Lecture 1: Introduction

Jaakko Meriläinen

5304 Econometrics @ Stockholm School of Economics

Why Is It Important That You Are Here?

- Econometrics focuses on statistical methods that are used to study economic phenomena
- The use of data to study economic phenomena is central to being an economist
- And so you need to understand—and be able to apply—the econometric tools that economists use
- But the use of these methods is also more general than just economics research
  - These skills are useful across a range of professions
  - Also, to make sense of empirical claims that are commonplace across sectors, in newspapers, in politics...

Plan for This Lecture

1. Introduction to the course
    - Objectives
    - Administration and logistics
2. Types of data and empirical questions
3. Introduction to potential outcomes framework
    - ATE and ATT
    - Selection bias
4. Random assignment as a solution
    - Bloom et al. (2013) on the importance of management practices
    - The Oregon Health Insurance Experiment
5. Internal and external validity

This Course: Objectives

- The goals for this course are for students to:
  - **Understand current empirical research methods**
    - To understand commonly-used econometric tools
    - To be able to read primary economic research directly
    - To be able to critique statistical methods employed in papers
  - **To be able to design empirical analyses**
    - How would you go about designing econometric analyses on a given research question?
  - **To be able to execute empirical research**
    - How to use statistical software to analyze data
    - How to interpret and present results
- Thus, this is primarily an applied econometrics course with requisite bits of theory

This Course: Prerequisites

- This course builds upon undergraduate training in econometrics
  - For instance, course 651 at SSE, built on Wooldridge's *Introductory Econometrics*
- We will cover some of this material again but faster than in an undergraduate-level econometrics course
  - If you need support, reach out!
- We will be extending this in several dimensions
  - A deeper look at topics you've seen before
  - Some new topics
  - A lot more direct engagement with economic research and data (prior exposure to Stata or R is helpful but not required)
- Our focus will be primarily on measuring causal effects (more on this later)

This Course: Course Material

- Cunningham (2021): *Causal Inference: The Mixtape*
  - A graduate-level text book, close to the level of this course
  - Freely available online **here**
  - All chapters come with code in Stata, R, and Python
- Angrist and Pischke (2009): *Mostly Harmless Econometrics: An Empiricist's Companion*
  - A graduate-level text book, one step more advanced than The Mixtape
  - Assumes previous grounding in econometrics, probability theory, matrix algebra
- Wooldridge (2015): *Introductory Econometrics: A Modern Approach*
  - An undergraduate text book, useful for the first engagement with core topics
  - However, it is not going to be sufficient for this class
- I will supplementing readings throughout with additional papers and online resources—and lecture notes will also be available on Canvas

This Course: Grading

- Grades for this course will be aggregated across the following different subcomponents:
  - 5 assignments, dispersed through the semester (20%)
  - Referee report and presentation in a sub-group (10%)
  - Final exam (70%)
  - Summary assignment due in January 2024 (0% but compulsory)

- More details on the final exam will follow later...
  - **All materials** covered in the course are up for inclusion
  - Similar questions as in previous years (application-centric)
  - Focus of the exam is on understanding methods, not memorizing proofs!

This Course: Lectures, Seminars, and Office Hours

- **Lectures:** This is where I will be teaching methods
  - You should try to attend lectures regularly
  - Lecture materials will be made available in advance—it is advisable to have a look at them beforehand
- **Seminars:** This is where you learn how to apply methods
  - TA: Petter Berg (3rd-year PhD student at SSE)
  - Direct engagement with real-world (messy) data
  - Also where the subgroup presentations etc. happen
  - Dedicated time each week for answering queries
- **Office hours:** I will set aside 1 hour every week
  - If you are feeling overwhelmed/not challenged, come to these!
  - I am also happy to discuss things not directly related to the course, such as thesis ideas or graduate studies
- **Email:** You can always reach me at `jaakko.merilainen@hhs.se`

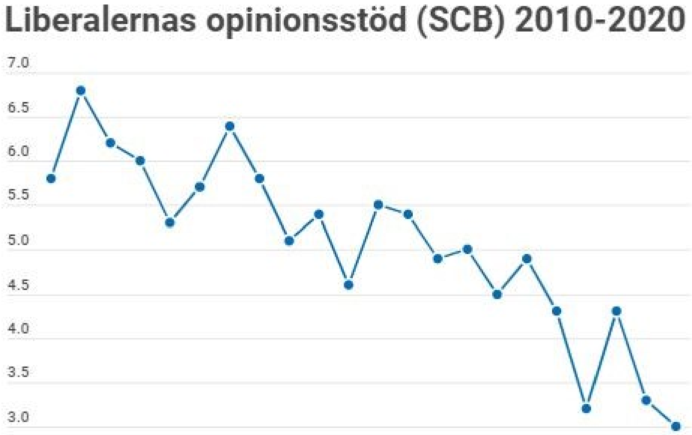
This Course: A Preliminary Road Map

Topic	Lectures	Seminars	Problem set
Introduction	1	1	1 (ungraded)
Regression	2-5	1, 2	2
Instrumental Variables	6-8	3	3
Regression Discontinuity	9, 10	4	4
Panel Data	11-13	5,6	5,6
Design of Experiments	14	7	—
LDV/Review	15	–	—

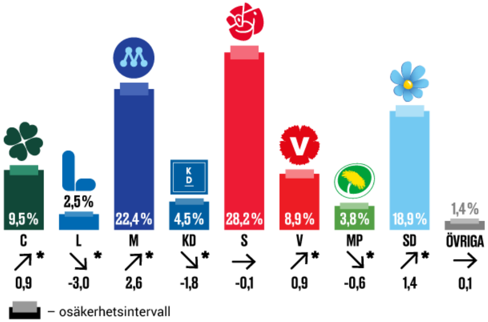
Types of Data

- **Time series**
  - Data for a single entity (person, firm, country) collected at multiple time periods
- **Cross-sectional data**
  - Data for multiple entities (workers, firms, countries) collected at a single time period
- **Panel data**
  - Data for multiple entities, observed two or more times
- This course will focus on cross-section and panel data analysis

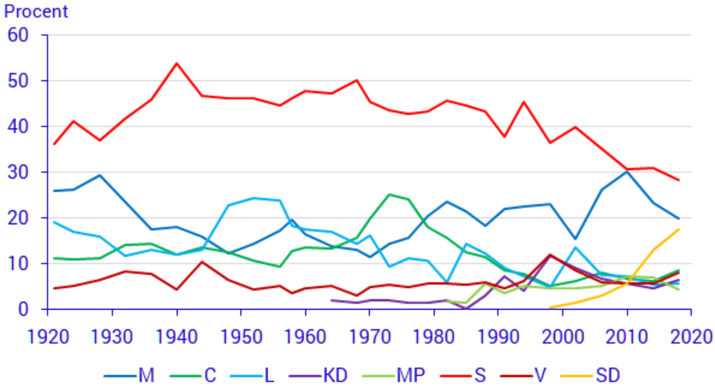
Time Series: Popularity of the Swedish Liberal Party Over Time



Cross-Section: Popularity of Swedish Parties in May 2021



Panel: Popularity of Swedish Parties Over Time



Causal and Non-Causal Questions

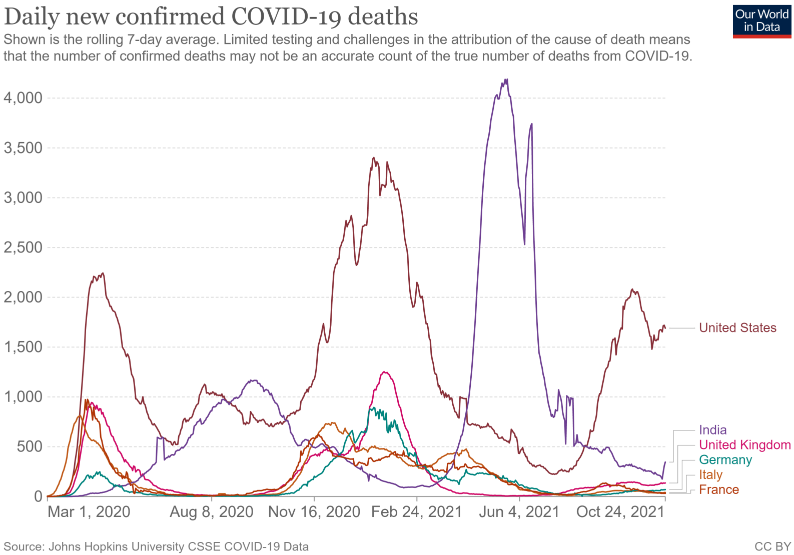
- Our core focus is going to be on **estimating causal effects**
  - What would happen to an outcome variable (Y) if we changed a variable/treatment (X)?
  - For instance, what would be the effect of raising the education of an individual by a year on wages?
  - Focus of most empirical applied microeconomics research
  - Closely related to the notion of *ceteris paribus* that you are familiar with

- Angrist and Pischke in *Mostly Harmless Econometrics*: "... we believe that the most interesting research in social science is about cause and effect."

- These are, however, not the only types of questions we should be asking!

- Non-causal questions can also be interesting and/or important:
  - Description of economic (and other societal) phenomena
  - Forecasting and prediction

Descriptive Question: COVID-19 Mortality



Descriptive Question: Trends in Income Inequality

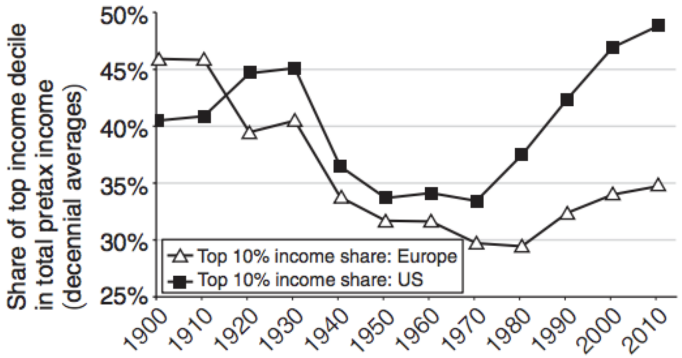
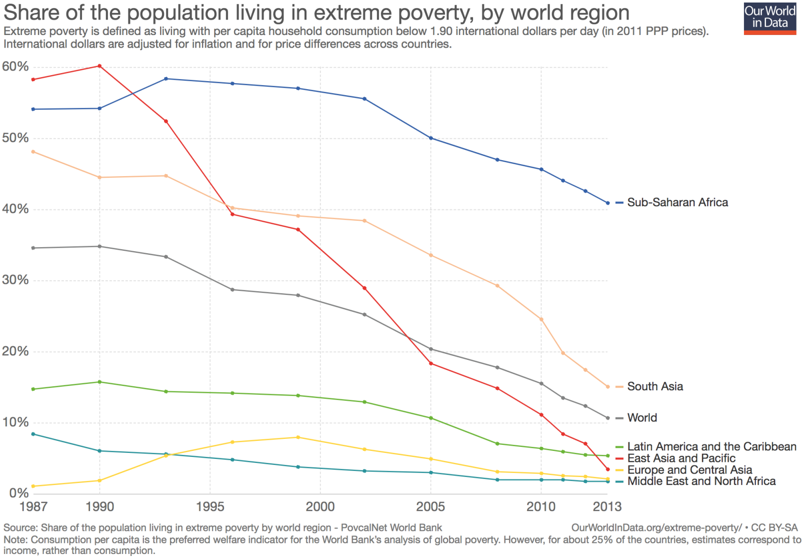


FIGURE 1. INCOME INEQUALITY: EUROPE AND THE UNITED STATES, 1900–2010

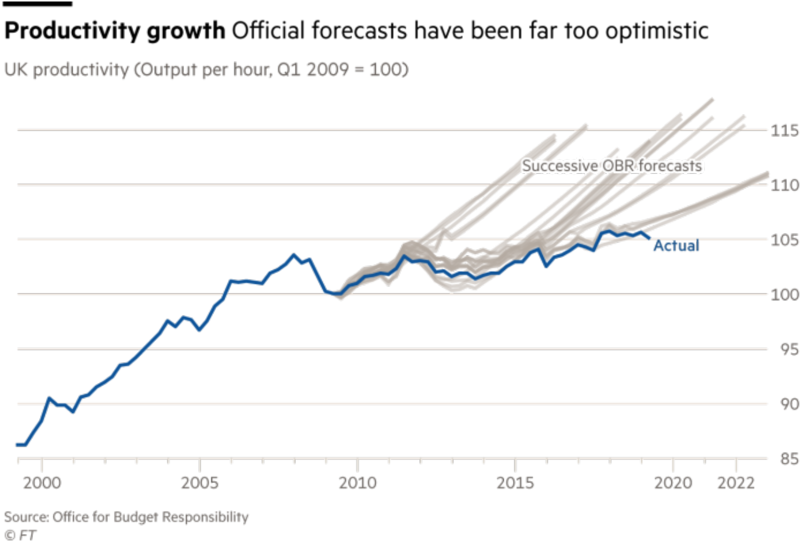
Descriptive Question: Quality of Politicians

	Population			Nominated (non-elected)			Politician (elected)		
	N	Raw	z	N	Raw	z	N	Raw	z
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Leadership motivation	269,748	14.7	-0.012	21,408	17.3	0.331	7,947	18.1	0.433
Dutifulness	270,588	10.9	-0.003	21,457	11.8	0.239	7,987	12.4	0.400
Achievement striving	270,716	13.4	-0.008	21,458	14.6	0.220	7,980	15.2	0.345
Activity-energy	270,921	16.1	-0.006	21,480	16.8	0.131	7,985	17.9	0.333
Sociability	270,879	20.5	-0.007	21,468	22.2	0.213	7,987	22.9	0.300
Self-confidence	270,772	22.5	-0.006	21,470	23.0	0.090	7,987	23.7	0.195
Deliberation	270,845	16.4	0.001	21,465	16.6	0.039	7,985	17.3	0.167

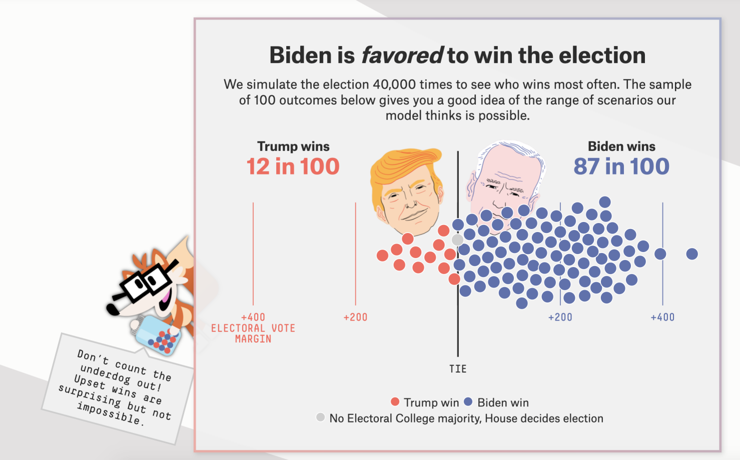
Descriptive Question: Share of Population in Extreme Poverty



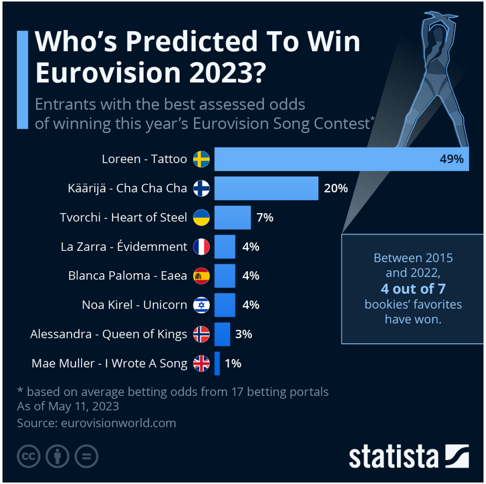
Question of Forecasting: Productivity Growth in the UK



Question of Forecasting: US Presidential Elections



Question of Forecasting: Winner of the Eurovision Song Contest



Back to Causal Effects

- Questions about causal effects are central to policy analyses and for testing economic theories
  - Do generous childcare provisions lead to more women in the labor force, holding all else fixed?
  - Will the demand of a good go down if the sellers raise price, holding all else fixed?

- Why is this hard? Because typically all else is not held constant...!

- Some of the most exciting developments in economic research have focused on estimating causal effects reliably

Like the Nobel Prize in Economics in 2021...

The Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel 2021

Ill. Niklas Elmehed © Nobel Prize Outreach.

David Card

Prize share: 1/2

Ill. Niklas Elmehed © Nobel Prize Outreach.

Joshua D. Angrist

Prize share: 1/4

Ill. Niklas Elmehed © Nobel Prize Outreach.

Guido W. Imbens

Prize share: 1/4

...Or the Nobel Prize in Economics in 2019

The Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel 2019

© Nobel Media. Photo: A. Mahmoud

Abhijit Banerjee

Prize share: 1/3

© Nobel Media. Photo: A. Mahmoud

Esther Duflo

Prize share: 1/3

© Nobel Media. Photo: A. Mahmoud

Michael Kremer

Prize share: 1/3

Do Hospitals Make People Healthier?

- National Health Interview Survey (Angrist and Pischke, p. 13):
  - During the past 12 months, was the respondent a patient in a hospital overnight?
  - Would you say your health in general is excellent (5), very good (4), good (3), fair (2), or poor (1)?

Group	Sample size	Mean Health Status	Std. Error
Hospital	7,774	3.21	0.014
No hospital	90,049	3.93	0.003

- But... is the health of people who go to hospitals otherwise comparable to those who do not?

Potential Outcomes Framework or Rubin's Causal Model

- For each individual, imagine two potential health variables:

$$Potential\ outcome = Y_i = \begin{cases} Y_{1i}\ if\ D_i = 1 \\ Y_{0i}\ if\ D_i = 0 \end{cases}$$

- $D_i = 1$ if hospitalized, 0 if not

- The above can be expressed as:

$$Y_i = Y_{0i} + (Y_{1i} - Y_{0i}) \times D_i$$

- The realized outcome Y_i depends on whether D_i is switched on/off

SUTVA Assumption

- There is an implicit assumption built in to this framework...
- The response of a particular unit depends only on the treatment to which it itself was assigned, not the treatments of others units
- This is called the **Stable Unit Treatment Value Assumption** or **SUTVA**
- It is important to think about and empirically relevant in many cases—we will return to this with examples later in the course

Treatment Effects in a Potential Outcomes Framework

- What do we really want to estimate? Maybe this:

$$\underbrace{E[Y_{1i}] - E[Y_{0i}]}_{\textit{Average Treatment Effect (ATE)}}$$

- Or maybe this:

$$\underbrace{E[Y_{1i}|D_i = 1] - E[Y_{0i}|D_i = 1]}_{\textit{Average Treatment Effect on the Treated (ATT)}}$$

- Note that we do not observe Y_{1i} and Y_{0i} for the same individual
- This is the fundamental missing data problem of causal inference!

Heterogeneous Potential Outcomes: An Example

- Imagine the intervention (D) is going to the gym, and we are studying five individuals whose potential outcomes (Y_{0i}, Y_{1i})—a fitness score measured on a 0-5 scale—are as follows

	Anna	Aron	Saloni	Eric	Gong
Y_{0i}	4	5	4	3	5
Y_{1i}	5	5	4	4	4
D_i	1	1	0	1	0

Heterogeneous Potential Outcomes: An Example

- What are $E(Y_{0i})$ and $E(Y_{1i})$? What are $E(Y_{0i}|D_i = 1)$ and $E(Y_{1i}|D_i = 1)$?
- What are ATE and ATT?

	Anna	Aron	Saloni	Eric	Gong
Y_{0i}	4	5	4	3	5
Y_{1i}	5	5	4	4	4
D_i	1	1	0	1	0

Heterogeneous Potential Outcomes: An Example

- What are $E(Y_{0i})$ and $E(Y_{1i})$? What are $E(Y_{0i}|D_i = 1)$ and $E(Y_{1i}|D_i = 1)$?
- What are ATE and ATT?

	Anna	Aron	Saloni	Eric	Gong
Y_{0i}	4	5	4	3	5
Y_{1i}	5	5	4	4	4
D_i	1	1	0	1	0
$E(Y_{0i})$			4.2		
$E(Y_{1i})$			4.4		

Heterogeneous Potential Outcomes: An Example

- What are $E(Y_{0i})$ and $E(Y_{1i})$? What are $E(Y_{0i}|D_i = 1)$ and $E(Y_{1i}|D_i = 1)$?
- What are ATE and ATT?

	Anna	Aron	Saloni	Eric	Gong
Y_{0i}	4	5	4	3	5
Y_{1i}	5	5	4	4	4
D_i	1	1	0	1	0
$E(Y_{0i})$			4.2		
$E(Y_{1i})$			4.4		
$E(Y_{0i}\|D_i = 1)$			4		
$E(Y_{1i}\|D_i = 1)$			4.67		

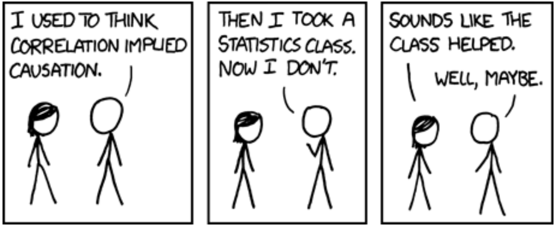
Selection Bias: Naïve Comparison of Averages

- Comparing the average health of hospitalized versus not tells us something, but not necessarily what we want to know

- Formally:

$$\begin{aligned}\underbrace{E[Y_{1i}|D_i = 1] - E[Y_{0i}|D_i = 0]}_{\textit{Observed difference in average health}} &= \underbrace{E[Y_{1i}|D_i = 1] - E[Y_{0i}|D_i = 1]}_{\textit{ATT}} \\ &\quad + \underbrace{E[Y_{0i}|D_i = 1] - E[Y_{0i}|D_i = 0]}_{\textit{Selection bias}}\end{aligned}$$

- Unless the hospitalized/non-hospitalized would have had the same expected outcomes without hospitalization, comparing averages can be misleading!


I USED TO THINK CORRELATION IMPLIED CAUSATION.
THEN I TOOK A STATISTICS CLASS. NOW I DON'T.
SOUNDS LIKE THE CLASS HELPED.
WELL, MAYBE.

What Does Estimating Causal Effects Require?

- To get at causal effects, we need to make sure there is no selection bias
- Effectively means that the treatment status is "ignorable"—we will see many ways of stating this in the course
- A lot of what you will be doing in this course will be to see how you can get there with different methods
  - Under which assumptions
  - Given the data you have
  - Given the nature of the problem
- Much applied research is focused foremost on evaluating the plausibility of causal claims from observational data

Three Steps to Analyzing Causal Effects

1. **Identification**
   - What assumptions are needed to answer the question I have, given the data I have?
   - You need to think about this even if you had the population distribution of the data!
2. **Estimation**
   - What is the estimation method by which you will identify the parameter?
3. **Inference**
   - How can you quantify the uncertainty in the estimate?

- The last two are often combined, the first is distinct (and begins before touching the data)

Random Assignment Solves the Selection Problem

- Random assignment makes treatment independent of potential outcomes
- Then:

$$\underbrace{E[Y_{1i}|D_i = 1] - E[Y_{0i}|D_i = 0]}_{\textit{Observed difference in average health}} = \underbrace{E[Y_{1i}|D_i = 1] - E[Y_{0i}|D_i = 1]}_{ATT}$$

- Further:

$$\underbrace{E[Y_{1i}|D_i = 1] - E[Y_{0i}|D_i = 1]}_{\textit{Observed difference in average health}} = \underbrace{E[Y_{1i} - Y_{0i}|D_i = 1]}_{ATT} = \underbrace{E[Y_{1i} - Y_{0i}]}_{ATE}$$

- The first equality is the key merit of random assignment—dealing with selection bias at source

Bloom et al. (2013): Does Management Matter?

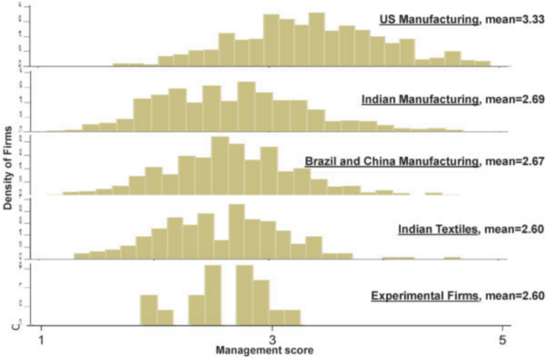


FIGURE I

Management Practice Scores across Countries

Histograms using Bloom and Van Reenen (2007) methodology. Double-blind surveys used to evaluate firms' monitoring, targets, and operations. Scores from 1 (worst practice) to 5 (best practice). Samples are 695 U.S. firms, 620 Indian firms, 1,083 Brazilian and Chinese firms, 232 Indian textile firms, and 17 experimental firms. Data from http://www.worldmanagementsurvey.com.

What Is the Selection Problem Here?

- Suppose I told you that firms which score higher on the management score also have higher productivity...
- Will that be enough to conclude that the differences in management cause higher productivity?
- The problem is the same as earlier: much else may differ between better-/worse-managed firms than just the quality of management
- Yet, the answer to this causal question underlies the vast international management consultancy world!
- Bloom et. al. (2013) randomly assign management consultants to plants in Indian textile firms to understan the consequences of management practices for productivity

DOES MANAGEMENT MATTER? EVIDENCE FROM INDIA*

NICHOLAS BLOOM
BENN EIFERT
APRAJIT MAHAJAN
DAVID MCKENZIE
JOHN ROBERTS

A long-standing question is whether differences in management practices across firms can explain differences in productivity, especially in developing countries where these spreads appear particularly large. To investigate this, we ran a management field experiment on large Indian textile firms. We provided free consulting on management practices to randomly chosen treatment plants and compared their performance to a set of control plants. We find that adopting these management practices raised productivity by 17% in the first year through improved quality and efficiency and reduced inventory, and within three years led to the opening of more production plants. Why had the firms not adopted these profitable practices previously? Our results suggest that informational barriers were the primary factor explaining this lack of adoption. Also, because reallocation across firms appeared to be constrained by limits on managerial time, competition had not forced badly managed firms to exit. *JEL* Codes: L2, M2, O14, O32, O33.

The Intervention

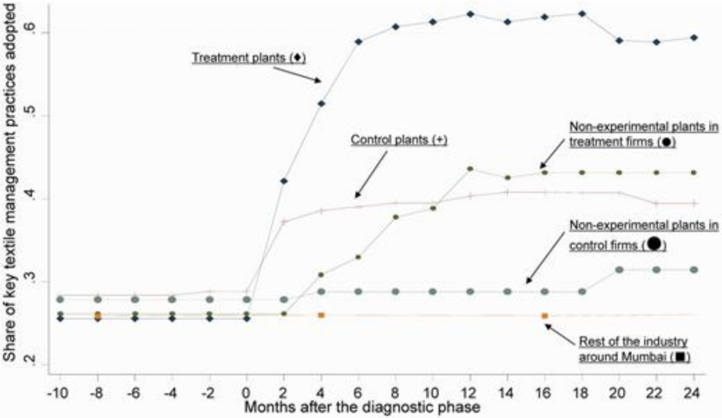
- **Sample** chosen randomly from the population of all textile firms around Mumbai with 100-1000 workers
  - Eventual sample: 17 firms willing to participate (out of 66); 11 treatment, 6 control
  - About 270 employees, annual sales about 13M USD

- **Randomization** at plant-level within treatment firms
  - 14 treatment plants, 6 "pure" control plants
  - Non-experimental plants: 5 at treatment firms, 3 at control firms
  - 96 plants from the rest of the industry

- **Intervention**
  - Diagnostic phase: All treatment and control plants were given a diagnostic assessment and report
  - Implementation phase: Detailed management support to improve practices given to treatment plants
    - Factory operations, quality control, inventory, human resources management, and sales and order management
  - Measurement phase: All treatment and control plants, light management advice

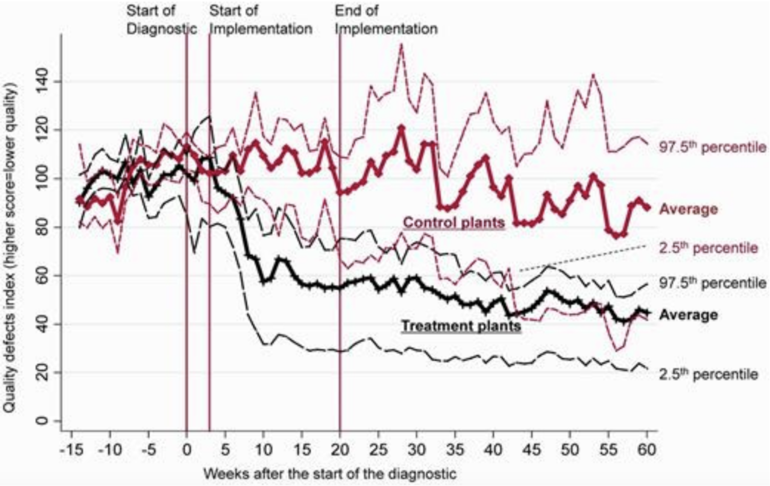
What Is the Impact of Management Consultancy? Management Practices

Figure V



What Is the Impact of Management Consultancy? Output Quality

Figure VI



Treatment and Control Plants Before the Intervention

TABLE I

THE FIELD EXPERIMENT SAMPLE

	All				Treatment	Control	Diff
	Mean	Median	Min	Max	Mean	Mean	*p*-value
Number of plants	28	n/a	n/a	n/a	19	9	n/a
Number of experimental plants	20	n/a	n/a	n/a	14	6	n/a
Number of firms	17	n/a	n/a	n/a	11	6	n/a
Plants per firm	1.65	2	1	4	1.73	1.5	0.393
Employees per firm	273	250	70	500	291	236	0.454
Employees, experimental plants	134	132	60	250	144	114	0.161
Hierarchical levels	4.4	4	3	7	4.4	4.4	0.935
Annual sales ($m) per firm	7.45	6	1.4	15.6	7.06	8.37	0.598
Current assets ($m) per firm	8.50	5.21	1.89	29.33	8.83	7.96	0.837
Daily mtrs, experimental plants	5,560	5,130	2,260	13,000	5,757	5,091	0.602
BVR management score	2.60	2.61	1.89	3.28	2.50	2.75	0.203
Management adoption rates	0.262	0.257	0.079	0.553	0.255	0.288	0.575
Age, experimental plant (years)	19.4	16.5	2	46	20.5	16.8	0.662
Quality defects index	5.24	3.89	0.61	16.4	4.47	7.02	0.395
Inventory (1,000 kilograms)	61.1	72.8	7.4	117.0	61.4	60.2	0.945
Output (picks, million)	23.3	25.4	6.9	32.1	22.1	25.8	0.271
Productivity (in logs)	2.90	2.90	2.12	3.59	2.91	2.86	0.869

What is the Effect of Getting Healthcare Insurance?

- Healthcare policy is a very contentious policy issue in the United States
  - No universal coverage health system
- Medicaid is a major (means-tested) social healthcare program which provides free health insurance to low-income and disabled Americans
- In 2008, Oregon expanded the Medicaid program in the state by 10,000 slots for residents who may otherwise not have received healthcare
  - Anticipating that demand will be significantly higher than the available slots, Oregon decided to allot the slots by lottery
  - This creates a **randomized trial at scale**
- Baicker et al. (2013) study the effects of the Oregon Health Insurance Experiment

Sample Characteristics

Table 1. Characteristics of the 12,229 Survey Respondents.*

Characteristic	Controls (N=5842)	Lottery Winners (N=6387)†	P Value
	percent		
Female sex	56.9	56.4	0.60
Age group‡			
19–34 yr	36.0	35.1	0.38
35–49 yr	36.4	36.6	0.87
50–64 yr	27.6	28.3	0.43
Race or ethnic group§			
Non-Hispanic			
White	68.8	69.2	0.68
Black	10.5	10.6	0.82
Other	14.8	14.8	0.97
Hispanic	17.2	17.0	0.82
Interview conducted in English	88.2	88.5	0.74

* Values for the control group (persons not selected in the lottery) are weighted means, and values for the lottery-winner group are regression-adjusted weighted means. P values are for two-tailed t-tests of the equality of the two means.

† Lottery winners were adults who were randomly selected in the lottery to be able to apply for Medicaid coverage.

‡ The data on age are for the age of the respondent at the time of the in-person interview. The study sample was restricted to persons who were between 19 and 64 years of age during the study period.

§ Race and ethnic group were self-reported. The categories of non-Hispanic race (white, black, and other) were not mutually exclusive; respondents could report as many races or ethnic groups as they wished.

Insurance and Self-Rated Health

Table 3. Mean Values and Absolute Change in Health-Related Quality of Life and Happiness with Medicaid Coverage.*

Variable	Mean Value in Control Group	Change with Medicaid Coverage (95% CI)†	P Value
Health-related quality of life			
Health same or better vs. 1 yr earlier (%)	80.4	7.84 (1.45 to 14.23)	0.02
SF-8 subscale‡			
Mental-component score	44.4±11.4	1.95 (0.03 to 3.88)	0.05
Physical-component score	45.5±10.5	1.20 (−0.54 to 2.93)	0.18
No pain or very mild pain (%)	56.4	1.16 (−6.94 to 9.26)	0.78
Very happy or pretty happy (%)	74.9	1.18 (−5.85 to 8.21)	0.74

* Plus–minus values are weighted means ±SD. Where means are shown without standard deviations, they are weighted means. The effect of Medicaid coverage was estimated with the use of two-stage least-squares instrumental-variable regression. All regressions included indicators for the number of household members on the lottery list, and all standard errors were clustered on household. All analyses were weighted with the use of survey weights. The sample was all 12,229 survey respondents.

† For variables measured as percentages, the change is expressed as percentage points.

‡ Scores on the Medical Outcomes Study 8-Item Short-Form Health Survey (SF-8) range from 0 to 100, with higher subscale scores indicating better self-reported health-related quality of life. The scale is normalized to yield a mean of 50 and a standard deviation of 10 in the general U.S. population.

Insurance and Spending on Healthcare

Table 4. Mean Values and Absolute Change in Financial Hardship with Medicaid Coverage.*

Variable	Mean Value in Control Group	Change with Medicaid Coverage (95% CI)†	P Value
Any out-of-pocket spending (%)	58.8	−15.30 (−23.28 to −7.32)	<0.001
Amount of out-of-pocket spending ($)	552.8±1219.5	−215.35 (−408.75 to −21.95)	0.03
Catastrophic expenditures (%)‡	5.5	−4.48 (−8.26 to −0.69)	0.02
Any medical debt (%)	56.8	−13.28 (−21.59 to −4.96)	0.002
Borrowed money to pay bills or skipped payment (%)	24.4	−14.22 (−21.02 to −7.43)	<0.001

* Plus–minus values are weighted means ±SD. Where means are shown without standard deviations, they are weighted means. The effect of Medicaid coverage was estimated with the use of two-stage least-squares instrumental-variable regression. All regressions include indicators for the number of household members on the lottery list, and all standard errors were clustered on household. All analyses were weighted with the use of survey weights. The sample was all 12,229 survey respondents.

† For variables measured as percentages, the change is expressed as percentage points.

‡ Persons with catastrophic expenditures had out-of-pocket medical expenses that exceeded 30% of their household income.

Insurance and Healthcare Utilization

Table 5. Mean Values and Absolute Change in Health Care Utilization and Spending, Preventive Care, Access to and Quality of Care, and Smoking and Obesity with Medicaid Coverage.*

Variable	Mean Value in Control Group	Change with Medicaid Coverage (95% CI)†	P Value
Utilization (no. of visits or medications)			
Current prescription drugs	1.8±2.8	0.66 (0.21 to 1.11)	0.004
Office visits in past 12 mo	5.5±11.6	2.70 (0.91 to 4.49)	0.003
Outpatient surgery in past 12 mo	0.1±0.4	0.03 (−0.03 to 0.09)	0.28
Emergency department visits in past 12 mo	1.0±2.0	0.09 (−0.23 to 0.42)	0.57
Hospital admissions in past 12 mo	0.2±0.6	0.07 (−0.03 to 0.17)	0.17
Estimate of annual health care spending ($)‡	3,257.3	1,171.63 (199.35 to 2,143.91)	0.018
Preventive care in past 12 mo (%)			
Cholesterol-level screening	27.2	14.57 (7.09 to 22.04)	<0.001
Fecal occult-blood test in persons ≥50 yr	19.1	1.26 (−9.44 to 11.96)	0.82
Colonoscopy in persons ≥50 yr	10.4	4.19 (−4.25 to 12.62)	0.33
Flu shot in persons ≥50 yr	35.5	−5.74 (−19.31 to 7.83)	0.41
Papanicolaou smear in women	44.9	14.44 (2.64 to 26.24)	0.016
Mammography in women ≥50 yr	28.9	29.67 (11.96 to 47.37)	0.001
PSA test in men ≥50 yr	21.4	19.18 (1.14 to 37.21)	0.037

Insurance and Clinical Outcomes

Table 2. **Mean Values and Absolute Change in Clinical Measures and Health Outcomes with Medicaid Coverage.***

Variable	Mean Value in Control Group	Change with Medicaid Coverage (95% CI)†	P Value
Blood pressure			
Systolic (mm Hg)	119.3±16.9	−0.52 (−2.97 to 1.93)	0.68
Diastolic (mm Hg)	76.0±12.1	−0.81 (−2.65 to 1.04)	0.39
Elevated (%)‡	16.3	−1.33 (−7.16 to 4.49)	0.65
Hypertension			
Diagnosis after lottery (%)§¶	5.6	1.76 (−1.89 to 5.40)	0.34
Current use of medication for hypertension (%)§‖	13.9	0.66 (−4.48 to 5.80)	0.80
Cholesterol**			
Total level (mg/dl)	204.1±34.0	2.20 (−3.44 to 7.84)	0.45
High total level (%)	14.1	−2.43 (−7.75 to 2.89)	0.37
HDL level (mg/dl)	47.6±13.1	0.83 (−1.31 to 2.98)	0.45
Low HDL level (%)	28.0	−2.82 (−10.28 to 4.64)	0.46
Hypercholesterolemia			
Diagnosis after lottery (%)§¶	6.1	2.39 (−1.52 to 6.29)	0.23

Internal Validity

- A study is said to have **internal validity** if it has plausibly solved the problem of identification
  - I.e. dealt with the issue of selection bias convincingly
  - Randomization is one such strategy (but we will study many others)
- Informally, can think of this as answering the following:
  - Given the data they have, and the models they estimate, have authors identified a causal estimate **in their population**?
- Internal validity is the most important characteristic on which we judge the **reliability** of a given study of causal effects
  - *Sine qua non* for taking treatment effects seriously
  - The focus of most of what we will do on this course

External Validity

- **External validity** relates to the extent the specific causal effect that a study identifies may be used as a guide to the parameter value in other populations
- For instance, is the impact of management consultancy in Indian textile firms likely to be a good guide to the effects in...
  - ...textile firms in Bangladesh?
  - ...software firms in India?
  - ...furniture makers in China?
- Or will the effects on low-income Oregon residents translate to other parts of the US? What about other countries?
- This is typically a matter of judgment (based on the features of the two populations), and it is relevant for both experimental and observational studies
- But it is of primary importance in judging **relevance** of a study

External Validity of the Oregon Health Insurance Experiment

"Our estimates of the effect of Medicaid coverage on health, health care utilization, and financial strain apply to able-bodied, uninsured adults with incomes below 100% of the federal poverty level who express interest in insurance coverage."

- Would the same results hold if e.g. this was a question of universal coverage?
  - Would e.g. non-poor households have similar outcomes?
- What if the policy proposal was to make universal coverage for the elderly only?
  - Would these results be externally valid for that population?

Good Papers (Usually) Discuss the Generalizability of Their Results

However, there are several important limits to the generalizability of our findings. First, the low-income uninsured population in Oregon differs from the overall population in the United States in some respects, such as the proportions of persons who are members of racial and ethnic minority groups. Second, our estimates speak to the effect of Medicaid coverage on the subgroup of people who signed up for the lottery and for whom winning the lottery affected their coverage status [...] Third, the newly insured participants in our study constituted a small share of all uninsured Oregon residents, limiting the system-level effects that insuring them might generate, such as strains on provider capacity or investment in infrastructure [...]

Recap: What Was This Lecture About?

1. Introduction to the course
   - Objectives
   - Administration and logistics

2. Types of data and empirical questions

3. Introduction to potential outcomes framework
   - ATE and ATT
   - Selection bias

4. Random assignment as a solution
   - Bloom et al. (2013) on the importance of management practices
   - The Oregon Health Insurance Experiment

5. Internal and external validity

What Should/Could You Read?

- To recap today's material, you can read the following:
  - Angrist and Pischke (2009): Chapters 1-2
  - Wooldridge (2015): Chapter 1
- In preparation for the coming lectures:
  - Cunningham (2021): Sections 2.7-2.10

Papers Discussed in This Lecture

- Baicker, K., Taubman, S. L., Allen, H. L., Bernstein, M., Gruber, J. H., Newhouse, J. P., ..., and Finkelstein, A. N. (2013). "The Oregon Experiment – Effects of Medicaid on Clinical Outcomes." *New England Journal of Medicine*, 368(18), 1713-1722.

- Bloom, N., Eifert, B., Mahajan, A., McKenzie, D., and Roberts, J. (2013). "Does Management Matter? Evidence from India." *Quarterly Journal of Economics*, 128(1), 1-51.