.	.
3	.	.
4	.	.
5	.	.

Unit	Y^1	Y^0	δ
1	7	3	4
2	6	5	1
3	1	1	0
4	10	3	7
5	4	9	-5

We can now calculate the ATE

Unit	Y^1	Y^0	δ
1	7	3	4
2	6	5	1
3	1	1	0
4	10	3	7
5	4	9	-5

$$\begin{aligned} ATE &= E[\delta_i] \\ &= E[Y_i^1 - Y_i^0] \\ &= E[Y_i^1] - E[Y_i^0] \end{aligned}$$

How do we calculate the ATE from the data?

$E[Y^1] = 5.6$

$E[Y^0] = 4.2$

$ATE = 1.4$

→ The treatment increases life span by 1.8 years on average (not estimated - calculated!!!).

→ Great! Right patient number 5?

Other treatment effects:

$ATT = \frac{4+1+0+7}{4} = 3$

$ATU = \frac{-5}{1} = -5$

Let’s step back to reality

We never observe both Y_i^1 and Y_i^0 for the same unit!

Unit	Y^1	Y^0	δ
1	7	3	4
2	6	5	1
3	1	1	0
4	10	3	7
5	4	9	-5

Unit	Y	D	δ
1	7	1	?
2	6	1	?
3	1	1	?
4	10	1	?
5	9	0	?

- We only observe the realised outcomes and the treatment assignment.
- How would our estimated ATT look like if we **naively compare the average outcome of treated and untreated units**?

$$\underbrace{E\left[Y_i^1 \mid D=1\right]}_{6} - \underbrace{E\left[Y_i^0 \mid D=0\right]}_{9} = -3$$

Why can we not rely on the simple difference in means estimator?

$$\widehat{ATT} = E\big[Y_i^1 \mid D=1\big] - E\big[Y_i^0 \mid D=0\big]$$

→ add and substract $E\big[Y_i^0 \mid D=1\big]$

$$\widehat{ATT} = \underbrace{E\big[Y_i^1 \mid D=1\big] - E\big[Y_i^0 \mid D=1\big]}_{ATT} + \underbrace{E\big[Y_i^0 \mid D=1\big] - E\big[Y_i^0 \mid D=0\big]}_{\text{Selection Bias}}$$

Let's break the link

The aim of causal inference is to break the link between treatment and potential outcomes!

- How can researchers achieve that?
- Gold standard: **Randomised controlled trials (RCTs)**
  - Through randomisation: $E\left[Y_i^0 \mid D=1\right] = E\left[Y_i^0 \mid D=0\right]$

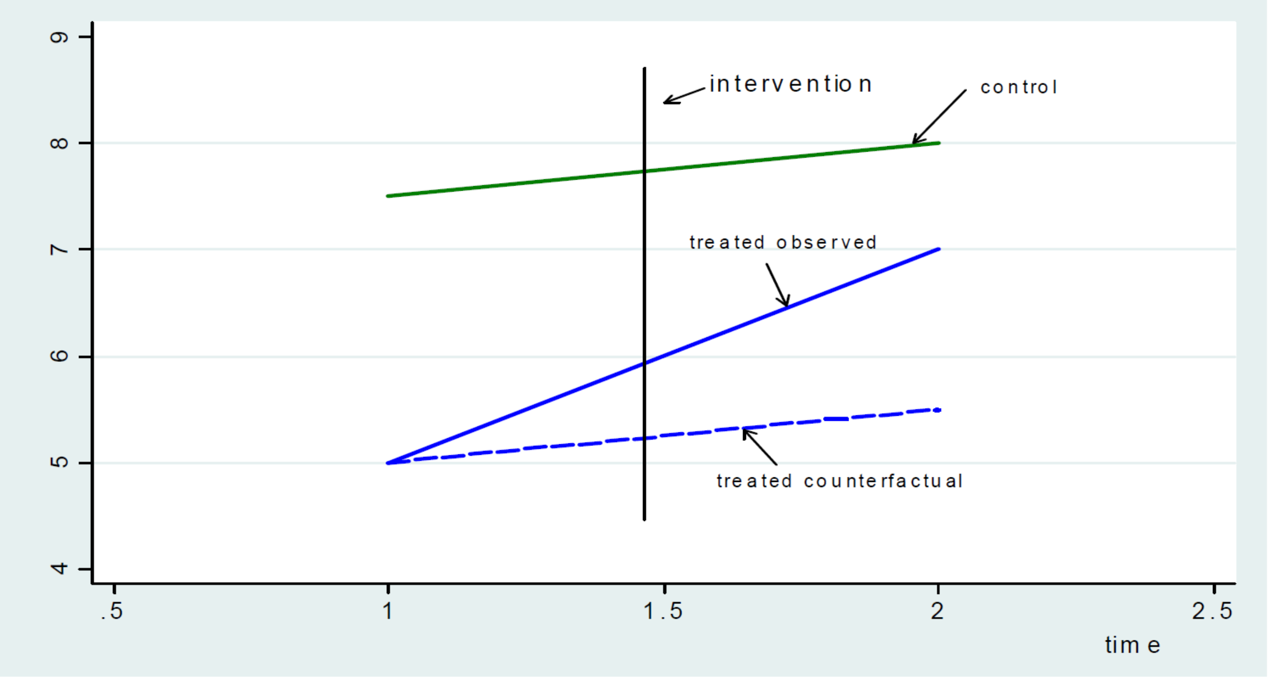
$$\widehat{ATT} = \underbrace{E\left[Y_i^1 \mid D=1\right] - E\left[Y_i^0 \mid D=1\right]}_{ATT} + \underbrace{E\left[Y_i^0 \mid D=1\right] - E\left[Y_i^0 \mid D=0\right]}_{\text{Selection Bias}=0}$$

→ In practice, RCTs are often not feasible due to ethical or practical reasons.

→ That is where **quasi-experimental methods** come into play.

The workhorse of applied environmental economics

Difference-in-differences (DID) and its many variants are one of the most widely used methodologies of modern applied environmental economics.



- What makes DID so attractive?
  - You do not have to assume there is no **selection bias**...
  - ... it just needs to be **stable across time**!

DID and potential outcomes

How we estimate the ATT with DID:

$$\hat{\delta}^{DID} = \underbrace{\Big(E[Y_T \mid Post] - E[Y_T \mid Pre]\Big)}_{\text{Change in treated outcome}} - \underbrace{\Big(E[Y_C \mid Post] - E[Y_C \mid Pre]\Big)}_{\text{Change in control outcome}}$$

Transform it into **potential outcomes**:

$$\hat{\delta}^{DID} = \Big(E[Y_T^1 \mid Post] - E[Y_T^0 \mid Pre]\Big) - \Big(E[Y_C^0 \mid Post] - E[Y_C^0 \mid Pre]\Big)$$

$$\hat{\delta}^{DID} = \Big(E[Y_T^1 \mid Post] - E[Y_T^0 \mid Pre]\Big) - \Big(E[Y_C^0 \mid Post] - E[Y_C^0 \mid Pre]\Big)$$
$$+ \underbrace{E[Y_T^0 \mid Post] - E[Y_T^0 \mid Post]}_{\text{Adding zero}}$$

The Parallel Trends Assumption

$$\hat{\delta}^{ATT} = \underbrace{E[Y_T^1 \mid Post] - E[Y_T^0 \mid Post]}_{\text{ATT}}$$
$$+ \underbrace{\underbrace{\Big[E[Y_T^0 \mid Post] - E[Y_T^0 \mid Pre]\Big]}_{\text{Change in } Y_0 \text{ (treated)}} - \underbrace{\Big[E[Y_C^0 \mid Post] - E[Y_C^0 \mid Pre]\Big]}_{\text{Change in } Y_0 \text{ (control)}}}_{\text{Non-parallel trends bias}}$$

Assumption: In the absence of treatment, treated and control outcomes would have evolved in parallel (constant selection bias).

→ not directly testable!

5 minutes break

Let’s now turn to the EU ETS.

Why is emissions trading (often portrayed as) a promising policy tool?

- Effectiveness?
- Political feasibility?
- Generating revenues to be recycled?
- Fairness?
- *Cost-effectiveness*: Achieving a given emissions target at least cost

What is required for an ETS to be cost-effective?

Necessary: **Equalising marginal abatement costs** among regulated units.

Independence property:

> "the market equilibrium in a cap-and-trade system will be cost-effective and **independent of the initial allocation of tradable rights**" (Hahn & Stavins, 2011)

This is a theoretical property of an ETS.

- How would you **test whether this property holds in the empirical context of the EU ETS**?
- What would be the perfect experiment? What would be a more realistic approach? What are the underlying assumptions?

Homework for next lecture

Read the paper by Zaklan (2023, AEJ): *Coase and Cap-and-Trade: Evidence on the Independence Property from the European Carbon Market*

Think about the following questions. Taking notes could be helpful as you will discuss them with your neighbour.

- What is the empirical approach taken by the author?
- Write down the assumptions underlying the author's empirical approach.
- What robustness tests does the author conduct?
- Do you think the empirical approach is convincing? What are the limitations?
- Think about additional robustness tests that could enhance the credibility of the paper.

You might find the paper by Roth et al. (2023) useful. Also feel free check out The Mixtape, on which this lecture builds.