Difference in Differences

Vinish Shrestha

11/8/2021

What causes cholera?

- Medical field oblivious as to how cholera is transmitted in 1800s
- Majority – microscopic particles floating throught the air
  - quaranting and isolation did not help
- John Snow, physician, believed that cholera transmitted through air
  - through empirical evidence, changed his mind
  - proved that its transmitted through WATER !!

Snow's Theory

1. living organism
2. interact with passed water
3. mixed with drinking water
4. contaminate water in Thames river
5. people drink contaminated water
6. cycle repeats

How to get to the truth and provide evidence?

- Two major water companies in London in 1800s
  - Southwark and Vauxhall (SV)
  - Lambeth
- Both of these water companies received water from a similar location of Thames
  - mixed with sewage and other waste
- In 1852, Lambeth changed their source
  - upwards, which avoided sewage of London
- Districts that receive water from Lambeth also receive water from SV

Two approaches

1. Compare closeby districts in 1854's epidemic – SV-only vs. Lambeth and SV.

- Lambeth and SV districts were treated in 1854
  - Lambeth changed its source -However, effects might be masked
  - Lambeth and SV districts are getting uncontaminated (Lambeth) and contaminated water (SV)
- How do we identify the effects of uncontaminated water?

$SVLambeth_{1854} - SVLambeth_{1849}$

- Shows improvements in SV and Lambeth districts
- *But could be that 1854's epidemic was less intense compared to 1849*

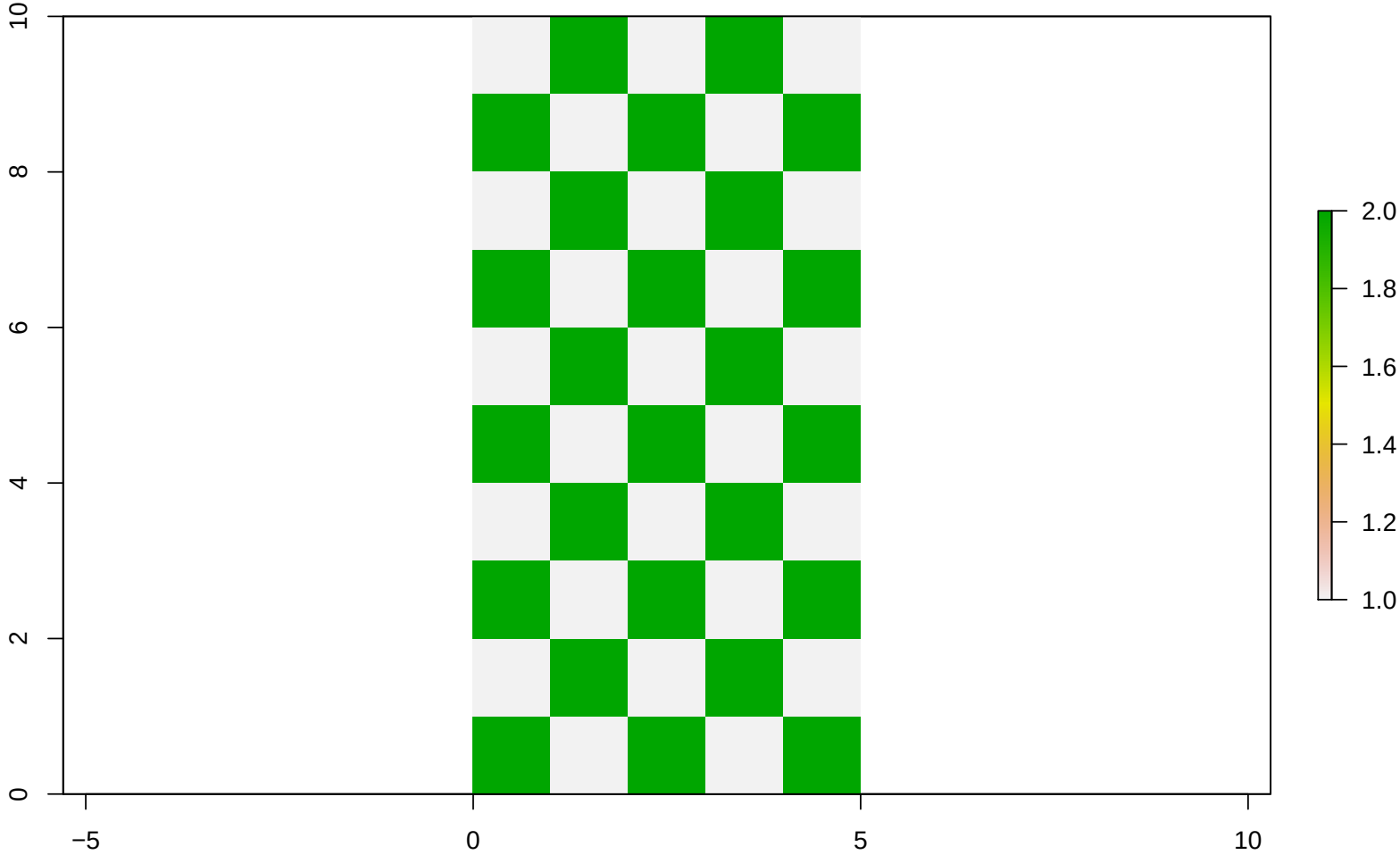
How to compare intensity of epidemic?

- See what is happening to SV-only districts
  - water quality unchanged (contaminated)

$SV_{1854} - SV_{1849}$

- **Any changes in cholera related mortality that is driven by factors other than water will be captured**
  - cholera intensity (1849 vs 1854)
  - advancement in medicine
  - awareness
- These are factors common to both SV-only and SV-Lambeth districts

A Simple Visual Example



- Green SV-Lambeth, white SV-only

Note two levels of variations

1. Time (before and after)
2. Group
   - treatment or affected (SV-Lambeth)
   - control or unaffected (SV-only)

Identify the treatment effects

$$SVLambeth_{1854} - SVLambeth_{1849} - [SV_{1854} - SV_{1849}]$$

Results: outcome variable – count of cholera related mortaility

source	cholera1849	cholera1854	change
SV only	2261	2458	197
SV and Lambeth	3905	2547	-1358
Lambeth only	162	37	-125

- $-1358 - [197] = -1555$
- raw; need to adjust this by population

Results: population adjusted per 100,000

source	cholera1849	cholera1854	change
SV only	135	85	-50
SV and Lambeth	147	19	-128

Case 2. More micro level evidence

2. Focus on districts that receive water from both SV and Lambeth. Compare households that receive water from Lambeth vs. SV