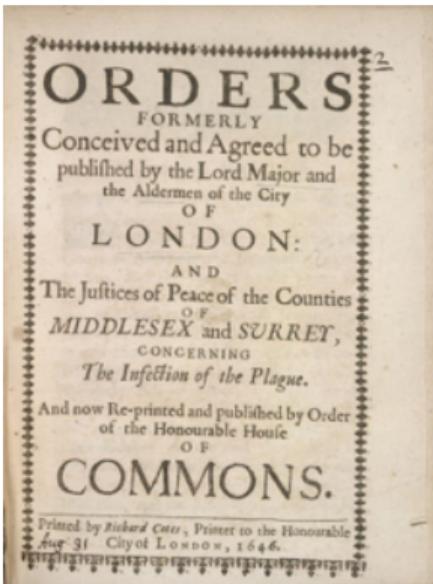


Introduction to Course

Why Causal Inference Matters for Public Policy

Two brief lessons in causation from London

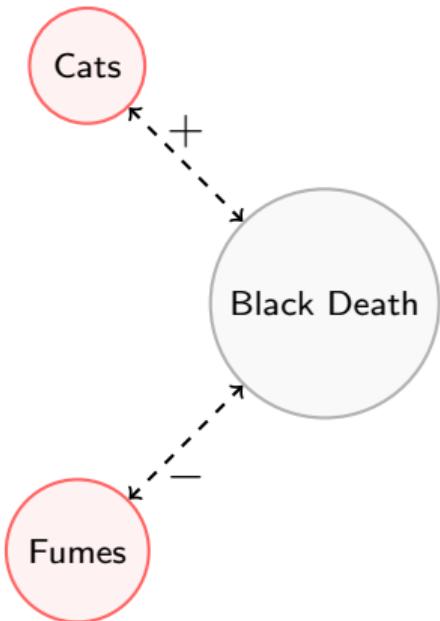


"Orders" enabled Lord Mayor Sir John Lawrence to take several drastic actions

- ▶ London erupted with the Bubonic Plague in 1665
 - ▶ Cause unknown
- ▶ Perhaps contact with animals led to culling of animals
 - ▶ "They talked of [slaughtering] forty thousand dogs, and five times as many cats!" (Defoe, 1722)

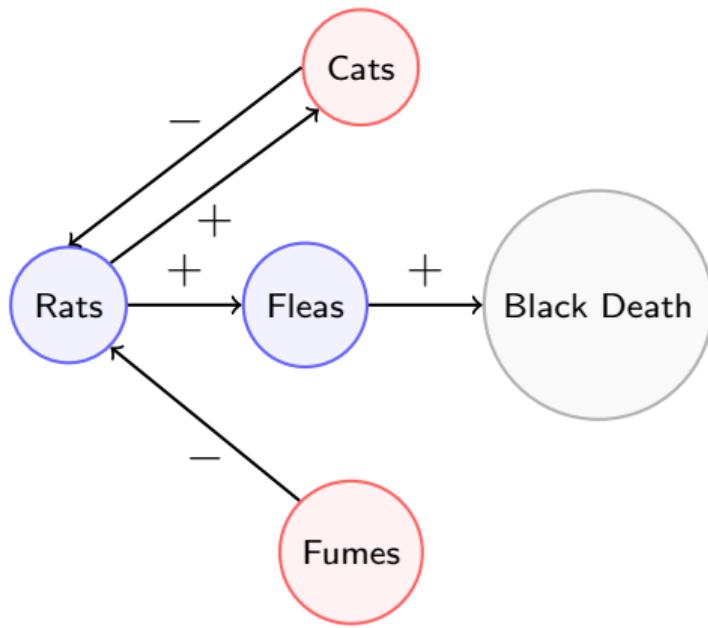


The 1665 bubonic plague



- ▶ Cats positively **correlated** with plague
- ▶ Fumes negatively **correlated** with plague

The 1665 bubonic plague



- ▶ Cats positively correlated with plague, but cause a decrease in deaths!
- ▶ Correlation due to relationship with rats
- ▶ Killing cats increased deaths

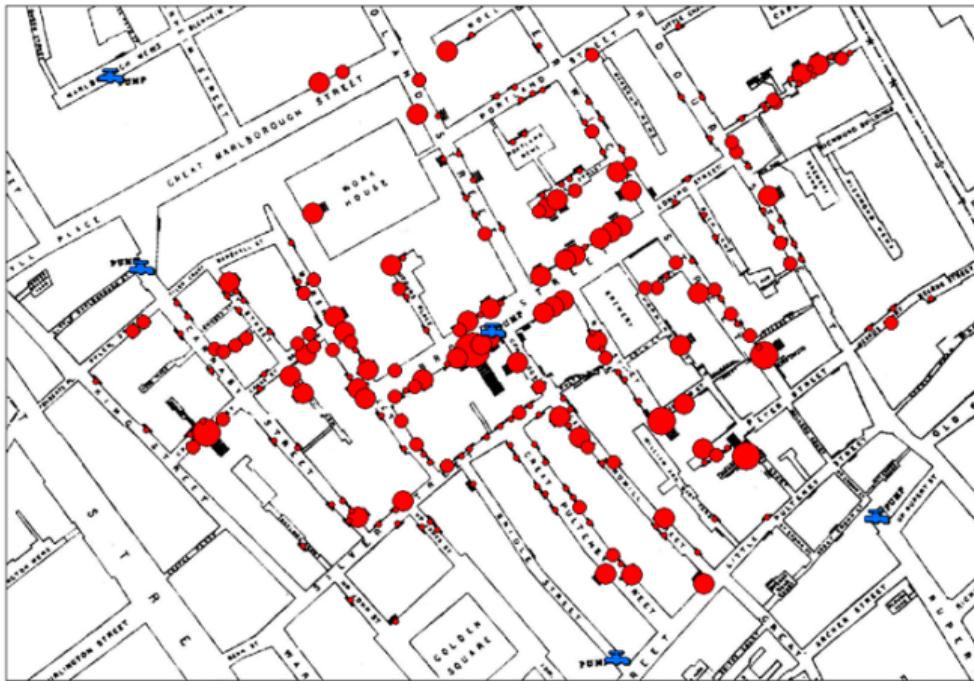
Course goal: prevent you from killing harmless cats

Some progress: the 1854 cholera outbreak



- ▶ Mid-19th century: source of cholera still unknown – miasma popular theory
- ▶ John Snow mapped deaths in 1854 outbreak

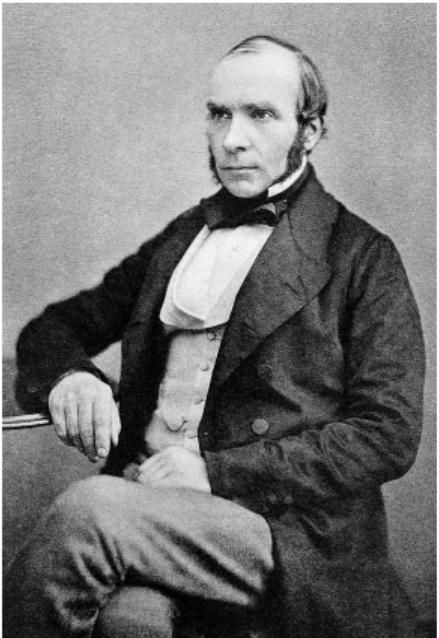
Some progress: the 1854 cholera outbreak



Some progress: the 1854 cholera outbreak



- ▶ Handle was removed to prevent further spread of Cholera



≠



Figure 1: Dr. John Snow, founder of modern epidemiology

Figure 2: Not to be confused with Jon Snow of Game of Thrones



Figure 3: Prof. Wright



Figure 4: Prof. Wright with Prof. Callen

More on the causes of cholera

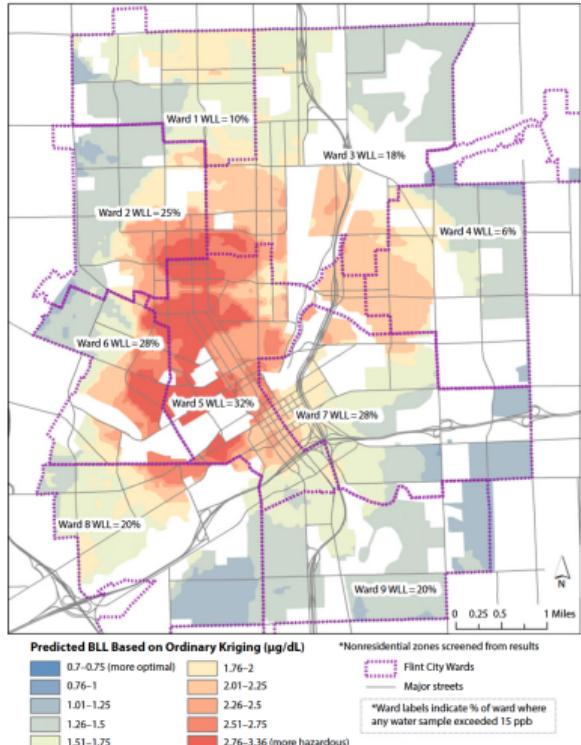
- ▶ John Snow reported death rates by water source in South London
 - ▶ In 1849, two companies drew water from the contaminated Thames in central London
 - ▶ Southwark & Vauxhall Co
 - ▶ Lambeth Co
 - ▶ Starting in 1852, the Lambeth changed the source of its water to the river at Thames Ditton, an uncontaminated water source upstream
- Natural Experiment!

More on the causes of cholera

I.—SYNOPSIS OF RESULTS.			
Death-Rates <i>per 1,000</i> of living Population in Two Epidemic Periods.	In Houses enumerated in 1854 as receiving their Water-supply—		
	from the LAMBETH Company.	from the SOUTHWARK and VAUXHALL Company.	
CHOLERA	1848-9..	12.5	11.8
	1853-4..	3.7	13.0

Source: Simon (1856). Perhaps, the very first difference in differences!

Flash forward to Flint Drinking Water Crisis, 2015



- ▶ We still use these same tools today!
- ▶ Flint, Michigan changed its water supply from Detroit-supplied Lake Huron to local Flint River
- ▶ Switching to Flint River water increased lead levels in drinking water
- ▶ Dr. Hanna-Attisha showed increases in water lead levels caused increases in blood lead levels in children

Parallels to COVID

The screenshot shows the USA Today homepage with a navigation bar at the top. Below it, a news article titled "Ohio judge orders hospital to treat COVID patient with Ivermectin despite CDC warnings" by Jake Zuckerman and Terry DeMio from the Ohio Capital Journal and The Enquirer. The article was published on Aug. 30, 2021, at 11:18 a.m. ET. Social sharing icons for Facebook, Twitter, and Email are visible below the headline.

NATION

Ohio judge orders hospital to treat COVID patient with Ivermectin despite CDC warnings

Jake Zuckerman and Terry DeMio Ohio Capital Journal and The Enquirer
Published 11:18 a.m. ET Aug. 30, 2021



There is 'no good evidence' Ivermectin helps treat COVID, says Oklahoma Center

Asai — Kouistas — Wright

UChicago Harris PPHA311



- ▶ Parallels to today's public health crisis
- ▶ Important to figure out what works and what doesn't. This is in part the role of the FDA/CDC.

Course Practicalities

Goals of the course

You will learn:

- ▶ That correlation is not causation
 - ▶ and what to do about it
- ▶ How to read a regression properly
- ▶ How to perform basic econometric analyses
- ▶ How to use R
- ▶ How to critically read empirical work

Course expectations

- ▶ I expect this course to be hard
 - ▶ You will have to exert effort to learn
 - ▶ I will try very hard to make assignments purposeful
 - ▶ You will have to learn by doing
- ▶ I will make mistakes
 - ▶ Typos will happen
 - ▶ We will make corrections
- ▶ Notation may vary (in course and in real life)
 - ▶ Important to understand ideas in more than one notation

Class logistics 1/3

► Assignments and Grading:

- The final grade for the course will be a function of the midterm (30%), final (40%) and five homework assignments (30%)
- You may work on the problems with others in the class, but you must turn in your own set of answers and indicate on the first page who you worked with
- No late problem sets will be accepted. The lowest problem set grade will be dropped at the end of the quarter.

► Gradescope Submission of Problem Sets:

- We will be using Gradescope via Canvas to manage assignments and grading this quarter (See syllabus for the details)
- Data needed for any assignments will be available on “Canvas” → “Modules”

Class logistics 2/3

- ▶ Textbooks:
 - ▶ *Mastering 'Metrics* by Angrist and Pischke
 - ▶ *Introductory Econometrics: A Modern Approach* (7th Ed.) by Wooldridge
- ▶ Supplemental textbook:
 - ▶ *Introduction to Econometrics* (4th Ed.) by Stock and Watson
- ▶ Multiple TA sessions (pick one), TA office hours, my office hours
 - ▶ May be helpful for problem sets.
- ▶ Problem sets:
 - ▶ You can work in groups, but must turn in your own assignment and indicate on the first page who you worked with.
 - ▶ You may **not** use any material from prior versions of this course

Class logistics 3/3

- ▶ Syllabus is a **guideline** for dates
 - ▶ Readings are also available on syllabus
- ▶ We will post slides on Canvas ahead of time
- ▶ We will cover a lot of material, quickly
 - ▶ Most of you will have to prepare for lecture to keep up
 - ▶ “Bonus” slides will not be on exams

Post to Ed Discussion with your questions

- ▶ For many questions, someone else may have that question too!
- ▶ Please post your question to the Ed Discussion messaging board, via Canvas (Positive Externalities!)

R

- ▶ You should be familiar with R from Stats I
- ▶ Problem sets will require you to use R
- ▶ R is a life skill – I want you to learn to teach yourself software
- ▶ Goal: if you are given some data later in life, you will be comfortable opening up R (or something similar) and **figuring out** how to open, manipulate and analyze it

TA Sessions and Office Hours

1 Weekly Recitation Sessions

- ▶ You are expected to attend the TA section you are registered for.
- ▶ Sessions will emphasize practice problems related to the week's lectures.

2 Optional Coding TA Session/Office Hours

- ▶ These sessions will go over some example coding exercises, and provide an opportunity for students to ask coding questions related to the course.
- ▶ No advance registration is necessary. See canvas.

3 Instructor Office Hours

- ▶ Drop-in and by appointment. See Canvas.

Conceptual course outline

- 1 Establish formal model of causal inference
- 2 Learn tools of bivariate and multivariate regression
- 3 Note **problems**: when will inference go awry?
- 4 Learn **solutions**: when can cleverness make problems go away?
- 5 Learn additional **tools**

This course is pessimistic: it is hard to uncover causal relationships. Goal is to leave you with enlightened pessimism.

Ethical Academic Conduct

We adhere to the official Harris School protocol for ethical violations:

- ▶ First violation:
 - ▶ The student will be summoned to meet with the Dean of Students and the instructor
 - ▶ The student will receive a grade of 0 on the assignment/exam in question
 - ▶ A formal letter of finding is sent to the student
- ▶ Second violation:
 - ▶ The case will be sent to the Harris Area Disciplinary Committee
 - ▶ The Area Disciplinary Committee can assign sanctions including suspension or expulsion from the University

Ethical Academic Conduct and Homework Assignments

- ▶ You may work on the homework assignments with others in the class
- ▶ However, you must turn in your own set of answers and code and indicate on the first page who you worked with
- ▶ (Demand-side) Copying the homework of another student/passing code from student to student in cheating
- ▶ (Supply-side) Providing another student with your assignment or sharing code to copy is considered cheating

Course Topics

- 1 Week of Jan 1: a) Intro, b) Causal inference/ randomized trials
- 2 Week of Jan 8: Bivariate regression
- 3 Week of Jan 15: Multivariate regression
- 4 Week of Jan 22: a) More multivariate regression, b) Binary dependent variables
- 5 Week of Jan 29: a) Midterm review
- * MIDTERM Feb 1st Thursday
- 6 Week of Feb 5: a) Functional form, b) Problems in practice
- 7 Week of Feb 12: Difference-in-differences
- 8 Week of Feb 19: Instrumental Variables
- 9 Week of Feb 26: a) Regression Discontinuity, b) Final review
- * FINAL EXAM March 6 Wednesday 9am-11am