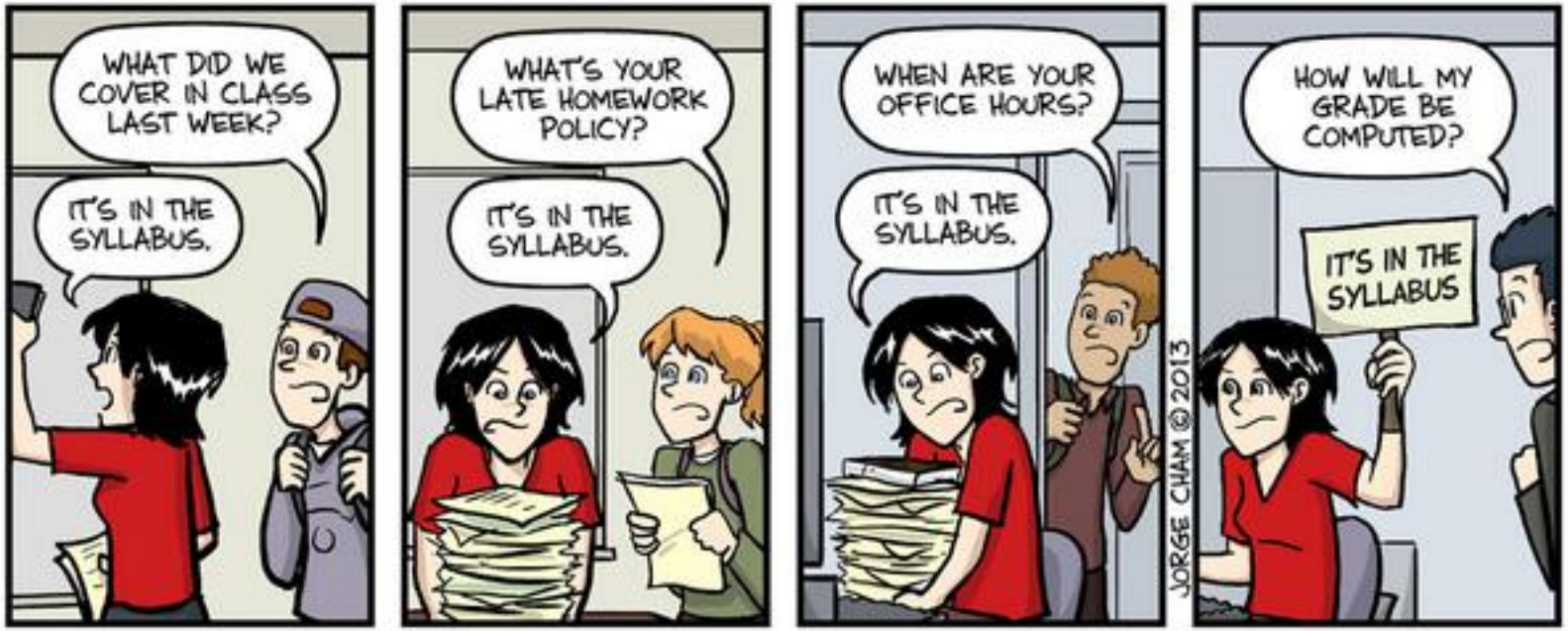
AIPP

1. Introduction and stylized facts

Correlation does not imply causation!

João Pereira dos Santos
E-mail: joao.santos@iseg.ulisboa.pt
ISEG, IZA, GLO

Everything you always wanted to know about this course and were (not) afraid to ask...



IT'S IN THE SYLLABUS

This message brought to you by every instructor that ever lived.

WWW.PHDCOMICS.COM

I am just kidding, you have my e-mail

Let me show you the main deadlines and what we will cover in this course

Kick off

You can learn more about me in
`https://sites.google.com/site/joaorpereirasantos/home`

Please send me an email with:
1. Your (econometric) background
2. What you would like to do after the MSc?
3. What would you like to learn with this course? Which topics would be more interesting for you?

Disclaimer: I borrowed many of these slides from the courses put forward by Raj Chetty, Henrik Kleven, Stefanie Stantcheva, Martin Halla, Muriel Niederle, Peter Hull, Paul Goldsmith-Pinkham, and many others

Why this course? Why are my slides written in English?

Companies like Amazon, Uber, Airbnb, Spotify,... have succeeded in solving major private market problems using technology and big data

Goal: show how same skills can be used to address important social problems and fine-tune public policies

We aim to provide an introduction to a broad range of topics, methods, and real-world applications

Start from the questions to motivate the methods rather than the traditional approach of doing the reverse

Many of these discussions are taken in English...

Why evaluate?

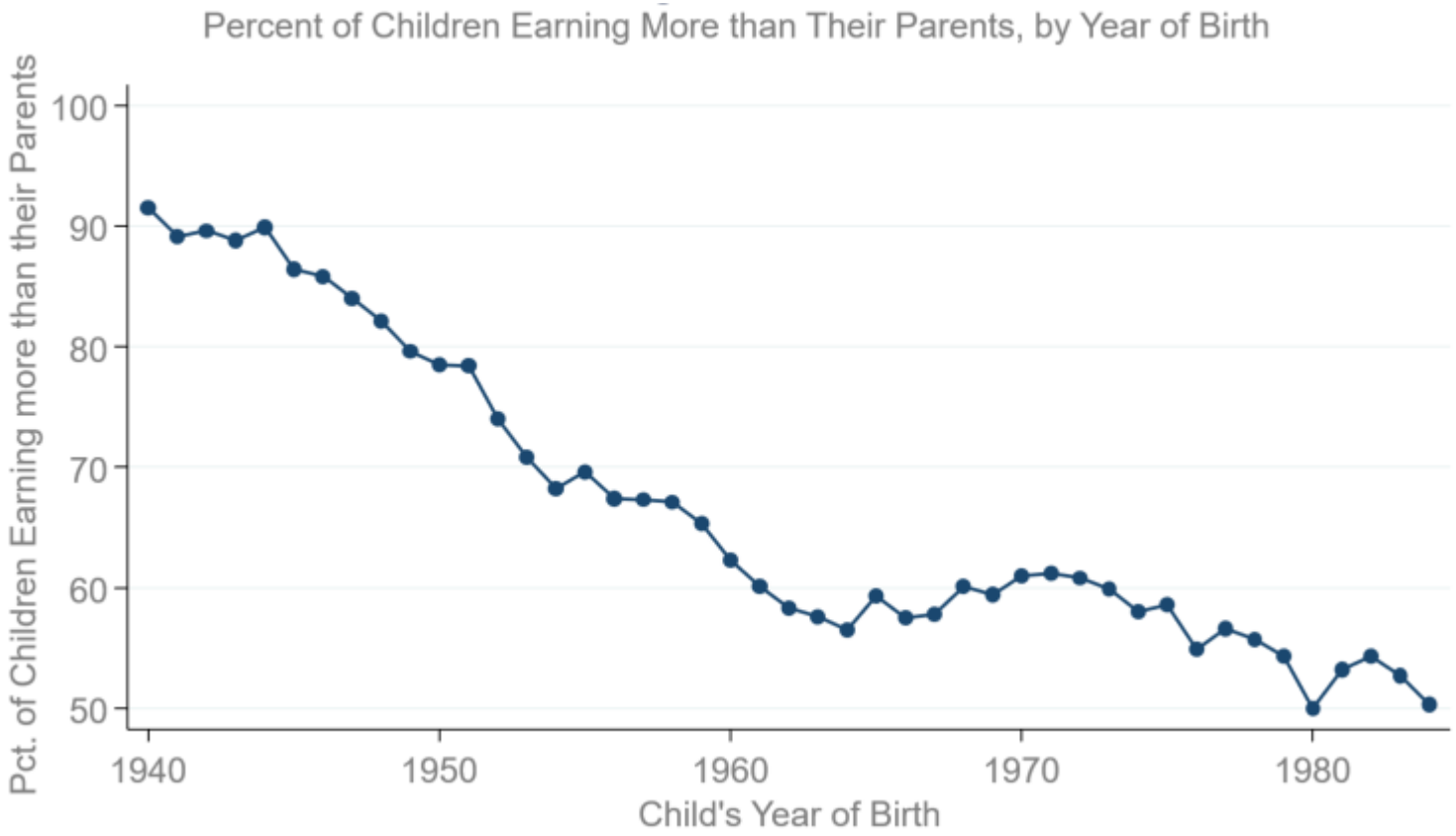
Policy is designed to change outcomes
E.g.: increase income, reduce mortality, improve test scores,...
In general: improve well-being!

Impact evaluation: Assessment of the changes in well-being of individuals that can be attributed to a particular policy
In practice: very often no impact evaluation, just a focus on inputs and immediate outputs: How much money is spent? How many textbooks are distributed? This is just monitoring!

Evidence-based policy making is part of broader agenda:
Comprises monitoring and evaluation
Shift in focus from inputs to outcomes & results

Goal: Enhance accountability, inform budget allocation, and guide policy decisions

The Fading American Dream? (Chetty et al., 2017 Science)



Why?

Central policy question: why are children's chances of climbing the income ladder falling in America?

And what can we do to reverse this trend. . . ?

Difficult to answer this question based solely on historical data on macroeconomic trends

Numerous changes over time make it hard to test between alternative explanations

Economics WAS a theoretical social science

Until recently, social scientists have had limited data to study policy questions like this

Social science has therefore been a theoretical field:

- Develop mathematical models (economics) or qualitative theories (sociology)

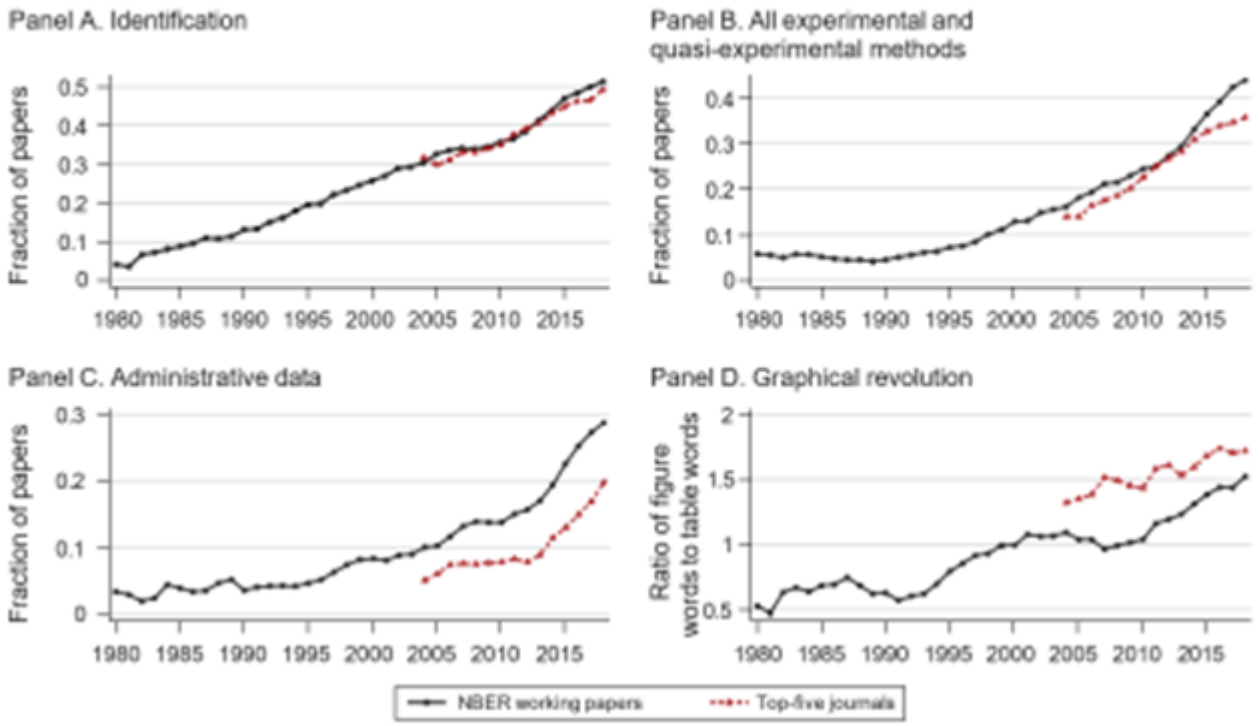
- Use these theories to explain patterns and make policy recommendations, e.g. to improve upward mobility

Problem: theories untested $\Rightarrow$ politicization of questions that in principle have scientific answers

"In God we trust, all others must bring data."


Economists: everything the light touches is our kingdom
law
public health
statistics
history
computer science
political science
math
geography
psychology
sociology

The Rise of Data and Empirical Evidence (Currie, Kleven, Zwiers, 2020 AER:P&P)



Social Science in the Age of Big Data

Large datasets are starting to transform social science, as they have transformed business

Examples:

- Government/ Administrative data: (historical) population census, tax records, linked employer-employee, balance sheet, housing transactions, invoice transactions, voting results, standardised national exams, ...

- Corporate data: Google, Uber, retailer data, ...

- Unstructured data: Twitter (recently, not so easy), newspapers, maps, ...

Why is Big Data Transforming Social Science?

1. Test theories with far greater reliability than with surveys (less selection problems)

2. Heterogeneity: Universal coverage allows us to "zoom in" to subgroups

3. Ability to measure new variables and combine knowledge from different perspectives (e.g., night-time light satellite imagery intensity, emotions, ...)

4. Better methodologies that can approximate scientific experiments and estimate causal parameters

What do Economists bring to the table?

An appreciation for the study of behavior, incentives, expectations, and constraints

An appreciation for the importance of equilibrium interactions and their often non-trivial unintended consequences

An appreciation for the importance of identification, causal statements, and counterfactuals

An appreciation for the fact that all participants maximize welfare, which makes it difficult to change an equilibrium

Why do we use Econometrics?

Statistical toolkit that economists use to answer questions with data

Examples of economic questions:

Has economic inequality increased since 1960?
Descriptive Q: asks about how things are (or were) in reality

How do increases in the minimum wage affect employment?
Causal Q: asks what would have happened in a counterfactual

What will the unemployment rate be next quarter?
Forecasting Q: asks what will happen in the future

We will focus on descriptive and causal questions in this course

Econ posits some underlying models that generate the data we see; to do science/policy, we'd like to know model parameters

Causality is an old philosophical question

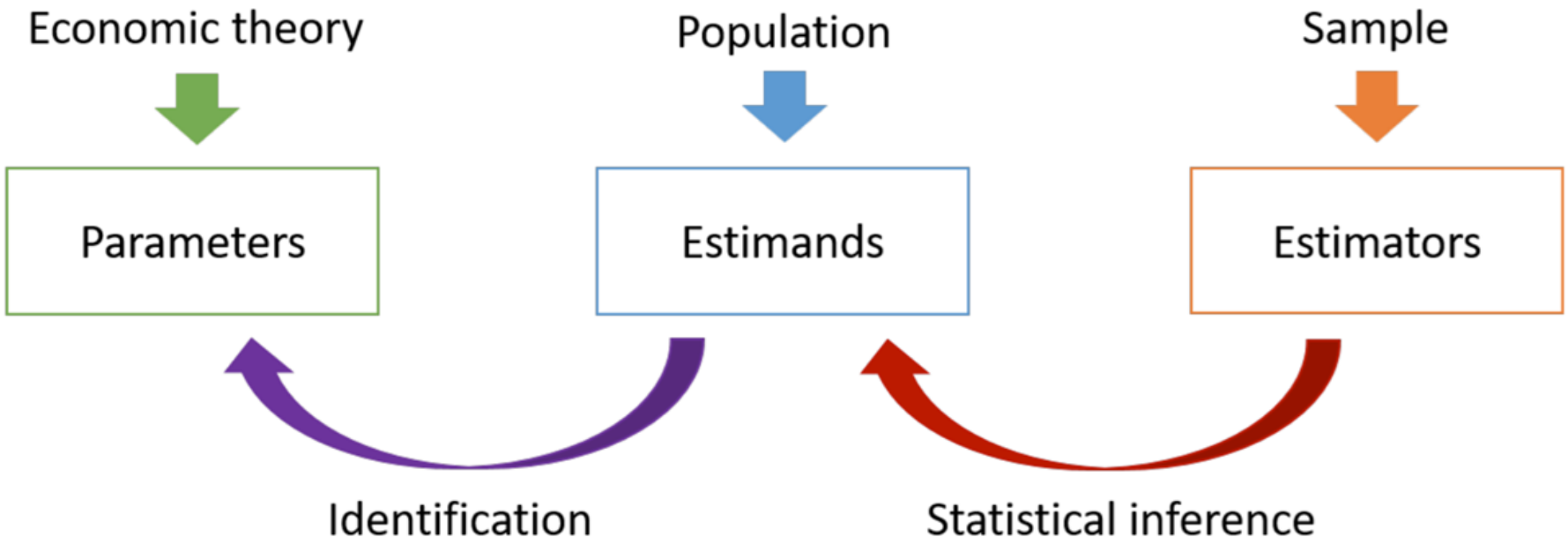
I would rather discover one causal law than be King of Persia.
— Democritus

We have knowledge of a thing only when we have grasped its cause.
— Aristotle, Posterior Analytics

We do not have knowledge of a thing until we have grasped its why, that is to say, its cause.
— Aristotle, Physics

We may define a cause to be an object followed by another, and where all the objects, similar to the first, are followed by objects similar to the second. Or, in other words, where, if the first object had not been, the second never had existed.
— David Hume (1748, Section VII)

The big picture



Key Concepts for This Course

Sample vs Population: the data we actually observe

Target Parameter: what we actually want to know about the population. In...

$$Y_{it} = \alpha_i + \gamma X_{it} + W'_{it}\beta + \epsilon_{it} \tag{1}$$

What happens to Y if we move/ implement X, after controlling for everything that matters (W)? The answer is γ!

When we don't know, we can use data to estimate γ: Estimator vs Estimand

Why is it Hard to Answer these Questions with Surveys?

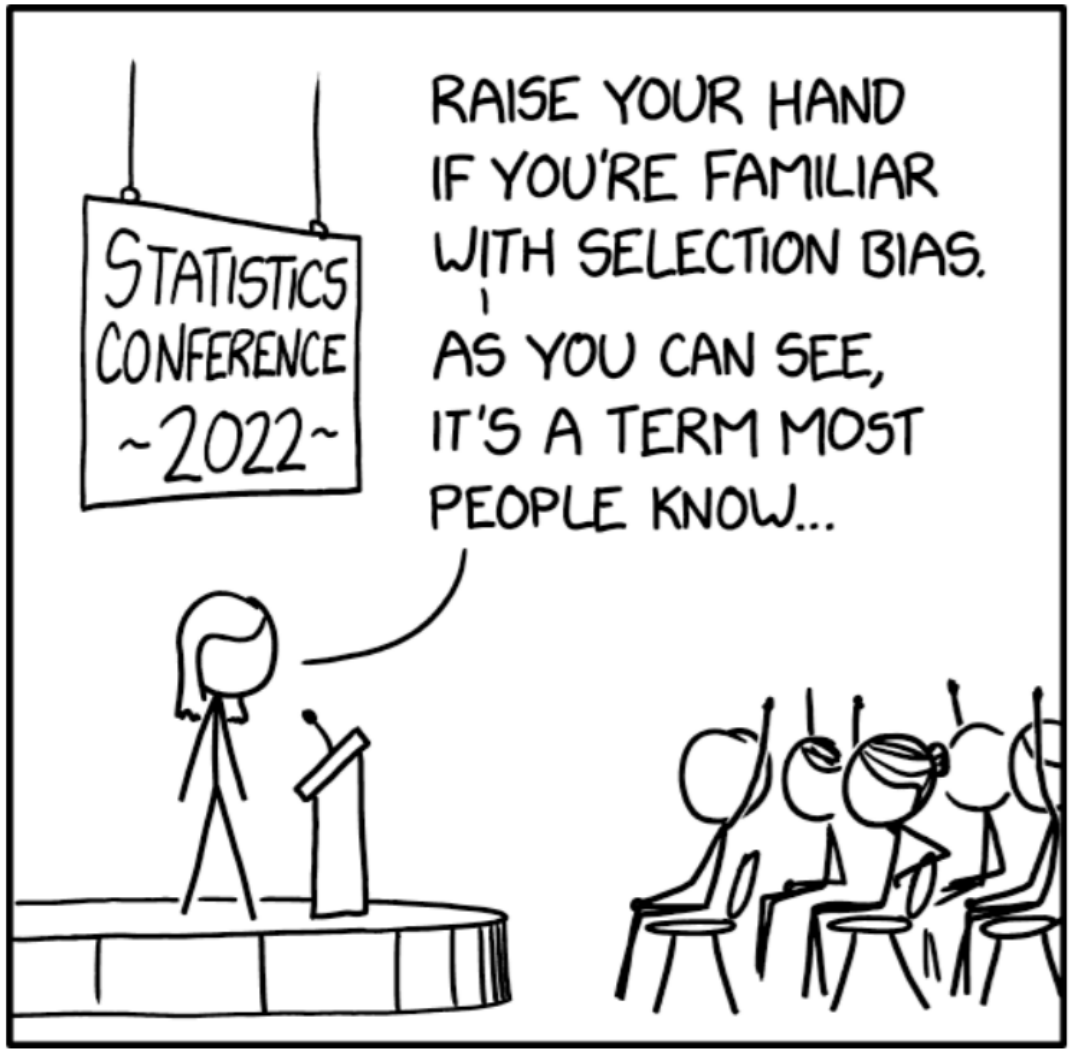
We often only observe a sample from the population we want to describe
E.g.: we want to know how the distributions of X or Y have changed. But we only observe them for a survey of workers

Best case scenario: Sample is randomly selected
E.g. surveyed workers drawn out of hat. If random, just need to account for the fact that the sample might have different characteristics from the population by chance: use weighting matrices

Worst case scenario: Sample is not representative of the population
E.g. certain types of individuals are more likely to respond to surveys

In 1948, the Chicago Tribune writes that Thomas Dewey defeats Harry Truman in the 1948 presidential election, based on a survey of voters. But their survey was conducted by phone. In 1948, only rich people had phones

Selection Bias



Why is it Hard to Answer Causal Questions even if we have Data on the Population?

Causal questions involve both a descriptive component (how are things in reality?) and a counterfactual component (how would things have been under different circumstances?)

E.g.: what is the effect on your earnings (or other outcomes) from going to ISEG instead of going to Harvard?
Descriptive: how much do ISEG students earn after graduation?
Counterfactual: how much would ISEG students have earned if they had instead gone to Harvard?

Counterfactuals can't be learned from seeing data alone. We need additional assumptions to learn about them! It involves "ceteris paribus" thinking

Is it a missing data problem? "Not everything that matters can be measured" (Einstein)

One classic example

Becker (1957)'s theory of human capital hypothesizes a positive causal effect of going to college on later-life earnings

Formalize this using potential outcomes notation (Rubin model)
Let D_i be a "dummy" variable for the college "treatment": D_i = 1 (0) if individual i went (did not go) to college

For each i, we imagine two potential earnings outcomes:
$Y_i(1)$ = outcome if "treated" (earnings if i went to college)
$Y_i(0)$ = outcome if "untreated" (earnings if i didn't go)
$Y_i(1) - Y_i(0)$: i's "treatment effect" (returns to college)

Observed earnings is $Y_i(1)$ if $D_i = 1$ and $Y_i(0)$ if $D_i = 0$

Our target parameter is the average effect: E[$Y_i(1)$-$Y_i(0)$]

Suppose we have a **sample** of earnings and schooling: (Y_i, D_i), $i = 1, \ldots, N$

Take as our **estimator** the difference in sample average earnings for the N_1 people who did go to college vs. the N_0 people who didn't go to college:

$$\underbrace{\frac{1}{N_1} \sum_{i:D_i=1} Y_i}_{\text{Avg. college earnings } \textit{in sample}} - \underbrace{\frac{1}{N_0} \sum_{i:D_i=0} Y_i}_{\text{Avg. no-college earnings } \textit{in sample}}$$

Our **estimand** is the corresponding difference in population means:

$$\underbrace{E[Y_i|D_i = 1]}_{\text{Avg. college earnings } \textit{in population}} - \underbrace{E[Y_i|D_i = 0]}_{\text{Avg no-college earnings } \textit{in population}}$$

Compare this to our **parameter**: average treatment effect $E[Y_i(1) - Y_i(0)]$

$$\underbrace{E[Y_i(1)]}_{\text{Avg. } \textit{potential} \text{ college earnings}} - \underbrace{E[Y_i(0)]}_{\text{Avg. } \textit{potential} \text{ no-college earnings}}$$

From Estimator to Estimand

Statistical inference in large (enough) samples:

- Sample means are "close to" population means (by the Law of Large Numbers)

- Deviations between sample and population means tend to follow a known distribution (by the Central Limit Theorem)

- We can use these facts to make inferences about population means from observed sample means (e.g. form 95% CIs)

From Estimand to Parameter

Suppose we were somehow able to see the entire population

People who do/don't go to college (D_i) have different potential earnings with/without college due to:
Differences in (academic) ability, personality, family background, career goals (ambition), etc.
Sometimes referred to as omitted variable bias (OVB) or confounding factors. If they are correlated with both Y and X, then we have an endogeneity problem!

Identification challenge: Only observe $Y_i(1)$ among those with $D_i = 1$ and $Y_i(0)$ among those with $D_i = 0$; counterfactual Y_i(d)'s are unobserved

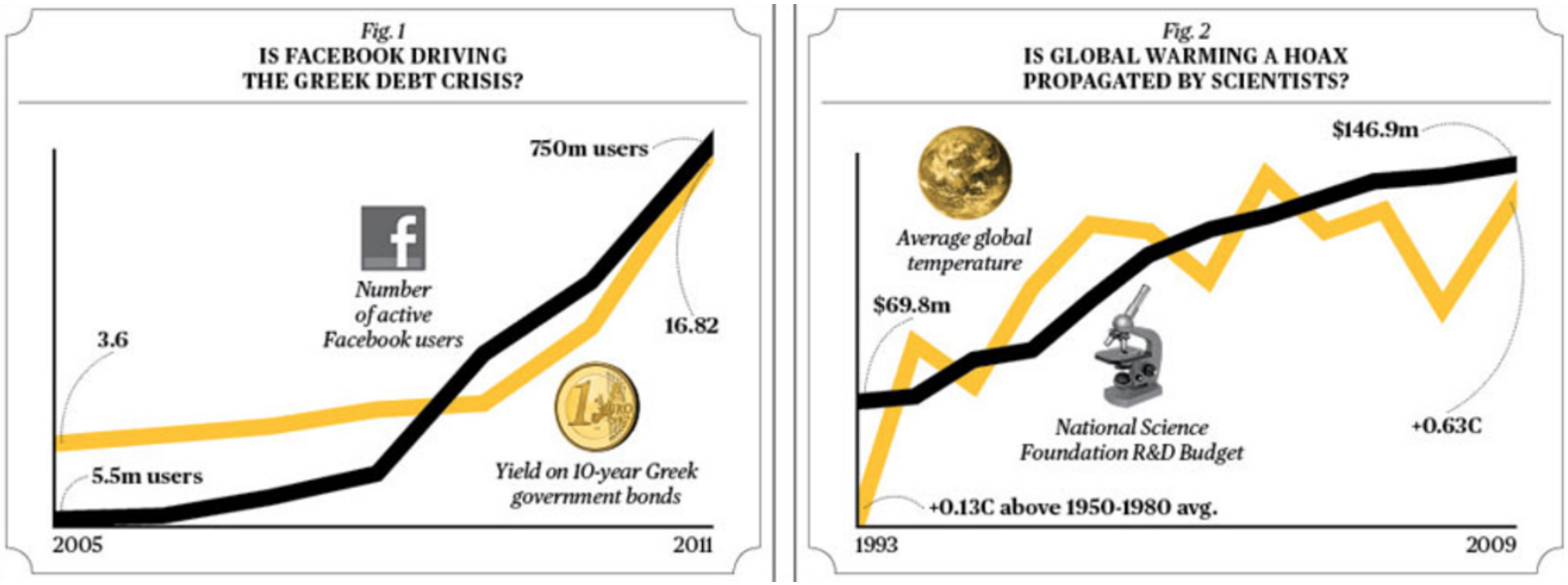
Side note: In Psychology, counterfactual thinking is a term to describe the tendency to imagine alternatives to reality

More on this from a Nobel Prize winner

https://www.youtube.com/watch?v=iPBV3BlV7jk&list=PL-uRhZ_p-BM5ovNRg-0
index=2

https://www.youtube.com/watch?v=6YrIDhaUQOE&list=PL-uRhZ_p-BM5ovNRg-0
index=3

Other source of endogeneity



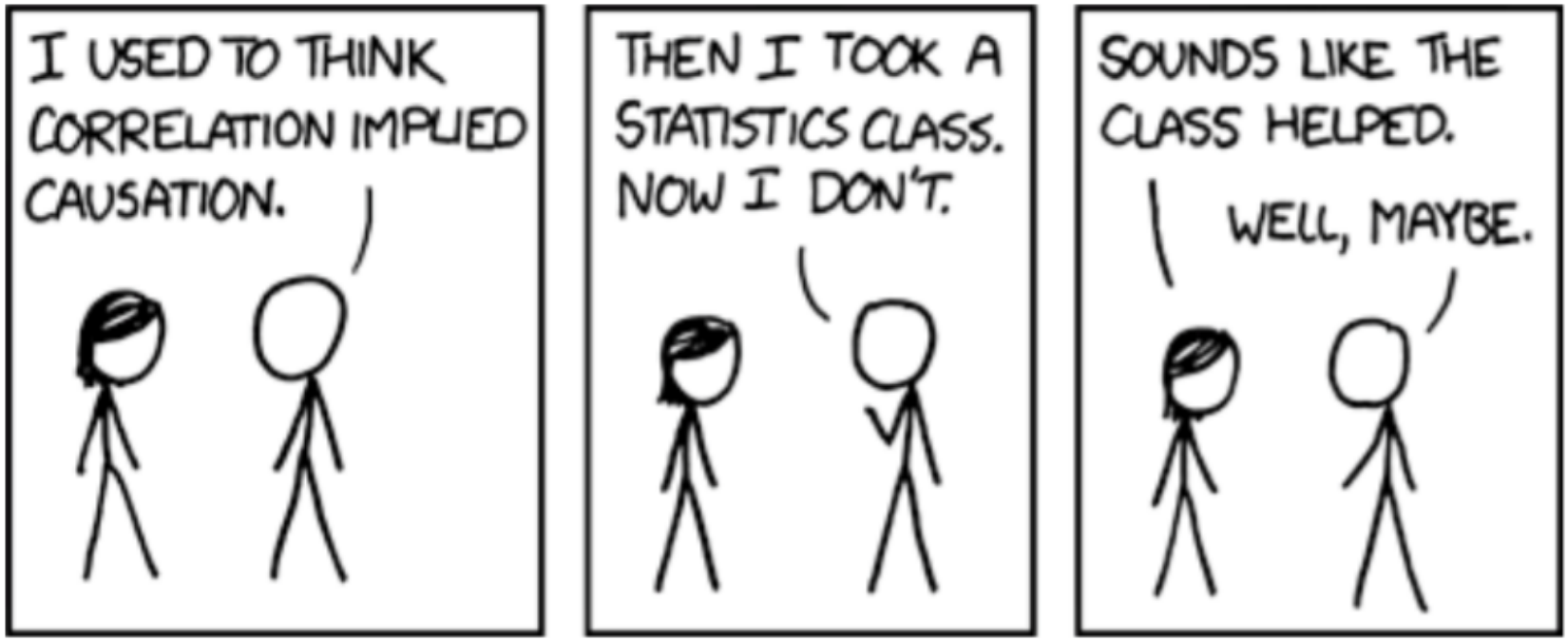
On top of selection bias and OVB, Simultaneity!

Classic example: monthly shark attacks and ice cream sales

Funnier examples in `https://www.tylervigen.com/spurious-correlations`

(but I am not so sure about the Nicholas cage example...)

Before-after conclusions can be very tricky!



https://x.com/KhoaVuUmn/status/1959614687147860025

Some historical perspective: "Let's take the con out of econometrics" Leamer (1983)

Econometricians would like to project the image of agricultural experimenters who divide a farm into a set of smaller plots of land and who select randomly the level of fertilizer to be used on each plot. If some plots are assigned a certain amount of fertilizer while others are assigned none, then the difference between the mean yield of the fertilized plots and the mean yield of the unfertilized plots is a measure of the effect of fertilizer on agricultural yields. The econometrician's humble job is only to determine if that difference is large enough to suggest a real effect of fertilizer, or is so small that it is more likely due to random variation.

This image of the applied econometrician's art is grossly misleading. I would like to suggest a more accurate one. The applied econometrician is like a farmer who notices that the yield is somewhat higher under trees where birds roost, and he uses this as evidence that bird droppings increase yields. However, when he presents this finding at the annual meeting of the American Ecological Association, another farmer in the audience objects that he used the same data but came up with the conclusion that moderate amounts of shade increase yields. A bright chap in the back of the room then observes that these two hypotheses are indistinguishable, given the available data. He mentions the phrase "identification problem," which, though no one knows quite what he means, is said with such authority that it is totally convincing. The meeting reconvenes in the halls and in the bars, with heated discussion whether this is the kind of work that merits promotion from Associate to Full Farmer; the Luminists strongly opposed to promotion and the Aviophiles equally strong in favor.

What can we do then to evaluate public policy?

One possibility: Identification through Randomization

The gold standard is a randomized controlled trial (RCT), aka an experiment (e.g., vaccine vs placebo; A-B tests in marketing). Suppose that we were somehow able to randomize who went to college

This makes D_i statistically independent of ($Y_i(0)$,$Y_i(1)$): same potential outcomes

Randomization eliminates selection bias:
E[$Y_i \mid D_i = 1$]-E[$Y_i \mid D_i = 0$]
= E[$Y_i(1) \mid D_i = 1$] - E[$Y_i(0) \mid D_i = 0$] (in potential outcomes)
= E[$Y_i(1)$]-E[$Y_i(0)$]
= E[$Y_i(1)$-$Y_i(0)$]

https://www.youtube.com/watch?v=eGRd8jBdNYg&list=PL-uRhZ_p-BM5ovNRg-0
index=4

Technical note: SUTVA

The Stable Unit Treatment Value Assumption implies no unmodeled spillovers, i.e., potential outcomes for a given observation respond only to its own treatment status; potential outcomes are invariant to random assignment of others

One could write out potential outcomes in a more extensive fashion, taking into account both one's own treatment status and the treatment status of other types of units
E.g., housemates, friends, relatives, neighbors, competitors, etc.

Implications for standard errors (cluster? bootstrap?)

Potential violations of SUTVA: spillovers

Contagion: the effect of being vaccinated also depends on other's being vaccinated

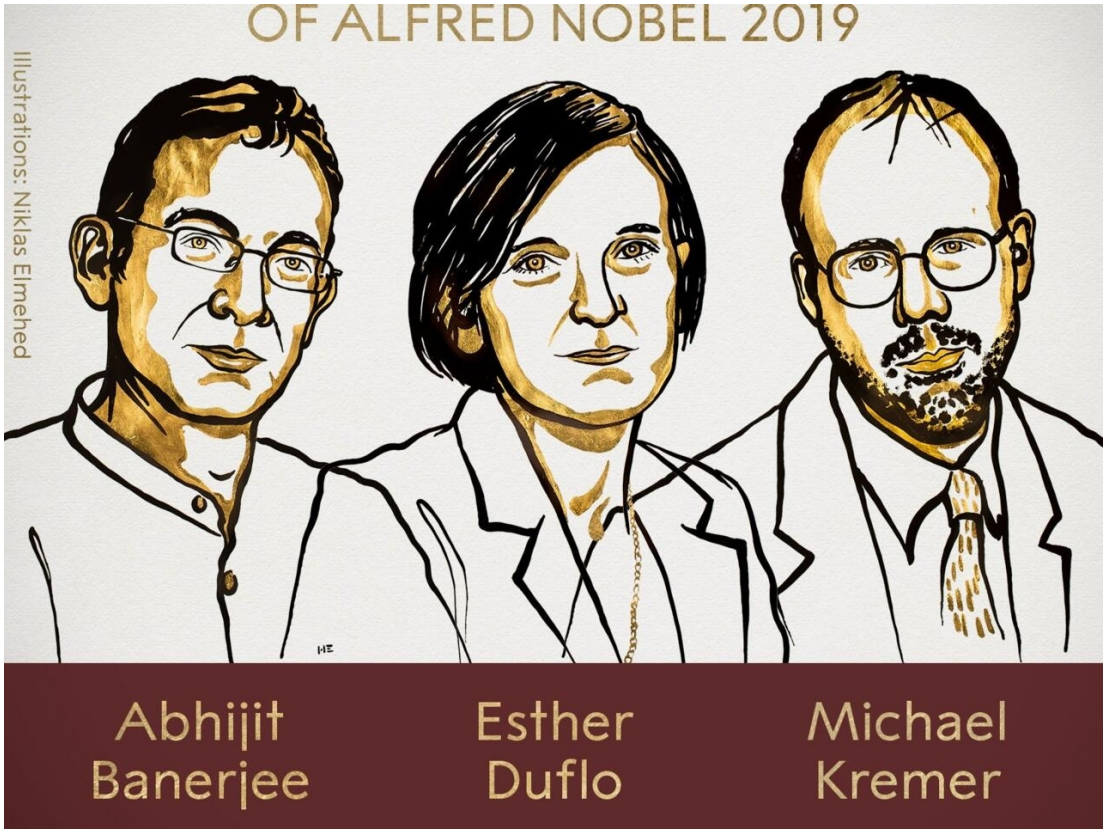
Displacement: interventions designed to affect one location may have effects on other locations (e.g., crime)

Communication from treated to control

Social comparison and Signalling

Anticipation effects

All these effects can be arguably more important when we scale interventions!



Abhijit Banerjee, Esther Duflo, and Michael Kremer have introduced a new approach to obtaining reliable answers about the best ways to fight global poverty. It involves dividing this issue into smaller, more manageable, questions. Since the mid-1990s, they have been able to test a range of interventions in different areas using field experiments.

What to do when we cannot randomize?

The Credibility Revolution in Empirical Economics: How Better Research Design Is Taking the Con out of Econometrics

Joshua D. Angrist

Jörn-Steffen Pischke

JOURNAL OF ECONOMIC PERSPECTIVES
VOL. 24, NO. 2, SPRING 2010
(pp. 3–30)

Download Full Text PDF
(Complimentary)

Article Information | Comments (0)

Abstract

Since Edward Leamer's memorable 1983 paper, "Let's Take the Con out of Econometrics," empirical microeconomics has experienced a credibility revolution. While Leamer's suggested remedy, sensitivity analysis, has played a role in this, we argue that the primary engine driving improvement has been a focus on the quality of empirical research designs. The advantages of a good research design are perhaps most easily apparent in research using random assignment. We begin with an overview of Leamer's 1983 critique and his proposed remedies. We then turn to the key factors we see contributing to improved empirical work, including the availability of more and better data, along with advances in theoretical econometric understanding, but especially the fact that research design has moved front and center in much of empirical micro. We offer a brief digression into macroeconomics and industrial organization, where progress – by our lights – is less dramatic, although there is work in both fields that we find encouraging. Finally, we discuss the view that the design pendulum has swung too far. Critics of design-driven studies argue that in pursuit of clean and credible research designs, researchers seek good answers instead of good questions. We briefly respond to this concern, which worries us little.

Many of the big questions in the social sciences deal with cause and effect. Some of these questions are possible to answer using natural experiments, in which chance events or policy changes result in groups of people being treated differently. Using natural experiments, David Card has analysed the labour market effects of minimum wages, immigration and education. The results showed, among other things, that increasing the minimum wage does not necessarily lead to fewer

jobs. Joshua Angrist and Guido Imbens showed what conclusions about causation can be drawn from natural experiments in which people cannot be either forced or forbidden to participate in the programme being studied.

Internal vs External validity

Internal validity: validity of a causal effect in a given setting
Degree to which an evaluation establishes the cause-and-effect relationship

External validity: generalizing from this causal effect to some other situation of interest
Informative for (i) larger or different pop, (ii) different time
"Will results obtained somewhere generalize elsewhere?" (Duflo)

Research: more vocabulary

First step: Formulate a clear study question, with a cause-and-effect focus

The results chain sets out the sequence of inputs, activities, and outputs that are expected to improve outcomes

Inputs: Resources (staff & budget)
Activities: Actions taken to convert inputs into outputs
Outputs: Tangible goods and services produced
Outcomes: Use of outputs by targeted population, typically achieved in the short-to-medium term

Results chain: an example

Consider that the Ministry of Education aims to introduce a new approach to teach math in high schools

Inputs: Staff from the ministry, high school teachers, federal budget, and the municipal training facilities

Activities: Designing new curriculum; teacher training program; commissioning, printing, and distributing textbook

Outputs: No. of teachers trained, no. of textbooks delivered, and the adaptation of standardized tests

Outcomes: Teachers' use of the new methods & textbooks; application of new tests
In the Medium-run: Improvements in student performance on the standardized mathematics tests
In the Long-run: Increased high school completion rates; higher employment rates and earnings for graduates

Formulate hypothesis and select indicators

Consider the example from above:
Is the new curriculum superior to the old one in imparting knowledge of mathematics?
Are Trained teachers using the new curriculum? In a more effective way than other teachers?

Good indicators (not just outputs!) are SMART
Specific: measure the info required as closely as possible
Measurable: information can be readily obtained
Attributable: indicator is linked to the policy's efforts
Realistic: data can be obtained in a timely fashion, with reasonable frequency, and at reasonable cost
Targeted: to the objective population

Think about this before the exam

https://www.youtube.com/watch?v=Da0FUc5miAk&t=52s