Lecture 5: When complete randomization fails

The non-compliance to treatment assignments

Xi Chen

Rotterdam School of Management
Erasmus University Rotterdam

June 10, 2024

Outline

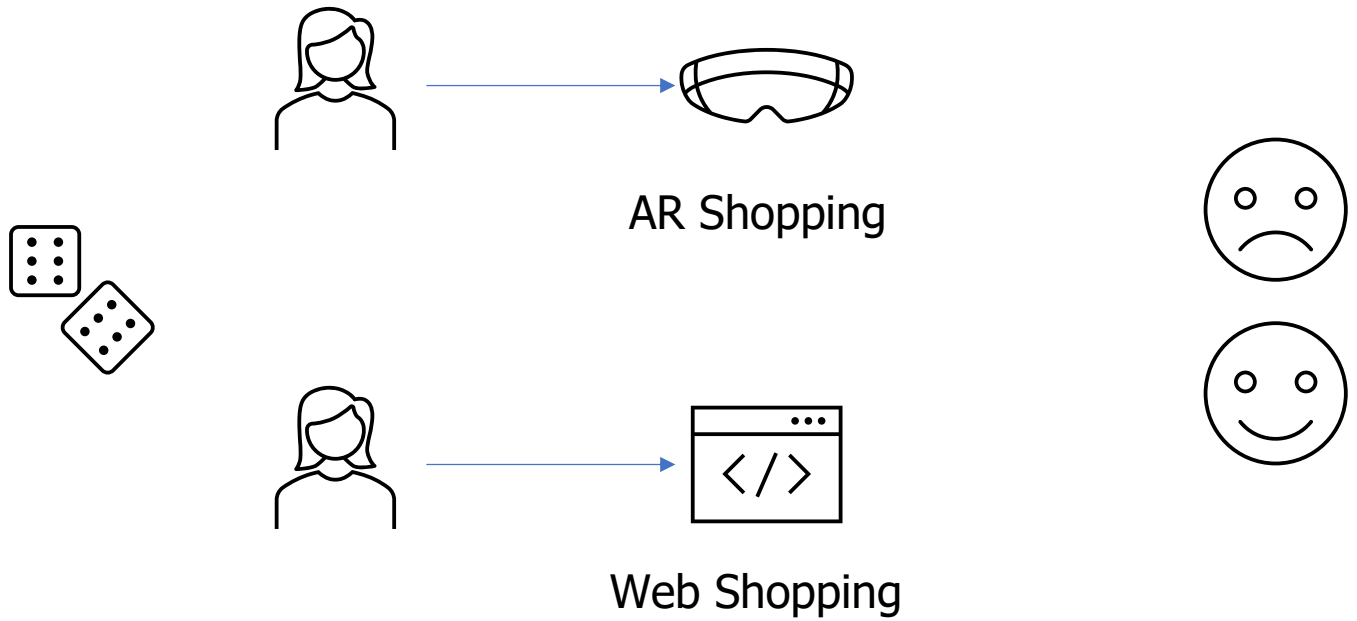
RSM Erasmus

Outline

RSM Erasmus

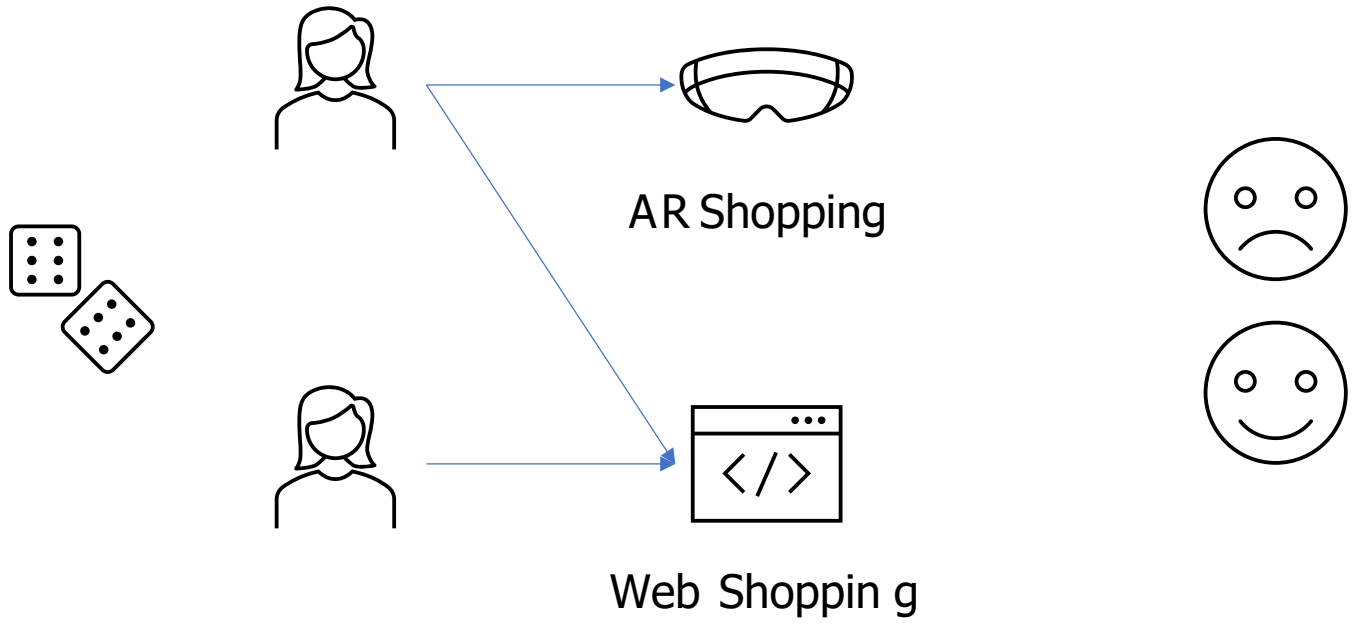
An example: the effect of AR in online furniture shopping

A lab experiment of AR shopping



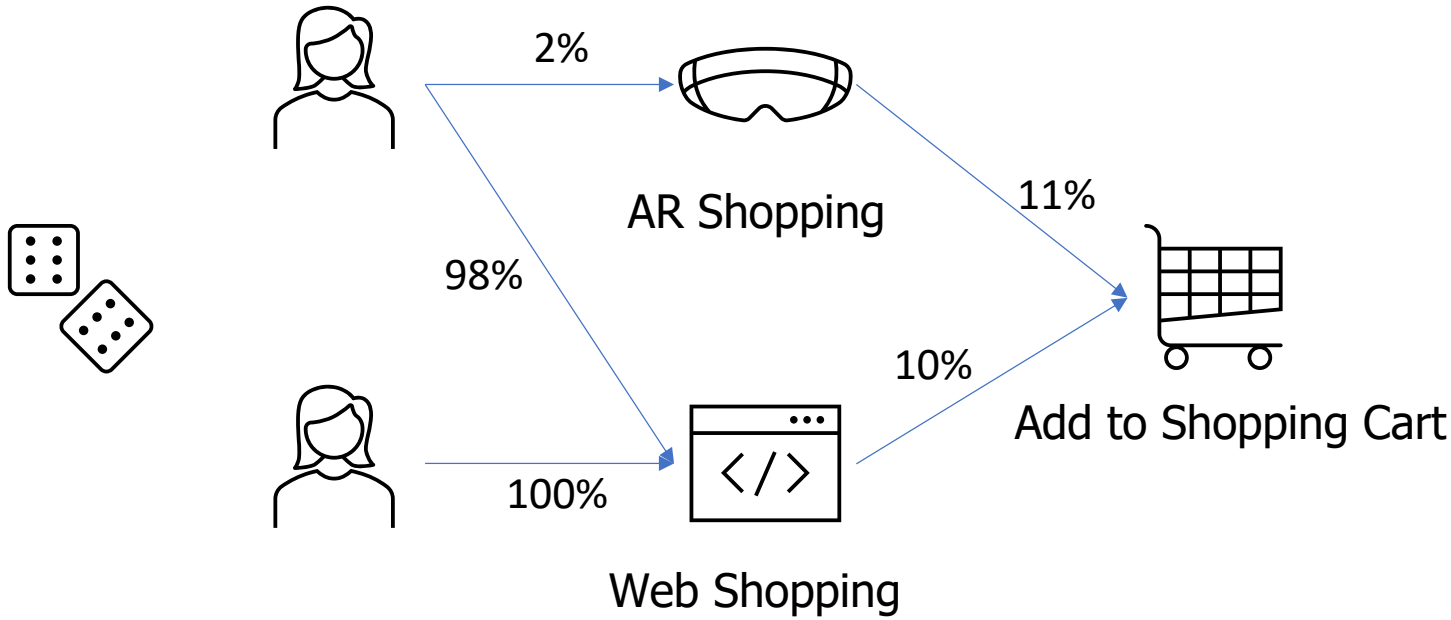
- Assignment to conditions and "force" people to use AR for shopping.

A field experiment of AR shopping



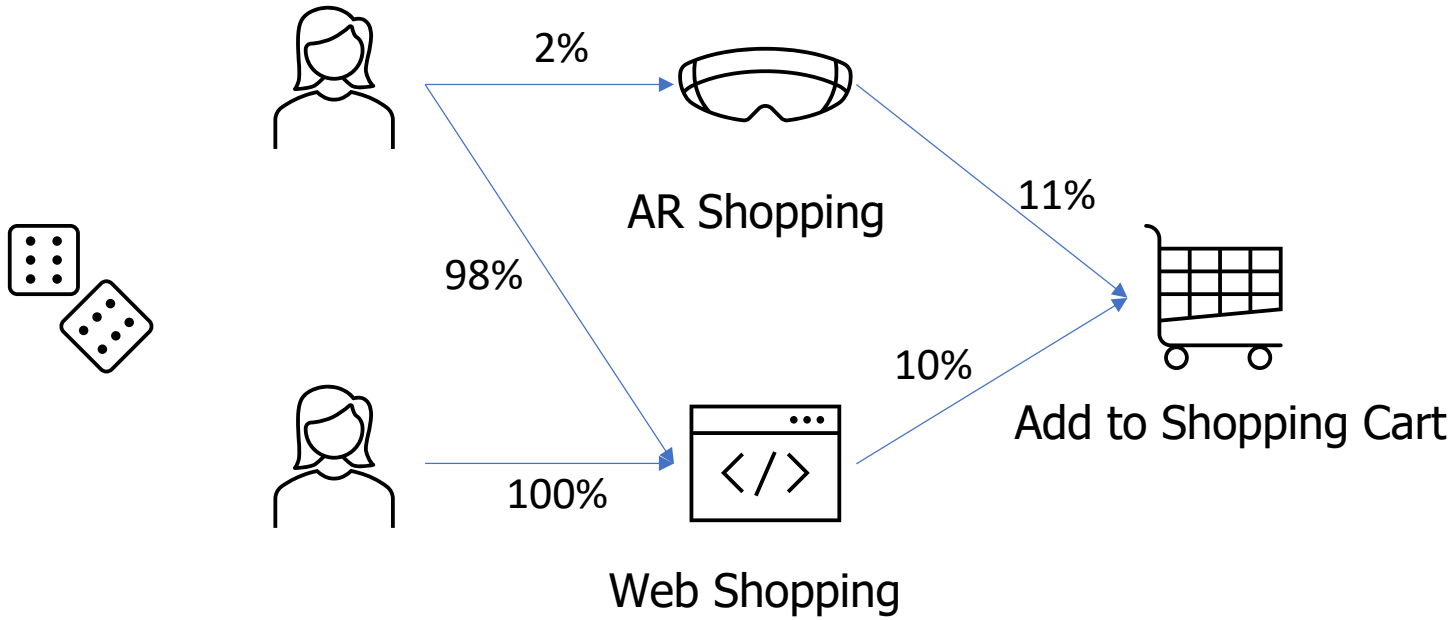
- Cannot "force" people to take the treatments;
- People will not fully comply.

A field experiment of AR shopping



- Treatment effect:
  $2\% \times 11\% + 98\% \times 10\% - 100\% \times 10\% = 0.02\%.$
- Local average treatment effect: $50 \times 0.02\% = 1\%.$

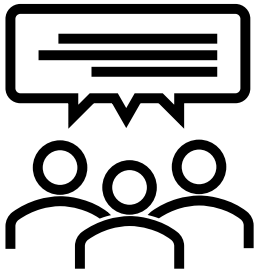
A field experiment of AR shopping



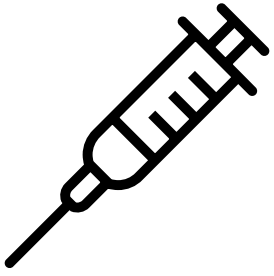
- Treatment effect:
  $2\% \times 11\% + 98\% \times 10\% - 100\% \times 10\% = 0.02\%.$
- Local average treatment effect: $50 \times 0.02\% = 1\%.$

RSM Erasmus

Many more examples...

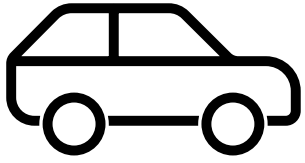


Loyalty programs on sales using coupons as incentives



Vaccinations on contagion with invitation letters

Military services on income with lottery drafting during Vietnam war



Car ownership on commute time under license plate lottery in Beijing

Compliance vs. Non-compliance

Compliance	Treatment $= 1$	Treatment $= 0$
Assignment $= 1$	100%	0%
Assignment $= 0$	0%	100%

$$\rho(A, D) = 1$$

Non-compliance	Treatment $= 1$	Treatment $= 0$
Assignment $= 1$	$<$100%	$>$0%
Assignment $= 0$	$>$0%	$<$100%

$$\rho(A, D) < 1$$

RSM Erasmus

Compliance vs. Non-compliance

Compliance	Treatment $= 1$	Treatment $= 0$
Assignment $= 1$	100%	0%
Assignment $= 0$	0%	100%

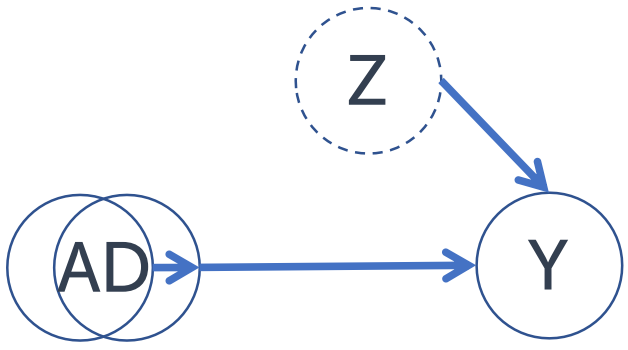
$$\rho(A, D) = 1$$

Non-compliance	Treatment $= 1$	Treatment $= 0$
Assignment $= 1$	$<$100%	$>$0%
Assignment $= 0$	$>$0%	$<$100%

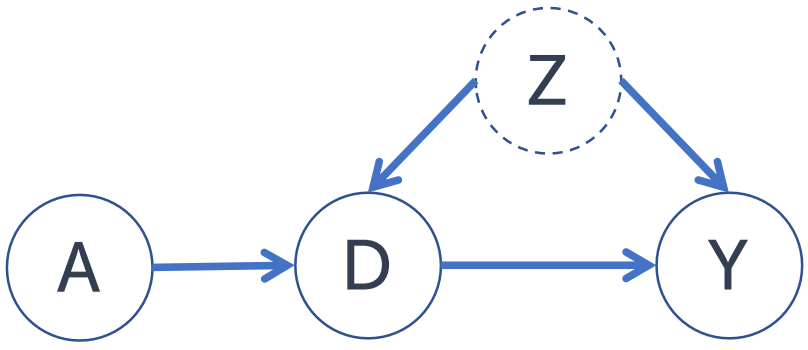
$$\rho(A, D) < 1$$

RSM Erasmus

How to think about the problem?



Complete Randomization

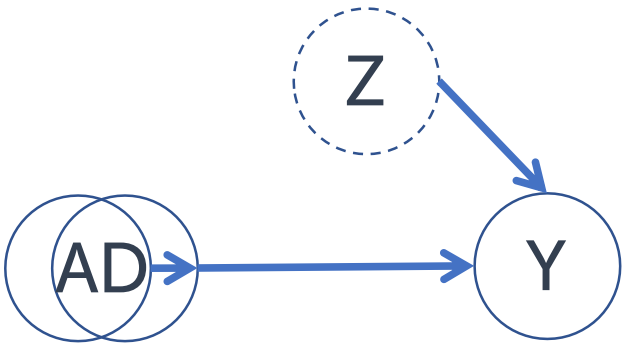


Non-compliance

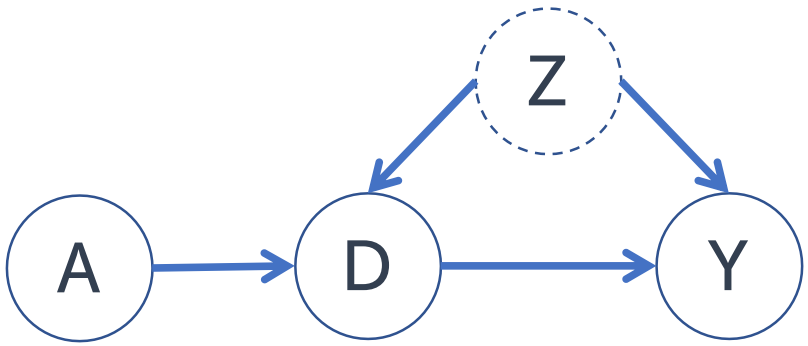
- The assignment A is still unconfounded, due to randomization.
- The actual treatment D is not fully aligned with A!
- For complete randomization: $P(D \mid A) = 1$, but for non-compliance: $P(D \mid A) < 1$.

RSM Erasmus

The DAG for non-compliance



Complete Randomization



Non-compliance

A closer look at the DAG for the non-compliance situation:

- A is an instrument for the effect $D \to Y$.
- The study of non-compliance $\Longleftrightarrow$ The study of instrumental variable.

RSM Erasmus

The seminal paper by Angrist, Imbens and Rubin (1996)

Identification of causal effects using instrumental variables

JD Angrist, GW Imbens, DB Rubin - Journal of the American ..., 1996 - Taylor & Francis

We outline a framework for causal inference in settings where assignment to a binary treatment is ignorable, but compliance with the assignment is not perfect so that the receipt ...

☆ Save Cite Cited by 7254 Related articles All 29 versions »

- Researchers had used instrumental variables (IV) for nearly 70 years without knowing the exact interpretation of the estimated effects.
- No one knew the exact meaning of IV estimates until the paper came out.

RSM Erasmus

Outline

RSM Erasmus

The basics

The assignment of treatment: $A_i \in \{0, 1\}$.
The actual treatment: $D_i \in \{0, 1\}$.
The potential outcomes of the receipt of treatment:

$$D_i = \begin{cases} D_i^0 & \text{if } A_i = 0 \\ D_i^1 & \text{if } A_i = 1 \end{cases}.$$

Fact

For an individual i, given an assignment A_i, we can only observe one of the potential outcomes for the receipt of assignment, with $D_i^{A_i} = 0$ or $D_i^{A_i} = 1$.

RSM Erasmus

The basics

The potential outcomes of Y_i are denoted as $Y_i\left(A_i, D_i^{A_i}\right)$, and we have 4 potential outcomes of Y_i:

$$\begin{cases} Y_i(0,0) & \text{if } A_i = 0 \text{ and } D_i^0 = 0 \\ Y_i(0,1) & \text{if } A_i = 0 \text{ and } D_i^0 = 1 \\ Y_i(1,0) & \text{if } A_i = 1 \text{ and } D_i^1 = 0 \\ Y_i(1,1) & \text{if } A_i = 1 \text{ and } D_i^1 = 1 \end{cases}.$$

Fact

*For an individual i, given the assignment A_i and the receipt of treatment D_i, we can only observe **one of four** potential outcomes on the final outcome Y_i.*

The basics

Comparing with complete randomization or full compliance:

$$\begin{cases} Y_i(0,0) & \text{if } A_i = 0 \text{ and } D_i^0 = 0 \\ \cancel{Y_i(0,1)} & \cancel{\text{if } A_i = 0 \text{ and } D_i^0 = 1} \\ \cancel{Y_i(1,0)} & \cancel{\text{if } A_i = 1 \text{ and } D_i^1 = 0} \\ Y_i(1,1) & \text{if } A_i = 1 \text{ and } D_i^1 = 1 \end{cases}.$$

Fact

In complete randomization, A_i and D_i are fully aligned, and only two potential outcomes of Y_i exist.

RSM Erasmus

Compliance behavior

With A and D both binary, we have 4 compliance types:

		D_i^1	
		0	**1**
D_i^0	**0**	Never-taker (nt) π_{nt}	Complier (co) π_{co}
	1	Defier (de) π_{de}	Always-taker (at) π_{at}

Compliance behavior

- It's a finite population behavior, not a super-population behavior (e.g., a never-taker in one study can be a complier in another).
- The "compliance type" is a pre-determined feature of people, and not influenced by the assignment in any ways.
- The specific compliance type of an individual is unobservable, and we assume the compliance types $\{nt, co, de, at\}$ are latent groups in the finite population.

RSM Erasmus

Compliance behavior

- It's a finite population behavior, not a super-population behavior (e.g., a never-taker in one study can be a complier in another).
- The "compliance type" is a pre-determined feature of people, and not influenced by the assignment in any ways.
- The specific compliance type of an individual is unobservable, and we assume the compliance types $\{nt, co, de, at\}$ are latent groups in the finite population.

RSM Erasmus

Compliance behavior

- It's a finite population behavior, not a super-population behavior (e.g., a never-taker in one study can be a complier in another).
- The "compliance type" is a pre-determined feature of people, and not influenced by the assignment in any ways.
- The specific compliance type of an individual is unobservable, and we assume the compliance types $\{nt, co, de, at\}$ are latent groups in the finite population.

RSM Erasmus

Failure to identify compliance types

From the observation of assignment A_i and the receipt of treatments D_i, we cannot infer the proportion of compliance types:

		D_i^{obs}	
		0	1
A_i	0	never-taker / complier	always-taker / defier
	1	never-taker / definer	always-taker / complier

It is a classic identification problem:

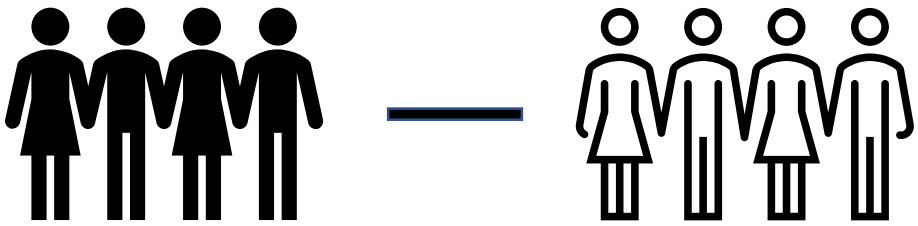
$$\begin{cases} b_1\pi_{nt} + (1-b_1)\,\pi_{co} &= \pi^{obs}(0,0) \\ b_2\pi_{at} + (1-b_2)\,\pi_{de} &= \pi^{obs}(0,1) \\ b_3\pi_{nt} + (1-b_3)\,\pi_{de} &= \pi^{obs}(1,0) \\ b_4\pi_{at} + (1-b_4)\,\pi_{co} &= \pi^{obs}(1,1) \end{cases}, \text{ with } b_j \in (0,1), \forall j$$

*No unique solution to the system of equations: 4 equations but 8 unkonwns.

Outline

RSM Erasmus

Naive analysis: "as-treated" analysis



$D_i^{obs} = 1$
$\bar{Y}_1^{obs}$

$D_i^{obs} = 0$
$\bar{Y}_0^{obs}$

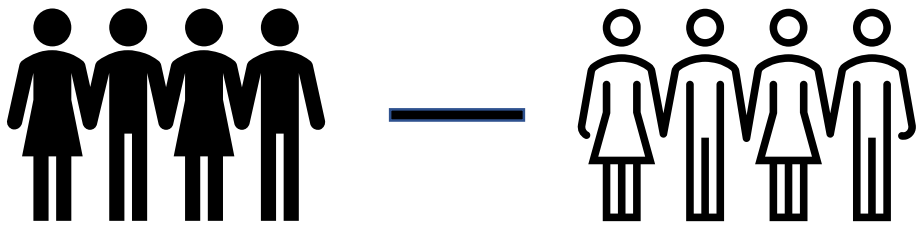
The as-treated analysis defines a treatment effect as **the difference in the average outcomes of people who actually receive the treatment and those who do not**.

$$\tau_{\text{as-treated}} = \frac{\sum_{i=1}^{N} Y_i^{obs} \cdot D_i^{obs}}{\sum_{i=1}^{N} D_i^{obs}} - \frac{\sum_{i=1}^{N} Y_i^{obs} \cdot \left(1 - D_i^{obs}\right)}{\sum_{i=1}^{N} \left(1 - D_i^{obs}\right)}$$

Problem: **people select into the treatment and the $\tau_{\text{as-treated}}$ is confounded!**

PSM Zafung

Naive analysis: "per-protocol" analysis



$A_i = 1$
$D_i^{obs} = 1$

$A_i = 0$
$D_i^{obs} = 0$

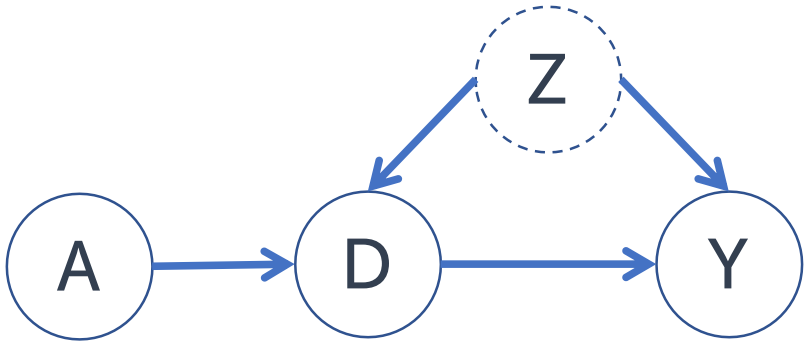
The per-protocol analysis defines a treatment effect as **the difference in the average outcomes of people who comply to the assignment**.

$$\tau_{\text{per-protocol}} = \frac{\sum_{i=1}^{N} Y_i^{obs} \cdot D_i^{obs} \cdot A_i}{\sum_{i=1}^{N} D_i^{obs} \cdot A_i} - \frac{\sum_{i=1}^{N} Y_i^{obs} \cdot \left(1 - D_i^{obs}\right) \cdot (1 - A_i)}{\sum_{i=1}^{N} \left(1 - D_i^{obs}\right) \cdot (1 - A_i)}$$

Problem: **people select into the treatment and the $\tau_{\text{per-protocol}}$ is confounded!** .

RSM Erasmus

What can we identify without further assumptions?



Non-compliance

Observation

- The assignment of treatment A_i is unconfounded!
- Therefore, we can identify $P(D \mid A)$ and $P(Y \mid A)$.

RSM Erasmus

The intention-to-treat effect (ITT)

Theorem (The identification of ITT)

Given the assumption that the assignment A_i is unconfounded, we can identify the ITT of A_i on the receipt of treatment D_i and Y_i.

$$\begin{cases} \widehat{ITT}_Y &= \frac{\sum_{i=1}^{N} Y_i^{obs} \cdot A_i}{\sum_{i=1}^{N} A_i} - \frac{\sum_{i=1}^{N} Y_i^{obs} \cdot (1-A_i)}{\sum_{i=1}^{N} (1-A_i)} \\ \widehat{ITT}_D &= \frac{\sum_{i=1}^{N} D_i^{obs} \cdot A_i}{\sum_{i=1}^{N} A_i} - \frac{\sum_{i=1}^{N} D_i^{obs} \cdot (1-A_i)}{\sum_{i=1}^{N} (1-A_i)} \end{cases}$$

RSM Erasmus

Interpreting ITT

The interpretation of ITT_Y is **the effect of assignment A on the final outcome of interest Y.**

Example (ITT for a field experiment on loyalty program)

An airline company intends to study the effect of their loyalty program on ticket revenue from customers. The design of the experiment is as such: the company sends randomly selected customers an incentive package (i.e., small gifts) for joining the loyalty program. The company realizes the non-compliance issue in the experiment and estimate the ITT. *Question: how to interpret the ITT effect in this case?*

RSM Erasmus

The limitation of ITT

A field experiment usually aims at "the effect of treatment", not "the effect of assignment."

A lot times the treatment effect is of interests:

Invitation letters to vaccinations on contagion.

But we want the effect of vaccinations on contagion!

Sometimes, the assignment effect is of interests:

Car plate lottery in Beijing on commute time.

It is still useful to learn the welfare implications of plate lottery!

RSM Erasmus

Can we know more than ITT?

Decomposing the ITT with $\text{ITT}_Y = E\left(Y_i\left(1, D_i^1\right) - Y_i\left(0, D_i^0\right)\right)$:

$$\begin{aligned}\text{ITT}_Y = &\text{ITT}_Y^{co} \times \pi_{co} \\ &+ \text{ITT}_Y^{nt} \times \pi_{nt} \\ &+ \text{ITT}_Y^{at} \times \pi_{at} \\ &+ \text{ITT}_Y^{de} \times \pi_{de}\end{aligned}$$

With

$$\text{ITT}_Y^g = E\left(Y_i\left(1, D_i^1\right) - Y_i\left(0, D_i^0\right) \mid i \in g\right)$$

$$\pi_g = P\left(i \in g\right)$$

RSM Erasmus

Take a closer look...

$$\begin{aligned} \text{ITT}_Y^{at} &= E\left(Y_i\left(1, D_i^1\right) - Y_i\left(0, D_i^0\right) \mid i \in at\right) \\ &= E\left(Y_i\left(1,1\right) - Y_i\left(0,1\right) \mid i \in at\right) \end{aligned}$$

$$\begin{aligned} \text{ITT}_Y^{nt} &= E\left(Y_i\left(1, D_i^1\right) - Y_i\left(0, D_i^0\right) \mid i \in nt\right) \\ &= E\left(Y_i\left(1,0\right) - Y_i\left(0,0\right) \mid i \in nt\right) \end{aligned}$$

Question: what it means if

$$E\left(Y_i\left(1,1\right) \mid i \in at\right) \neq E\left(Y_i\left(0,1\right) \mid i \in at\right)$$
$$E\left(Y_i\left(1,0\right) \mid i \in nt\right) \neq E\left(Y_i\left(0,0\right) \mid i \in nt\right) \textbf{?}$$

RSM Erasmus

Assumptions: exclusion restrictions for AT and NT

The assignment A_i for always-takers (AT) and non-takers (NT) will **not influence the final outcomes Y_i directly**.

$$E\left(Y_i\left(1,1\right) \mid i \in at\right) = E\left(Y_i\left(0,1\right) \mid i \in at\right)$$
$$E\left(Y_i\left(1,0\right) \mid i \in nt\right) = E\left(Y_i\left(0,0\right) \mid i \in nt\right)$$

Under the exclusion restrictions, we have $\text{ITT}_Y^{at} = 0$ and $\text{ITT}_Y^{nt} = 0$, so

$$\text{ITT}_Y = \text{ITT}_Y^{co} \times \pi_{co} + \text{ITT}_Y^{de} \times \pi_{de}$$

RSM Erasmus

Understanding the exclusion restrictions for AT and NT

Exclusions for AT and NT

The key feature of the exclusion restrictions for AT and NT is that they are, at their core, substantive assumptions, and cannot be statistically tested!

- A physical randomization is inadequate!
- Sufficient condition: "**double-blindness**" – both researchers and participants do not know the treatment.

Understanding the exclusion restrictions for AT and NT

- The exclusion restrictions need careful scrutiny outside of double blinding!
- In most settings, people are aware of their assignments and treatments.

Exclusions require substantive knowledge

In the loyalty program for an airliner, let's consider the never-takers, or people who would never join the loyalty program irrespective of the incentives. The incentives might change their purchases.

Never-takers	Normal Passengers	Business Travelers
Reasons?	Only travel few times	Care about privacy
Exclusions?	Reasonable	Less likely to hold

What else is needed?

$$\text{ITT}_Y = \text{ITT}_Y^{co} \times \pi_{co} + \text{ITT}_Y^{de} \times \pi_{de}$$

Monotonicity or No Defiers

There are no defiers in the finite population, with $D_i^1 > D_i^0$.

Implications for the monotonicity assumption:

$$\pi_{de} = 0 \Rightarrow \text{ITT}_Y = \text{ITT}_Y^{co} \times \pi_{co}$$

Again, this assumption is untestable and requires substantive knowledge!

RSM Erasmus

The local average treatment effect

Under exclusion restrictions, we can identify the local / complier average treatment effect (LATE or CATE).

$$\text{ITT}_Y^{co} = \frac{\text{ITT}_Y}{\pi_{co}}$$

Intuitively, it's the average treatment effect for a subgroup - compliers in the experiment (hence, **the name "local"**).

Exclusion Restriction for Compliers

For compliers, or with $i \in g = co$, we have $\boldsymbol{Y_i(0, D_i) = Y_i(1, D_i)}$. The implication is we can fully attribute the LATE to the receipt of treatment instead of the assignment of treatment.

Outline

RSM Erasmus

How to estimate?

Start with the definition of the LATE:

$$\text{ITT}_Y^{co} = \frac{\text{ITT}_Y}{\pi_{co}}$$

It is straightforward to estimate the overall ITT on Y

$$\text{ITT}_Y = \frac{\sum_{i=1}^{N} Y_i^{obs} \cdot A_i}{\sum_{i=1}^{N} A_i} - \frac{\sum_{i=1}^{N} Y_i^{obs} \cdot (1 - A_i)}{\sum_{i=1}^{N} (1 - A_i)}$$

Or it is the mean difference of people who are assigned the treatment ($A_i = 1$) vs. who are not ($A_i = 0$) on the final outcome.

RSM Erasmus

How to estimate?

How about the proportion of the latent complier group?
Observe:

$$\text{ITT}_D = \text{ITT}_D^{co} \times \pi_{co} + \text{ITT}_D^{nt} \times \pi_{nt} + \text{ITT}_D^{at} \times \pi_{at} + \text{ITT}_D^{de} \times \pi_{de}$$

With,

$$\text{ITT}_D^g = E\left(D_i^1 - D_i^0 \mid i \in g\right), \forall g \in \{co, nt, at, de\}$$

For always-takers and never-takers, $D_i^1 = D_i^0 = 1$ or 0

$$\text{ITT}_D^{at} = E\left(D_i^1 - D_i^0 \mid i \in at\right) = 0$$
$$\text{ITT}_D^{nt} = E\left(D_i^1 - D_i^0 \mid i \in nt\right) = 0$$

RSM Erasmus

How to estimate?

How about the proportion of the latent complier group?

$$\text{ITT}_D = \text{ITT}_D^{co} \times \pi_{co} + \text{ITT}_D^{de} \times \pi_{de}$$

For compliers, $D_i^1 - D_i^0 = 1$ and for defiers $D_i^1 - D_i^0 = -1$.

$$\text{ITT}_D^{co} = E\left(D_i^1 - D_i^0 \mid i \in co\right) = 1$$
$$\text{ITT}_D^{de} = E\left(D_i^1 - D_i^0 \mid i \in de\right) = -1$$

With these observations,

$$\begin{aligned} \text{ITT}_D &= \text{ITT}_D^{co} \times \pi_{co} + \text{ITT}_D^{de} \times \pi_{de} \\ &= \pi_{co} - \pi_{de} \\ &\underbrace{=}_{\textbf{No Defiers}} \pi_{co} \end{aligned}$$

RSM Erasmus

How to estimate?

With the derivation,

$$\tau_{\text{LATE}} = \text{ITT}_Y^{co} = \frac{\text{ITT}_Y}{\text{ITT}_D}$$

To estimate ITT_Y and ITT_D:

$$\widehat{\text{ITT}}_Y = \frac{\sum_{i=1}^N Y_i^{obs} \cdot A_i}{\sum_{i=1}^N A_i} - \frac{\sum_{i=1}^N Y_i^{obs} \cdot (1 - A_i)}{\sum_{i=1}^N (1 - A_i)}$$

$$\widehat{\text{ITT}}_D = \frac{\sum_{i=1}^N D_i^{obs} \cdot A_i}{\sum_{i=1}^N A_i} - \frac{\sum_{i=1}^N D_i^{obs} \cdot (1 - A_i)}{\sum_{i=1}^N (1 - A_i)}$$

ITT_Y and ITT_D are the mean difference people who are assigned the treatment ($A_i = 1$) vs. who are not ($A_i = 0$) on the final outcome Y and the treatment receipt D.

RSM Erasmus

The inference

Please check Imbens and Rubin (2015), p. 531 for the derivation.

$$V\left(\tau_{\text{LATE}}\right) = \frac{1}{\text{ITT}_D^2} \cdot V\left(\text{ITT}_Y\right) + \frac{\text{ITT}_Y^2}{\text{ITT}_D^4} \cdot V\left(\text{ITT}_D\right) - 2 \cdot \frac{\text{ITT}_Y}{\text{ITT}_D^3} \cdot COV\left(\text{ITT}_Y, \text{ITT}_D\right)$$

Define:

$$N_0 = \sum_{i=1}^{N} \left(1 - A_i\right) \text{ and } N_1 = \sum_{i=1}^{N} A_i$$

The sample variance of a random variable X as,

$$S_X^2 = \frac{1}{N-1} \sum_{i=1}^{N} \left(X_i - \overline{X}\right)^2$$

RSM Erasmus

The inference

The Neyman variance of the intention-to-treat on the receipt of the treatment is:

$$V\left(\mathrm{ITT}_D\right) = \frac{S^2_{D,0}}{N_0} + \frac{S^2_{D,1}}{N_1}$$

With $S^2_{D,0}$ and $S^2_{D,1}$ the sample variance of the receipt of treatment D of people of $A_i = 0$ and $A_i = 1$, respectively.

The Neyman variance of the intention-to-treat on the final outcome is:

$$V\left(\mathrm{ITT}_Y\right) = \frac{S^2_{Y,0}}{N_0} + \frac{S^2_{Y,1}}{N_1}$$

With $S^2_{Y,0}$ and $S^2_{Y,1}$ the sample variance of the final outcome Y of people of $A_i = 0$ and $A_i = 1$, respectively.

RSM Erasmus

The inference

The covariance between ITT_D and ITT_Y is:

$$COV\left(\text{ITT}_Y, \text{ITT}_D\right) = \frac{1}{N_1\left(N_1 - 1\right)} \sum_{i:A_i=1} \left(Y_i^{obs} - \overline{Y}_i^{obs}\right)\left(D_i^{obs} - \overline{D}_i^{obs}\right)$$

Alternative estimation method:

The traditional 2SLS method for IV estimation can also be applied, but makes unnecessary linear assumptions on the model.
But the method is useful when covariates are considered.

RSM Erasmus

An example

		Percentages of Groups			
Population	Assigned (A)	Control arm (A=0)		Treatment arm (A=1)	
	Receipt (D)	Control (D=0)	Treatment (D=1)	Control (D=0)	Treatment (D=1)
50	Complier	25			25
30	Never Taker	15		15	
20	Always Taker		10		10
100	Observed	40	10	15	35
		50		50	

		Mean Outcomes			
Population	Assigned (A)	Control arm (A=0)		Treatment arm (A=1)	
	Receipt (D)	Control (D=0)	Treatment (D=1)	Control (D=0)	Treatment (D=1)
50	Complier	500			600
30	Never Taker	480		480	
20	Always Taker		550		550
100	Observed	492.5	550.0	480.0	585.7
		504.0		554.0	

RSM Erasmus

An example

We can calculate the intention-to-treat effects and the LATE as below,

$$\text{ITT}_D = \pi_{co} = \frac{35}{50} - \frac{10}{50} = 0.5$$
$$\text{ITT}_Y = 554.0 - 504.0 = 50.0$$
$$\tau_{\text{LATE}} = \text{ITT}_Y^{co} = \frac{50.0}{0.5} = 100.0$$

To compare with the naive approaches,

$$\tau_{\text{As-treated}} = \frac{585.7 \cdot 35 + 550.0 \cdot 10}{35 + 10} - \frac{492.5 \cdot 40 + 480.0 \cdot 15}{40 + 15} = 88.7$$
$$\tau_{\text{Per-protocol}} = 585.7 - 492.5 = 93.2$$

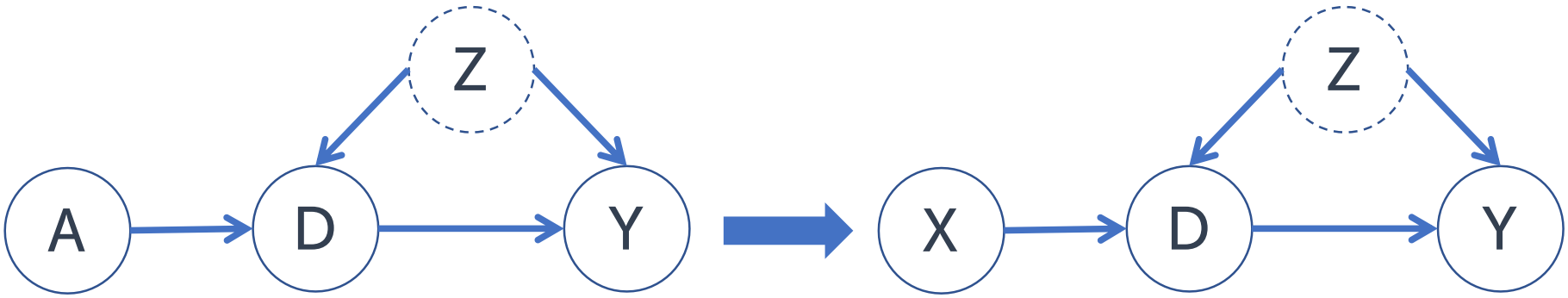
Both as-treated and per-protocol underestimate the LATE, as well as the intention-to-treat.

RSM Erasmus

Outline

RSM Erasmus

From non-compliance to instrumental variables

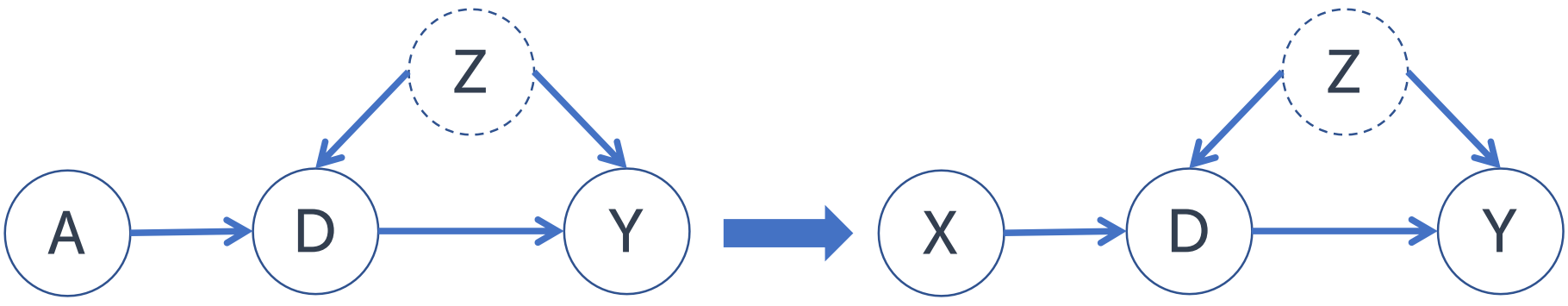


Non-compliance **Instrumental Variable**

- **Exogeneity**: X is "as-good-as-randomly assigned."
- **Exclusion**: X is only influences Y through D.
- **Relevance**: X influences D.

RSM Erasmus

From non-compliance to instrumental variables



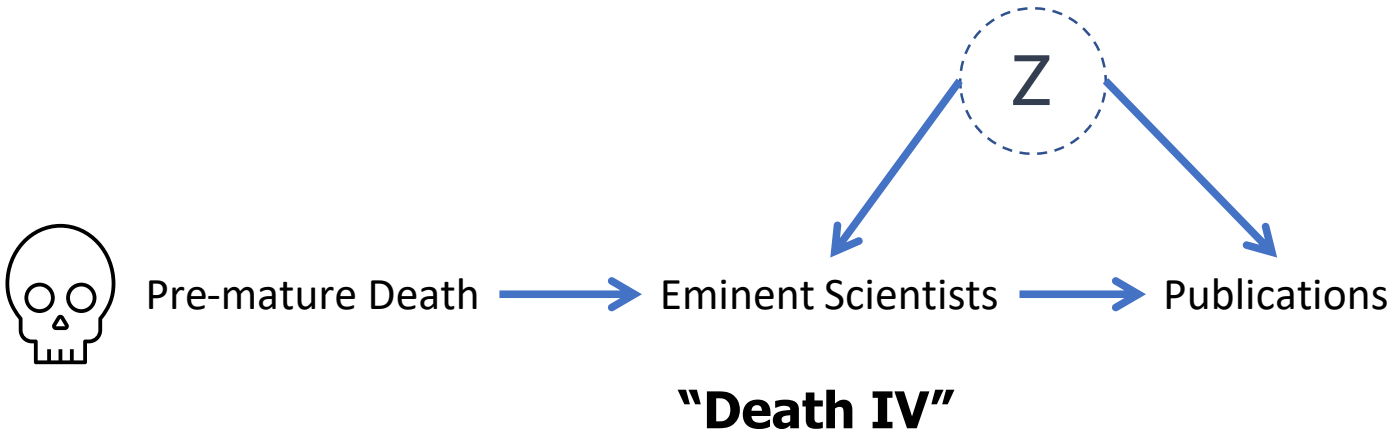
Non-compliance **Instrumental Variable**

- **Untestable**: each assumption, on its own, cannot be statistically tested.
- **Local**: the estimated effects from IV design are LATE (i.e., how much X shifts D).

RSM Erasmus

Where do good instruments come from?

Natural occurrences (a.k.a. natural experiments)
Azoulay et al. (2019)



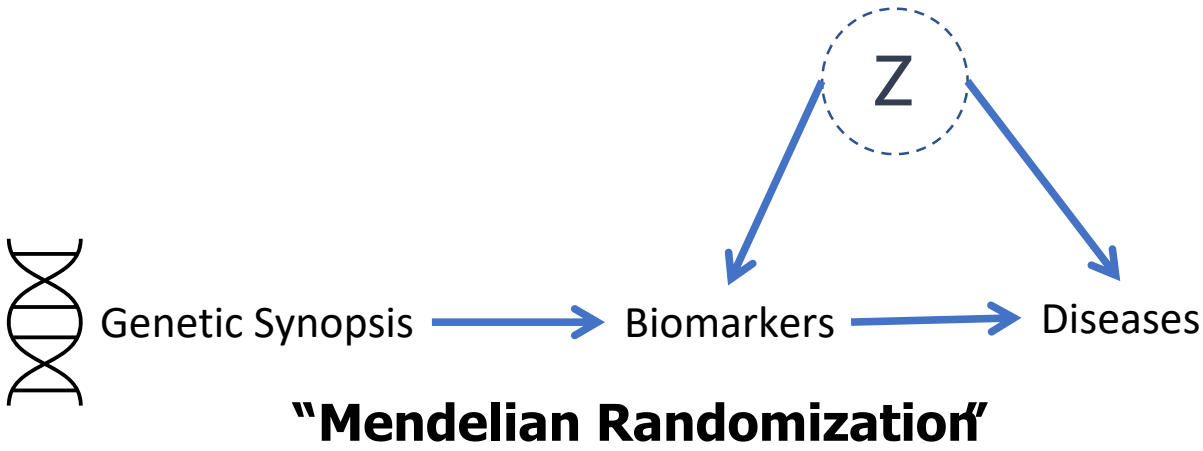
"Death IV"

- Exogeneity
- Exclusion
- Relevance

RSM Erasmus

Where do good instruments come from?

Natural occurrences (a.k.a. natural experiments)
Emdin et al. (2017)



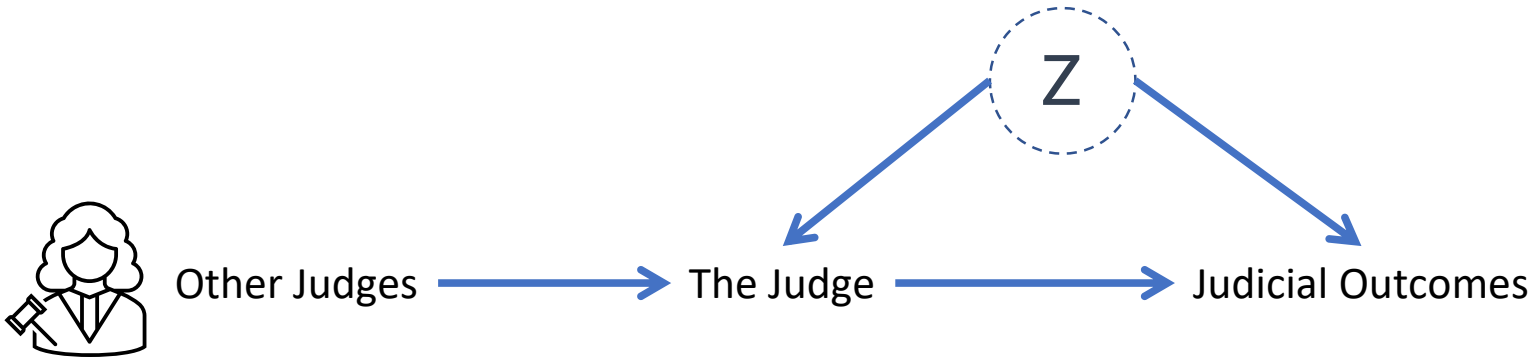
"Mendelian Randomization"

- Exogeneity
- Exclusion
- Relevance

RSM Erasmus

Where do good instruments come from?

Policies / regulations / institutional processes...
Arnold et al. (2018)

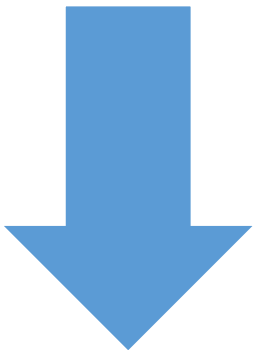


"Judge IV"

- Exogeneity: quasi-random assignments of judges to cases.
- Exclusion: other judges have no direct influence on the outcome of the focal judge's case.
- Relevance: judges share similarities.

RSM Erasmus

Finding instruments: exogeneity vs. relevance

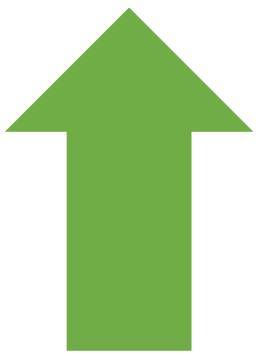


Relevance

e.g., Is weather relevant for the treatment?

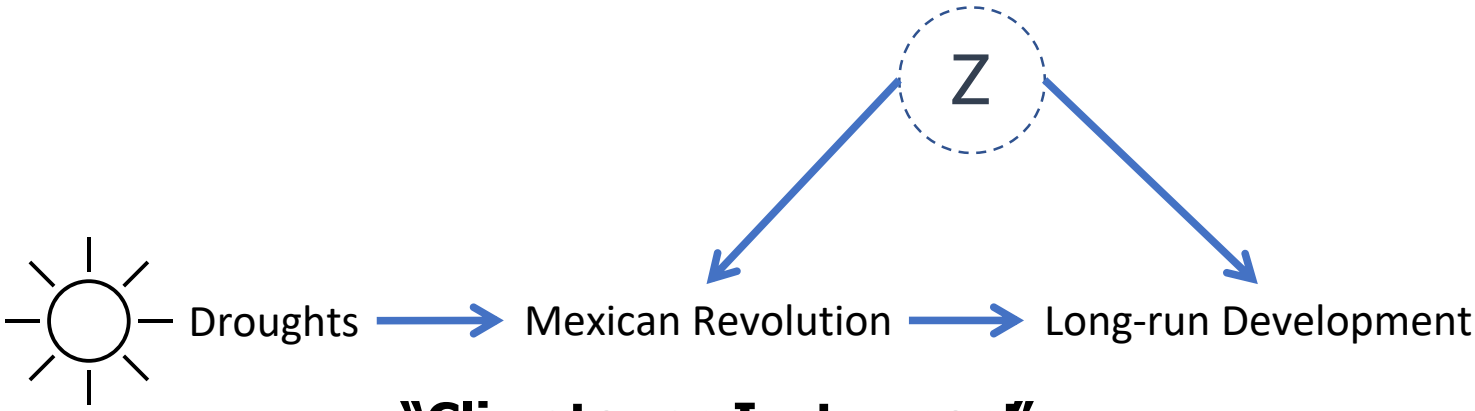
Exogeneity

e.g., Weather is for sure exogenous



Finding instruments: exogeneity vs. relevance

Using droughts as the instrument
Dell (2012)



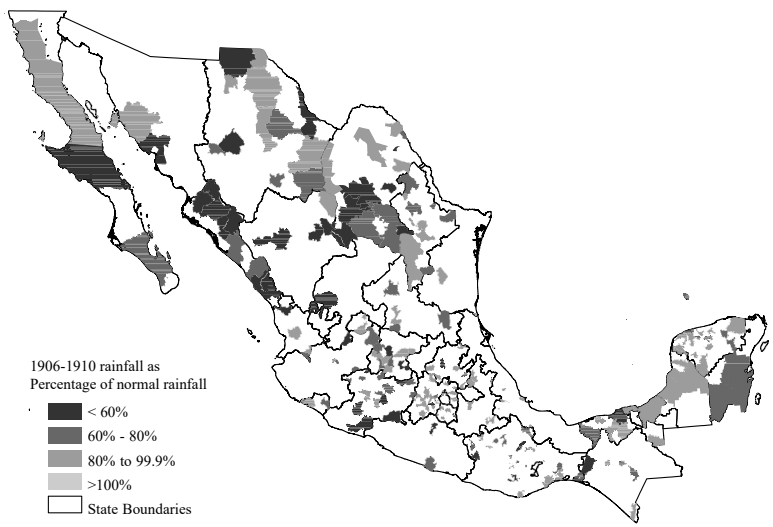
"Climates as Instruments"

- Exogeneity: climate conditions are exogenous.
- **Relevance**: Is climate driving the insurgence of revolutions?

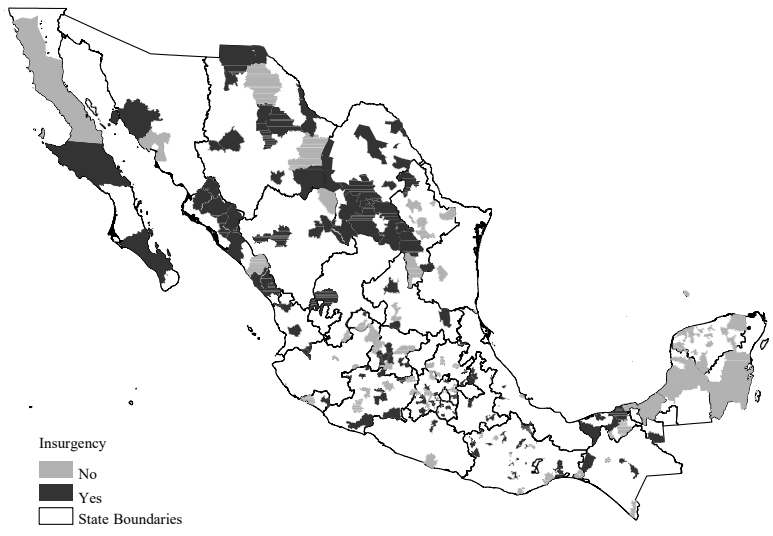
RSM Erasmus

Finding instruments: exogeneity vs. relevance

Using droughts as the instrument
Dell (2012)



(a) Droughts



(b) Revolutions

Finding instruments: a guide

01

Exogeneity is the first order problem.

02

Relevance can be partially offset with the sheer number.

- Many weak instruments

03

Exclusion should be motivated conceptually.

RSM Erasmus

References I

- Imbens, G. W., & Rubin, D. B. (2015). Causal inference in statistics, social, and biomedical sciences. Cambridge University Press.
- Imbens, G. W. (2014). Instrumental variables: an econometrician's perspective. Statistical science, 29(3), 323-358.
- Azoulay, P., Fons-Rosen, C., & Zivin, J. S. G. (2019). Does science advance one funeral at a time? American Economic Review, 109(8), 2889-2920.
- Emdin, C. A., Khera, A. V., & Kathiresan, S. (2017). Mendelian Randomization. JAMA, 318(19), 1925-1926.

RSM Erasmus

References II

- Arnold, D., Dobbie, W., & Yang, C. S. (2018). Racial bias in bail decisions. The Quarterly Journal of Economics, 133(4), 1885-1932.
- Dell, M. (2012). Path dependence in development: Evidence from the Mexican Revolution. Harvard University, mimeograph.

RSM Erasmus