Fiscal Policy and Inequality
5. Intro to Causal Inference

Elliott Ash (ashe@ethz.ch)

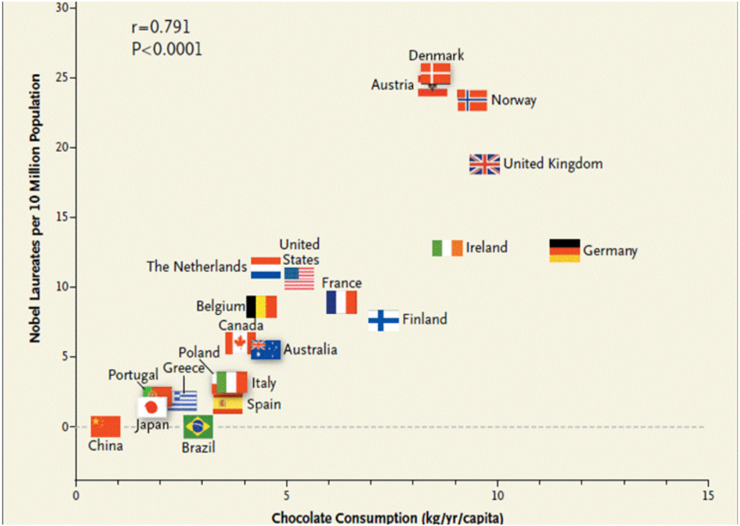
ETH Zurich

Oct 1, 2018

Causal effects

- Economists very often are motivated by **why** questions but, when they conduct their research, they tend proceed by addressing **what if** questions.
- Examples:
  - How does taking this course affect the grade that you will obtain in your master thesis?
    - Note that this is different from the predictive question: "What is the grade that students taking this course will obtain with their master thesis?"
  - If Zurich imposed a special tax on Uber drivers, how would that effect the supply of Uber rides?
  - Does the death penalty decrease crime rates?

Correlation does not imply causation



y can cause x even if y occurs after x

- In economics, this is often due to behavioral responses to an expectation on y:
  1. Rain:
     1.1 When people carry umbrellas in morning, that would predict that it will rain in the afternoon
     1.2 but subsidizing umbrellas would not increase rain
  2. Stock market might go down before an extremist party wins an election

How to estimate a causal relationship?

- In the physical sciences one can often estimate a causal relationship with a controlled experiment.
  - E.g.: Galileo's Leaning Tower of Pisa experiment
- Properties:
  - Temporal stability: the response does not change if the time when a treatment is applied is varied slightly.
  - Causal transience: the response of one treatment is not affected by prior exposure of the unit to the other treatment.
  - Unit homogeneity: units are homogeneous with respect to the treatments and responses.
- In Social Sciences none of these assumptions is plausible.

Hallsworth et al (2017)

www.hmrc.gov.uk

Dear Sir/Madam

Date of issue 4 August 2011
Reference REFERENCE NUMBER

Please pay £999999999999.99

Our records show that your Self Assessment tax payment is overdue.

It is easy to pay. Please call the phone number above to pay by debit card, credit card, or Direct Debit.

You can also pay using internet and telephone banking. For more information on when and how to pay, go to **www.hmrc.gov.uk/payinghmrc**

www.hmrc.gov.uk

Dear Sir/Madam

Date of issue 4 August 2011
Reference REFERENCE NUMBER

Please pay £999999999999.99

Our records show that your Self Assessment tax payment is overdue.

Nine out of ten people pay their tax on time.

It is easy to pay. Please call the phone number above to pay by debit card, credit card, or Direct Debit.

You can also pay using internet and telephone banking. For more information on when and how to pay, go to **www.hmrc.gov.uk/payinghmrc**

Six Different Experiment Groups

Group name	Test phrase	N
Control		17,038
Basic norm	*Nine out of ten people pay their tax on time.*	17,026
Country norm	*Nine out of ten people in the UK pay their tax on time.*	16,926
Minority norm	*Nine out of ten people in the UK pay their tax on time. You are currently in the very small minority of people who have not paid us yet.*	16,515
Gain-framed public good	*Paying tax means we all gain from vital public services like the NHS, roads, and schools.*	16,807
Loss-framed public good	*Not paying tax means we all lose out on vital public services like the NHS, roads, and schools.*	17,159

Treatment Effects by Group

Type of letter	% paid in first seven days
Control early letter	0.092
Control late letter	0.025
Difference	**0.067***** **(0.006)**
Basic norms early letter	0.099
Basic norms late letter	0.021
Difference	**0.078***** **(0.006)**
Country norms early letter	0.095
Country norms late letter	0.024
Difference	**0.071***** **(0.006)**
Minority norms early letter	0.101
Minority norms late letter	0.024
Difference	**0.078***** **(0.000)**
Gain-public early letter	0.090
Gain-public late letter	0.031
Difference	**0.059***** **(0.006)**
Loss-public early letter	0.098
Loss-public late letter	0.022
Difference	**0.076***** **(0.006)**
All letters early letter	0.096
All letters late letter	0.025
Difference	**0.071***** **(0.002)**

Basics

- Represent a **treatment** as a binary random variable $D_i = 0, 1$.
  - $D_i = 1$ is treatment, $D_i = 0$ is control.
  - e.g., receive a medicine or not
  - equals zero or one with some probability.
- Define an **outcome** Y_i for individual i.
  - e.g., life expectancy.
- Define "**potential outcomes**" (counterfactuals) as:

$$Y_i = \begin{cases} Y_{0i} & \text{if } D_i = 0 \\ Y_{1i} & \text{if } D_i = 1 \end{cases}$$

  - for an individual i, the outcome Y_i depends on treatment assignment.
- The **causal effect** of treatment for individual i is $Y_{1i} - Y_{0i}$.
  - the difference in the outcome between treatment and control.

The Causal Inference Problem

- The **causal effect** of treatment (or **treatment effect**) for individual i is $Y_{1i} - Y_{0i}$.
- The problem:
  - For i, we can observe Y_{1i} (the individual takes medicine) or Y_{0i} (the individual does not), **but not both.**
- This is the fundamental problem of causal inference.
  - The empirical methods we discuss in this course are designed to deal with this problem in the context of fiscal policy.

Illustration

		Leo	Mia
Y_{0i}	life expectancy without medicine	3	5
Y_{1i}	life expectancy with medicine	4	5

- Let's say Leo gets medicine ($D_{\text{Leo}} = 1$) and Mia does not ($D_{\text{Mia}} = 0$).
  - This is imaginary data: we would not observe Y_{0i} for Leo, and would not observe Y_{1i} for Mia

Illustration

		Leo	Mia
Y_{0i}	life expectancy without medicine	3	5
Y_{1i}	life expectancy with medicine	4	5
$Y_{1i} - Y_{0i}$	**treatment effect**	**1**	**0**

- In this imaginary data, the medicine would work for Leo, but not for Mia.

Illustration

		Leo	Mia
Y_{0i}	life expectancy without medicine	3	5
Y_{1i}	life expectancy with medicine	4	5
$Y_{1i} - Y_{0i}$	treatment effect	1	0
D_i	actual treatment assignment	1	0
Y_i	**actual health outcome**	**4**	**5**

- Note that

$$Y_{\text{Leo}} - Y_{\text{Mia}} = -1$$

  - based on these outcomes, one would be led to believe that the medicine actually harms the patient!

Defining Selection Bias

$$
\begin{aligned}
Y_{\text{Leo}} - Y_{\text{Mia}} =& Y_{1,\text{Leo}} - Y_{0,\text{Mia}} \\
= & \underbrace{(Y_{1,\text{Leo}} - Y_{0,\text{Leo}})}_{\substack{=1 \\ \text{Treat Effect on Leo}}} + \underbrace{(Y_{0,\text{Leo}} - Y_{0,\text{Mia}})}_{\substack{=-2 \\ \text{Selection Bias}}}
\end{aligned}
$$

From Individuals to Populations

- We will use expectations a lot in this course:
  - $\mathbb{E}[Y_i] = \frac{1}{n}\sum_{i=1}^{n} Y_i$, the mean or expected value for Y_i given a population of size n, where $i \in \{1, ..., n\}$
- The **average treatment effect** (ATE) is

$$\mathbb{E}[Y_{1i} - Y_{0i}] = \frac{1}{n}\sum_{i=1}^{n} Y_{1i} - \frac{1}{n}\sum_{i=1}^{n} Y_{0i}$$

Conditional Expectations

- $\mathbb{E}[Y_i|X_i = x]$, the mean or expected value of Y_i given some characteristic X_i taking a value x.
- In particular:
  - $\mathbb{E}[Y_i|D_i = 1]$, the average outcome for the treated
  - $\mathbb{E}[Y_i|D_i = 0]$, the average outcome for the controls
- The difference between treatment and control is

$$\mathbb{E}[Y_i|D_i = 1] - \mathbb{E}[Y_i|D_i = 0]$$

- Since we cannot observe the same individual with and without the medicine, this is equal to

$$\mathbb{E}[Y_{1i}|D_i = 1] - \mathbb{E}[Y_{0i}|D_i = 0]$$

Selection Bias

We have the difference in observed outcomes:

$$\mathbb{E}[Y_{1i}|D_i = 1] - \mathbb{E}[Y_{0i}|D_i = 0]$$

$$= \underbrace{\mathbb{E}[Y_{1i} - Y_{0i}|D_i = 1]}_{\substack{=1 \\ \text{Treat Effect on Treated}}} + \underbrace{\mathbb{E}[Y_{0i}|D_i = 1] - \mathbb{E}[Y_{0i}|D_i = 0]}_{\substack{=-2 \\ \text{Selection Bias}}}$$

- If there is no selection bias ($\mathbb{E}[Y_{0i}|D_i = 1] = \mathbb{E}[Y_{0i}|D_i = 0]$), the difference in observed outcomes is equal to the **average treatment effect on the treated** (ATT).

ATE versus ATT

- We have to distinguish the Average Treatment Effect (ATE)

$$\mathbb{E}[Y_{1i} - Y_{0i}]$$

  from the Average Treatment Effect on the Treated (ATT)

$$\mathbb{E}[Y_{1i} - Y_{0i}|D_i = 1]$$

- These will be the same only if there is no selection bias and the treatment effect would be the same for the control group.

Random Assignment

Random assignment $\rightarrow D_i$ independent of potential outcomes:

$$\mathbb{E}[Y_{1i}|D_i = 1] = \mathbb{E}[Y_{1i}|D_i = 0] = E[Y_{1i}]$$
$$\mathbb{E}[Y_{0i}|D_i = 1] = \mathbb{E}[Y_{0i}|D_i = 0] = E[Y_{0i}]$$

Therefore, the difference in observed outcomes

$$\mathbb{E}[Y_{1i}|D_i = 1] - \mathbb{E}[Y_{0i}|D_i = 0]$$

is equal to average treatment effect on treated (ATT),

$$\begin{aligned} &= \mathbb{E}[Y_{1i}|D_i = 1] - \mathbb{E}[Y_{0i}|D_i = 1] \\ &= \mathbb{E}[Y_{1i} - Y_{0i}|D_i = 1] \end{aligned}$$

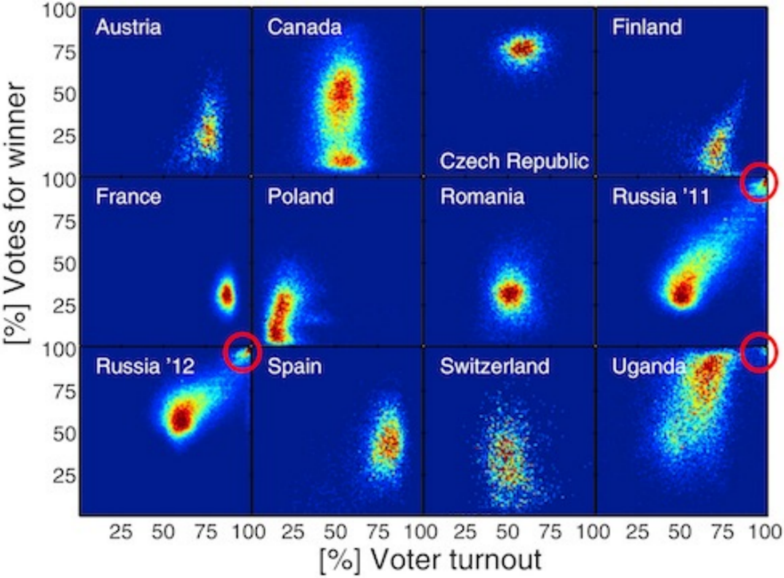
average treatment effect on non-treated,

$$\begin{aligned} &= \mathbb{E}[Y_{1i}|D_i = 1] - \mathbb{E}[Y_{0i}|D_i = 0] \\ &= \mathbb{E}[Y_{1i} - Y_{0i}|D_i = 0] \end{aligned}$$

and the population average treatment effect (ATE):

$$\mathbb{E}[Y_{1i} - Y_{0i}|D_i = 1] = \mathbb{E}[Y_{1i} - Y_{0i}|D_i = 0] = \mathbb{E}[Y_{1i} - Y_{0i}]$$

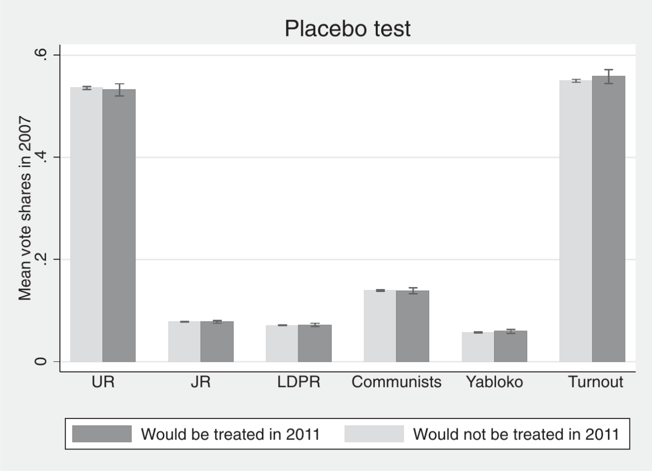
Circumstantial Evidence



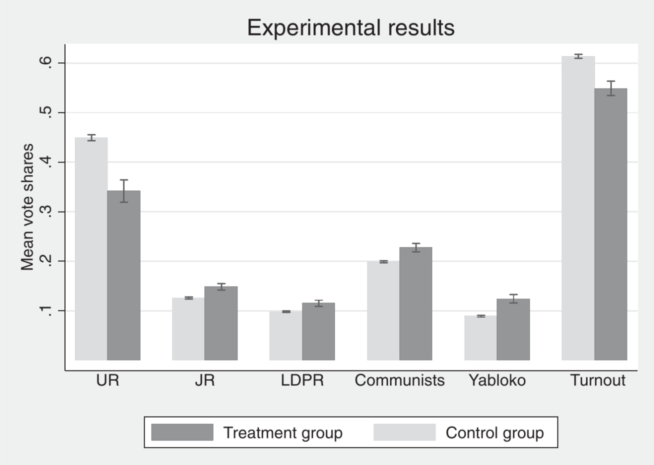
Electoral Fraud Field Experiment

- Let us rephrase slightly our question in a *treatment effects* fashion:
  - Would electoral results change if there were independent observers in the polling stations?
  - Many NGOs do this to address election fraud.
- Enikolopov, Korovkina, Petrova, Sonin and Zakharov (2013) also do this, but they randomly assign the observers to stations.
  - randomly assigned to 156 stations, out of 3,164 in Moscow, for 2011 elections

Placebo Test (effect on 2007 votes)



Main Results (Effect on 2011 votes)



Replicating Enikolopov et al

- Replication materials (data and stata code) available at sites.google.com/site/rubenenikolopov/

```
* load data
use 05_russia_data, clear

* summarize data
su _share all_observers

* party tabs
tab party

* Incumbent party vote share, by treatment status
tabstat _share if party == "United Russia", by(all_observers)
```