Natural Experiments

Overview and Applications

Today's Agenda

Natural Experiments

- **Overview**: What natural experiments are and why we should care.
- **Application #1**: Elections.
- **Application #2**: Income redistribution.
- **Application #3**: Political attitudes.

Overview

Overview

The nature of the treatment

- What's so "***natural***" about **natural experiments**?
  - It's all about the ***nature* of the treatment**:
    - Is ***not*** "**designed/implemented**" by the researcher (**?**)
      Examples.
    - Is "**unknown**" ***and*** "**unknowable**" to the researcher (**?**)
      Examples.
    - Is "**probabilistic**" and it depends on an "**external factor**" (**?**)
      Examples.
- If you can't design one, and are unsure about the treatment, and the treatment doesn't depend on you, **how do *you* study causal effects using one?**

Overview

The nature of the treatment

Titiunik emphasizes: we should ***not*** focus on the **assumption** of "**as-if-randomness**."

➡Scholars usually state that their "**natural experiments**" are (1) natural and (2) experimental **based on "as-if-randomness" claims**.
Mimicking RCTs.

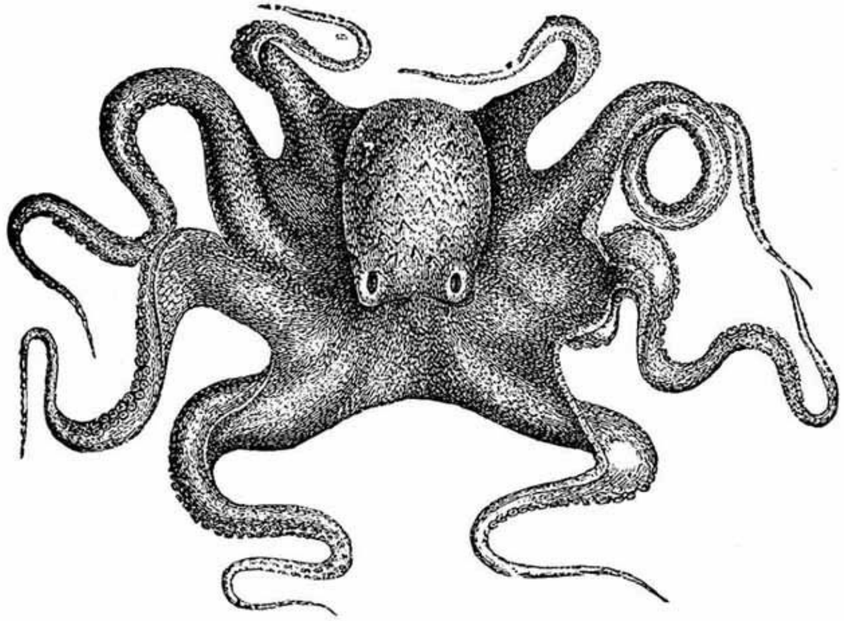
- Problem**?**
  - Natural experiments ***can't*** be defined as "**imperfect**" **RCTs**.
    RCTs are truly randomized designs.

✓As discussed, the **nature of the treatment is *radically* different**.
For ex., the distribution of the treatment assignment mechanism is unknown and unknowable.

Overview

Randomness and unconfoundedness

- “**Randomness does *not* imply unconfoundedness**” (**?**)
  - What does **randomization accomplish?**
  - What does **unconfoundedness guarantee?**



Can an octopus pick the papers I stacked on my desk “*at random*”?

Overview

Randomization devices and unconfoundedness

- The ***only*** **way to ensure unconfoundedness** is via a "**randomization device/ procedure**" (**?**)
  - A **set of rules** that allows the researcher to **assign the treatment** according to a ***known*** "**probability function**" (**?**)
  - What does a **known probability function ensure?**
  - Notice that the **researcher** must be able to **identify** the **probability distribution**, why**?**
    - **Probabilistic does *not* ensure unconfoundedness**.
      It could be for example a probabilistic distribution but one that always/never picks certain cases—e.g., octopus example.

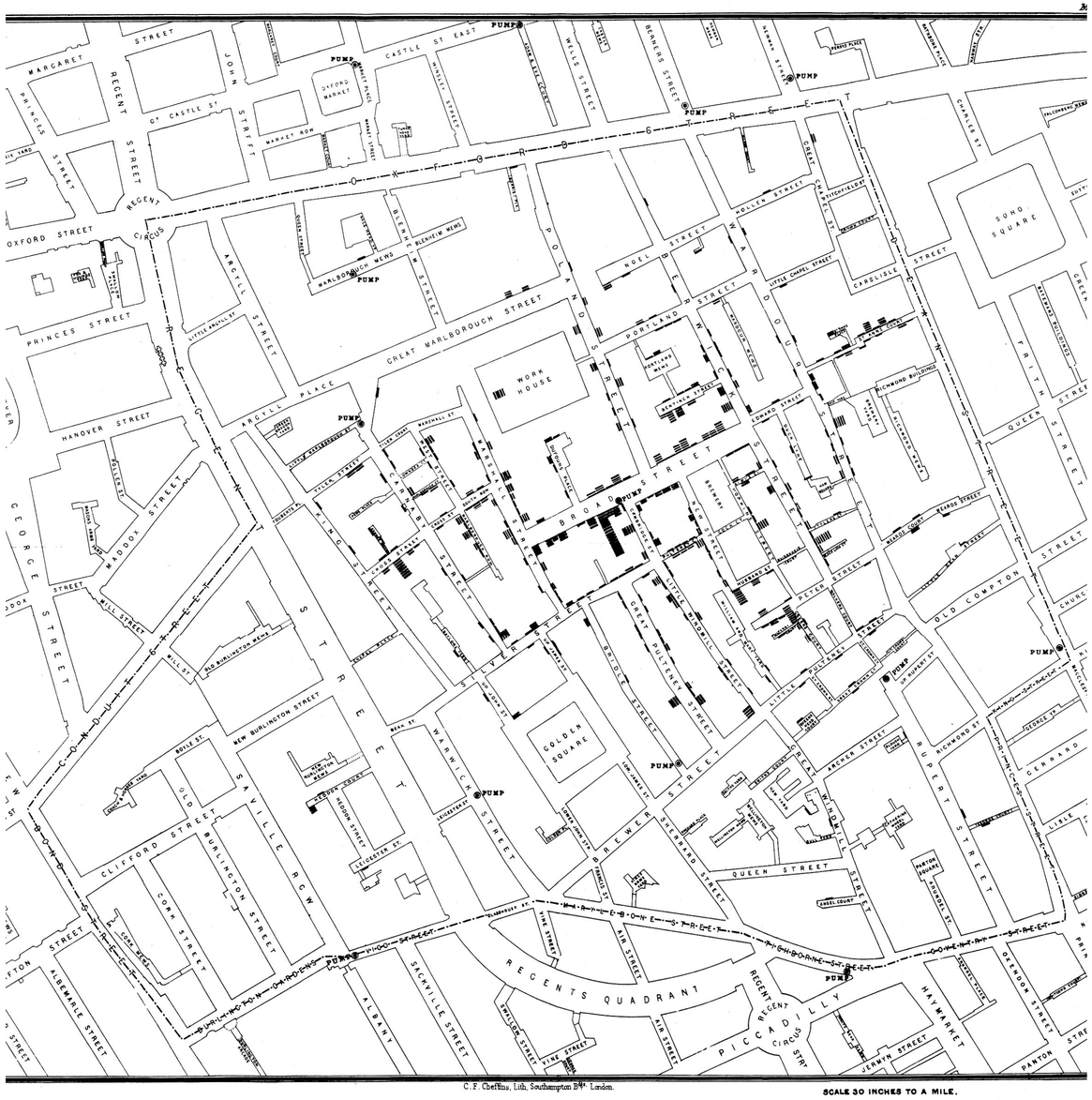
Overview

Defining a natural experiment

Randomization mechanism (recap):

- ✓Is not designed/implemented by the researcher.
- ✓Is unknown and unknowable.
- ✓Is probabilistic by virtue of an external “device.”

◉ **What’s this map?**



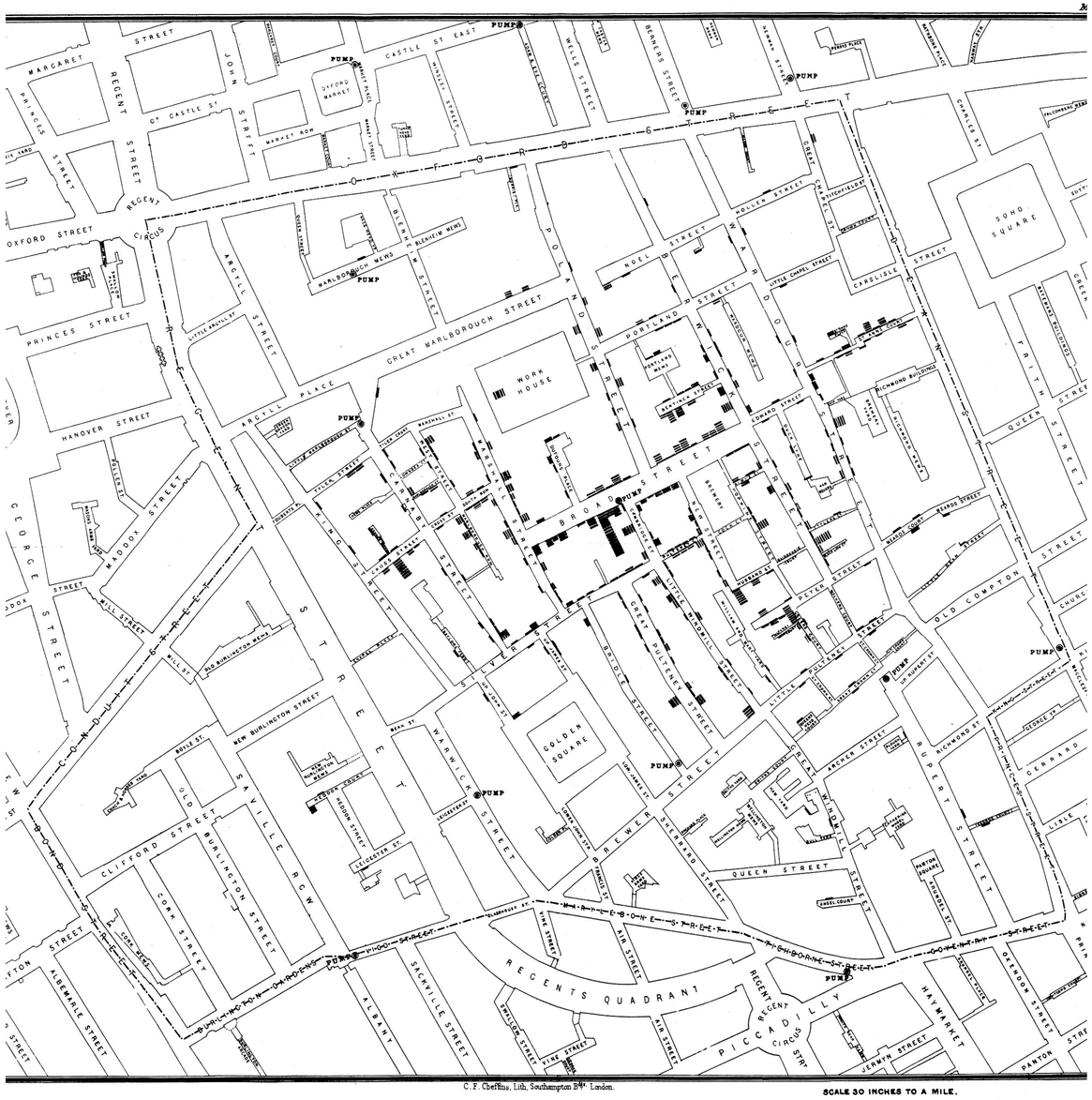
Overview

Defining a natural experiment

Randomization mechanism (recap):

- ✓Is not designed/implemented by the researcher.
- ✓Is unknown and unknowable.
- ✓Is probabilistic by virtue of an external “device.”

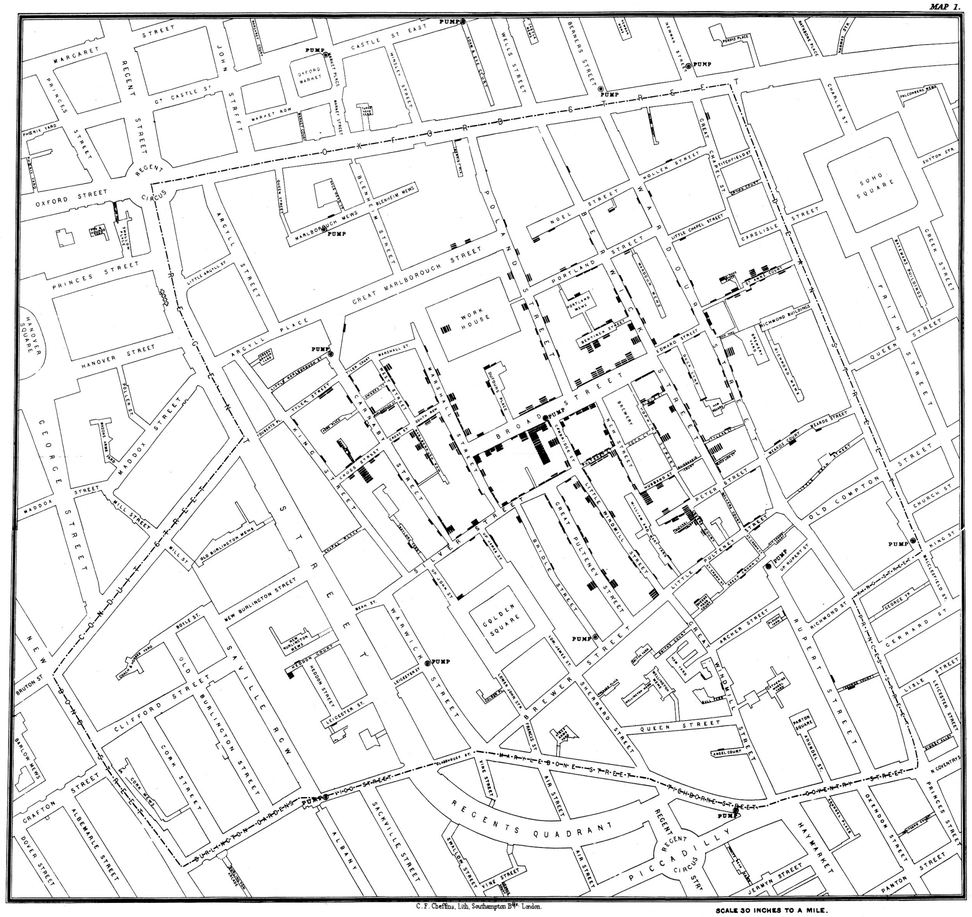
◉ **What’s this map?**



John Snow's map showing the clustering of cholera cases in Soho during the London epidemic of 1854

What *cause*s cholera?

A study of water supply and miasma in London, 1854

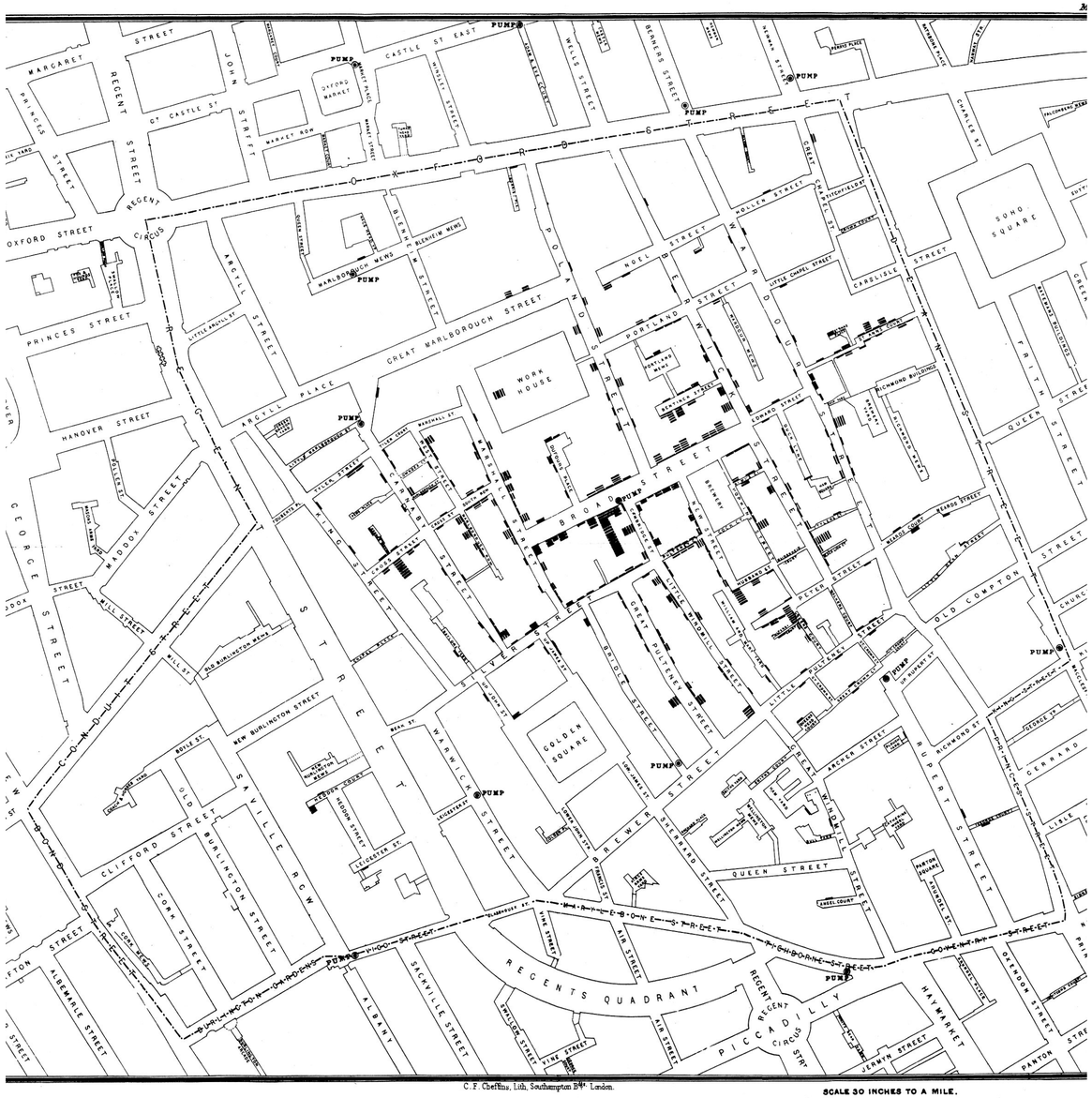


Overview

Was Snow's design a natural experiment?

Randomization mechanism:

- **Was it designed/implemented by the researcher?**
- **Was it unknown and unknowable?**
- **Was it probabilistic by virtue of an external "device"?**



John Snow's map showing the clustering of cholera cases in Soho during the London epidemic of 1854

Rocío Titiunik
Professor, Dpt. of Politics, Princeton University

"In a natural experiment […] the researcher finds some intervention that has been implemented and also finds some subjects. She then constructs treatment and control groups to address a particular hypothesis. But the treatment and control groups are constructed *post hoc*."

Rocío Titiunik
Professor, Dpt. of Politics, Princeton University

Natural Experiments are a subclass of observational study

Overview

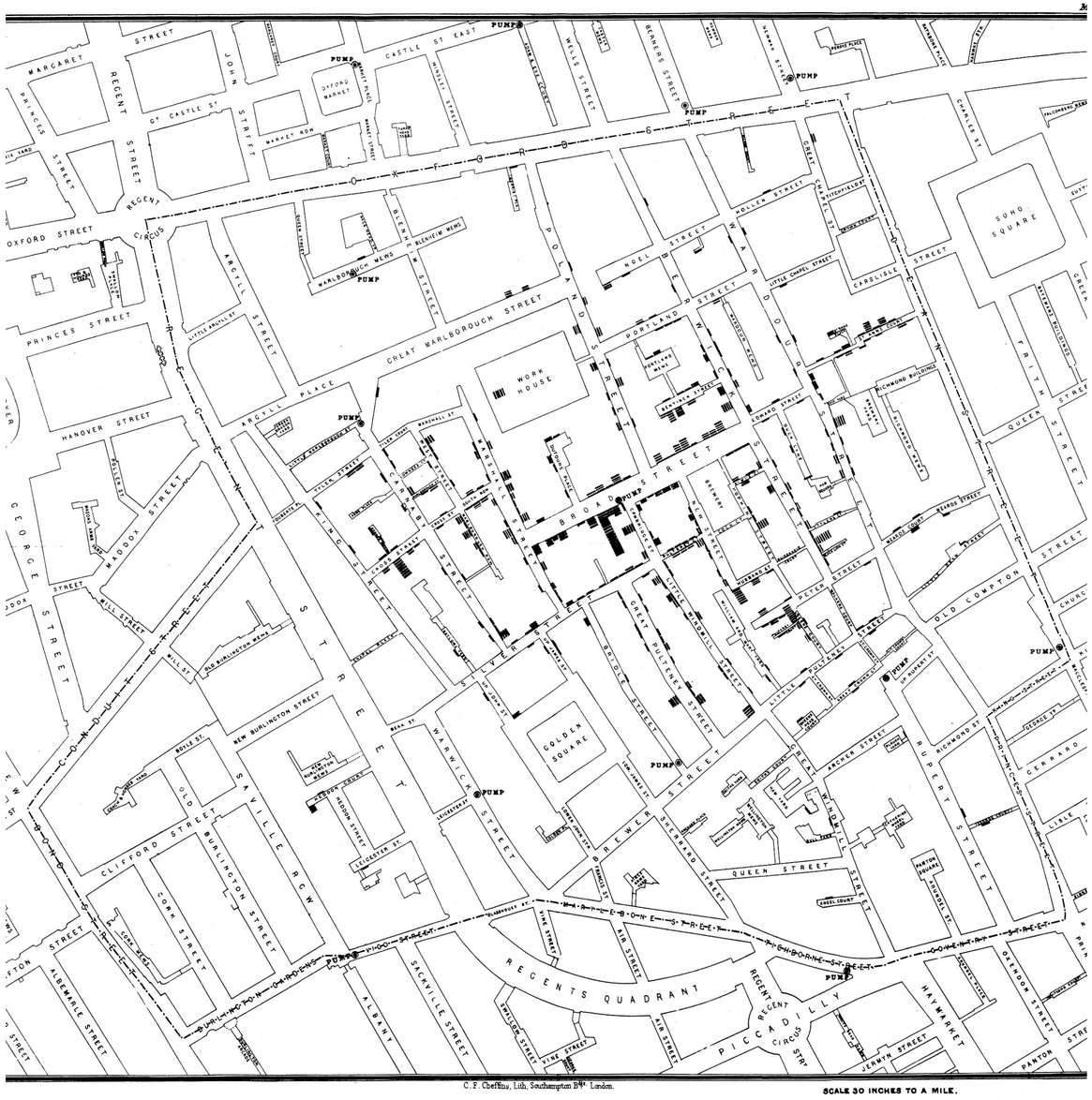
Natural experiments: A subclass of observational study

If certain **assumptions** **hold**, we can claim we have a natural experiment.

But never based on the usually claimed "**as-if random**" assumption.

Randomization mechanism:

- Is not designed/implemented by the researcher.
- Is unknown and unknowable.
  Researcher can find out.
- Is probabilistic by virtue of an external "device."
  Researcher can find out.



John Snow's map showing the clustering of cholera cases in Soho during the London epidemic of 1854

Overview

Extended techniques: Achieving Covariate Balance

- Covariate balance is important because it ensures that the **treatment (T) and control (C) groups are ("statistical") identical** (**?**)
  Except that T is treated and C is not.

- As discussed, however, **just *claiming*** that a natural experiment is "natural and experimental" is *not* enough—**certain assumptions must hold**.

- There are extended **techniques that might help** the researcher in making the case his/her design is indeed a **natural experiment**. "**Matching**."

Why care?

Scholars use matching to support their claims that this “unknown” and “unknowable” treatment distribution indeed achieved **unconfoundedness**, which is **reflected in covariate balance** (the ultimate goal of randomization).

Elections

Application #1

Elections

Overview

- **Gist**: What do the authors find**?**
  - Working-class **appearance** negatively affects the number of **votes** received in Finnish municipal elections.
    This relationship affects *only* female candidates.
- Is this a natural experiment**?**
  Is there any randomization process, and if so, what's being randomized?
- Is this design "better" than a non-experimental (i.e. observational) design**?**
  Think of a survey.

Elections

Design: empirical advantages

What's to like about this piece? Possible random (*natural*) assignment of "physical appearances" to political candidates, indicating that this is indeed a **natural experiment**.

- What do the following features achieve**?**
  - Extensive usage of **public posters** as the ***only*** means of campaigning.
  - The design exploits a **real situation**, with **real outcomes** and **actual behavior**.

Elections

Design: what could go wrong?

- **Is this piece perfect, though? What could go wrong?**
- You can see the authors **are really trying** to make the case that:
- There was no "**omitted variable bias**" (**?**)
  - **Assumption** that there is no "**private information**" (i.e. "gossip effects").
    Voters get "treated" *only* with posters (and no other alternatives sources of campaigning).
    - What could happen if there is **private information?**
    - Do the authors **solve** this potential **source of omitted variable bias?**
- There was no "**self-selection bias**" (**?**)
  - Only attractive individuals could **self-select** into political campaigning.
    What do the authors say?

Income Redistribution

Application #2

Income Redistribution

Overview

- **Gist**: What do the authors find**?**
  - **Lottery money**:
    1. **Increases hostility** towards **taxation**.
    2. **Increases hostility** towards **redistribution**.
    3. **Does not affect** attitudes concerning **socio-economic stratification**.
- Is this a natural experiment**?**
  What's being randomized?
- Is this design "**better**" than a non-experimental (i.e. observational) design**?**
  Think of a survey.

Income Redistribution

Design: empirical advantages

What's to like about this piece? **Income** is **widely used** in social sciences. However, **it is never random**. Thus, income is usually **correlated** with something else. In this design, however, we can ask a super interesting question, ***Does randomly assigned income causally change attitudes towards redistribution?***

What do the following features achieve**?**

- Lotteries *are* **random**.
- Authors exploit amount of the prize.
  Unequal and random alterations in income.

Income Redistribution

Design: what could go wrong?

Is this piece perfect, though? What could go wrong?

- You can clearly see the authors struggled with the following:
- "**Demand bias**": the authors had to **interview** prize winners and **ask them directly** about their **attitudes towards redistribution**.
  What could go wrong here?
- "**Sample attrition**": ***some*** prize winners **did not want to be interviewed**.
  Is that a problem?
- **Different *marginal utilities* of prize winners**: Does the same dollar awarded satisfies you the same, regardless if your poor or rich?
  How can you control for that?

Political Attitudes

Application #3

Political Attitudes

Overview

- **Gist**: What do the authors find**?**
  - The "**vulnerability**" (**?**) of serving in the **Vietnam war causally changed political attitudes**:
    - Became more **antiwar**.
    - Became more **liberal/Democrat** (i.e., left-wing).
- Is this a natural experiment**?**
  What's being randomized?
- Is this design **better** than a non-experimental (i.e. observational) design**?**
  Think of a survey.

Political Attitudes

Design: empirical advantages

What's to like about this piece? Super smart empirical strategy of **quantifying** the "**risk of being drafted**": it introduces a **continuous treatment** (draft lottery), and it connects it with an **important question** (political attitudes).

- Lotteries *are* **random**, assigning 366 **lottery numbers** according to the **citizen's birthday**.
  Low lottery number (drafted), high lottery number (not drafted).

- Authors exploit the "risk" of being drafted ("vulnerability").
  Random alterations in vulnerability.

Political Attitudes

Design: what could go wrong?

Is this piece perfect, though? What could go wrong?

- You can clearly see the authors struggled with the following:
- **Between-subjects design**: Sure—it was an unpopular war. But does that ***necessarily*** mean that ***all*** draftees did ***not*** **want** to go to Vietnam?
  Let's think of an alternative design: within-subjects.
- Actually, the "***vulnerability***" **trick** seemed like a lot of **hand-waving** to me.
  Thoughts?

Thank you