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*

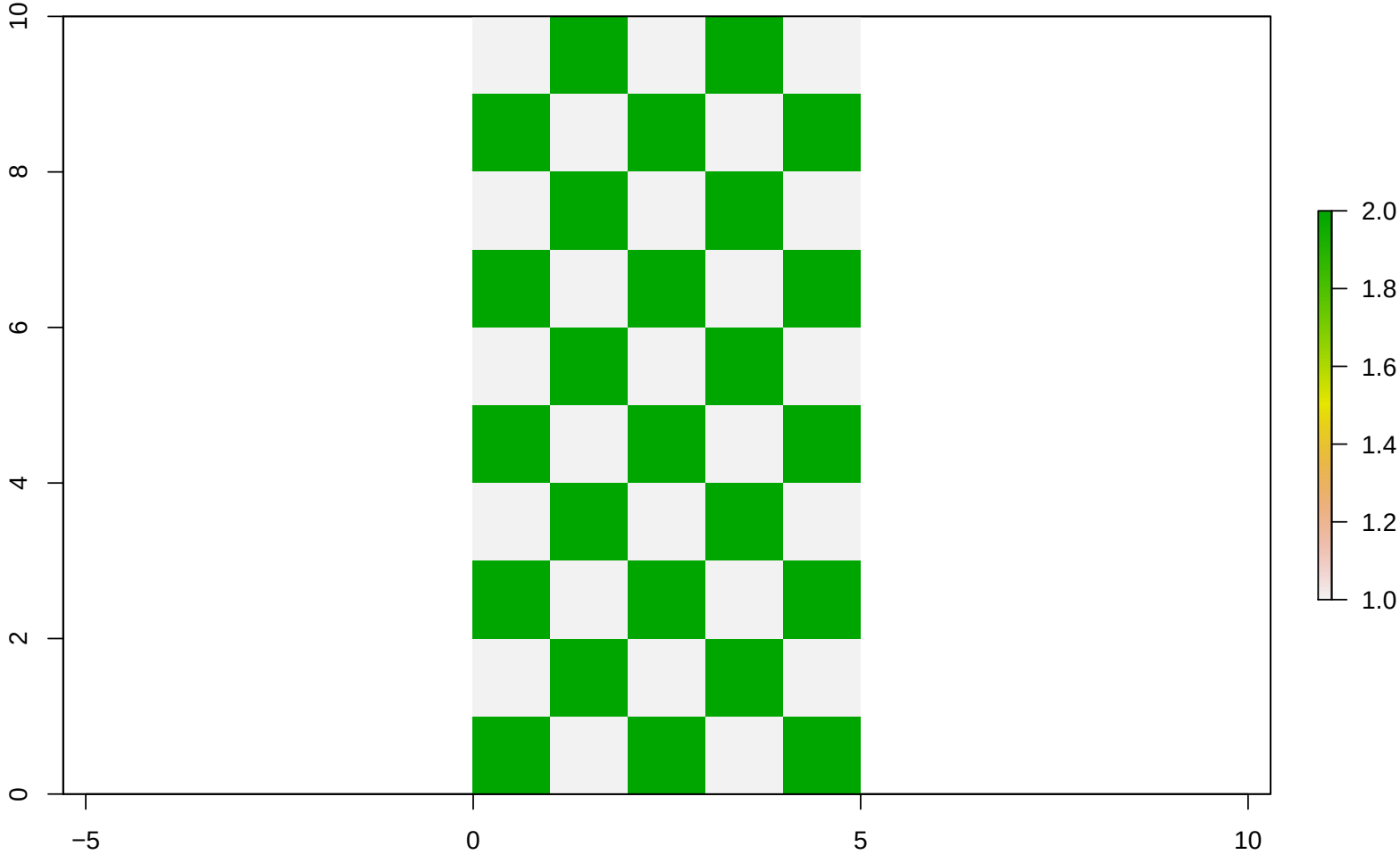
How to compare intensity of epidemic?

- 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 sub-districts that receive water from both SV and Lambeth. Compare households that receive water from Lambeth vs. SV

- Sub-District Comparison

Mortality from Cholera in the seven weeks ending 26th August.

Sub-Districts.	Population in 1851.	Deaths from Cholera in the seven weeks ending 26th August.	Water Supply.				
			Southwark & Vauxhall.	Lambeth.	Pump-wells.	River Thames and ditches.	Unascertained.
*St. Saviour, Southwark	19,709	125	115	—	—	10	—
*St. Olave, Southwark	8,015	53	43	—	—	5	5
*St. John, Horsleydown	11,360	51	48	—	—	3	—
*St. James, Bermondsey	18,899	123	102	—	—	21	—
*St. Mary Magdalen .	13,934	87	83	—	—	4	—
*Leather Market .	15,295	81	81	—	—	—	—
*Rotherhithe . .	17,805	103	68	—	—	35	—
*Battersea . . .	10,560	54	42	—	4	8	—
Wandsworth . .	9,611	11	1	—	2	8	—
Putney . . .	5,280	1	—	—	1	—	—
*Camberwell . .	17,742	96	96	—	—	—	—
*Peckham . . .	19,444	59	59	—	—	—	—
Christchurch, Southwk.	16,022	25	11	13	—	—	1
Kent Road . .	18,126	57	52	5	—	—	—
Borough Road . .	15,862	71	61	7	—	—	3
London Road . .	17,836	29	21	8	—	—	—
Trinity, Newington .	20,922	58	52	6	—	—	—
St. Peter, Walworth .	29,861	90	84	4	—	—	2
St. Mary, Newington .	14,033	21	19	1	1	—	—
Waterloo Road (1st) .	14,088	10	9	1	—	—	—
Waterloo Road (2nd)	18,348	36	25	8	1	2	—
Lambeth Church (1st)	18,409	18	6	9	—	1	2
Lambeth Church (2nd)	26,748	53	34	13	1	—	5
Kennington (1st) .	24,261	71	63	5	3	—	—
Kennington (2nd) .	18,848	38	34	3	1	—	—
Brixton . . .	14,610	9	5	2	—	—	2
*Clapham . . .	16,290	24	19	—	5	—	—
St. George, Camberwell	15,849	42	30	9	2	—	1
Norwood . . .	3,977	8	—	2	1	5	—
Streatham . . .	9,023	6	—	1	5	—	—
Dulwich . . .	1,632	—	—	—	—	—	—
Sydenham . . .	4,501	4	—	1	2	—	1
	486,936	1514	1263	98	29	102	22

Water Source and Cholera Deaths – Aggregate Summary

Company	Number of Houses	Deaths	per 10,000 houses
SV Company	40046	1263	315
Lambeth Company	26107	98	37
Rest of London	256423	1422	59

*Deaths from cholera and per 10,000 houses

A 2×2 design

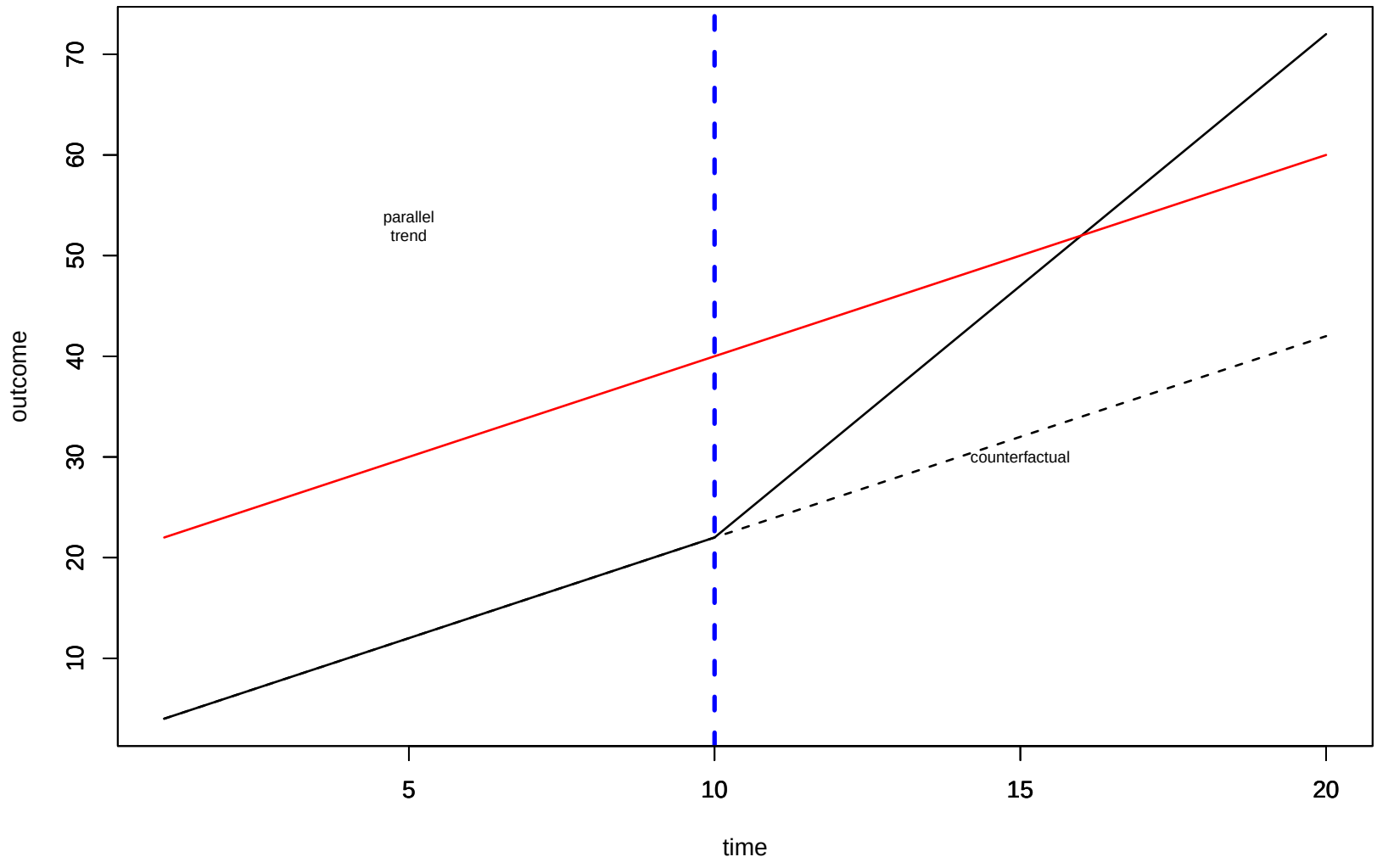
Time	Treated or Exposed	Control or Unexposed
before	$\hat{Y}_{L1849}$	$\hat{Y}_{SV1849}$
after	$\hat{Y}_{L1953}$	$\hat{Y}_{SV1953}$

- First Difference: $\hat{Y}_{L1853} - \hat{Y}_{L1849}$ – within exposed unit across time
- Second Difference: $\hat{Y}_{SV1853} - \hat{Y}_{SV1849}$ – within unexposed unit across time

- $\hat{Y}_{L1853} - \hat{Y}_{L1849} - [\hat{Y}_{SV1853} - \hat{Y}_{SV1849}]$

- **Assumption: outcome in exposed unit will have been similar to the outcome in unexposed unit if not for the treatment (change in water source).**
  - any other unobserved factors leading to changes in cholera deaths in SV-Lambeth districts would also affect cholera deaths in SV-only districts - magnitude would be the same

Parallel Trend Assumption



Treatment Effects