# Quasi-experiments in epidemiology: Prep

Lee Kennedy-Shaffer, PhD

2025-06-10

# Standard Difference-in-Differences

# Motivating Example: Cholera, London, 1850s

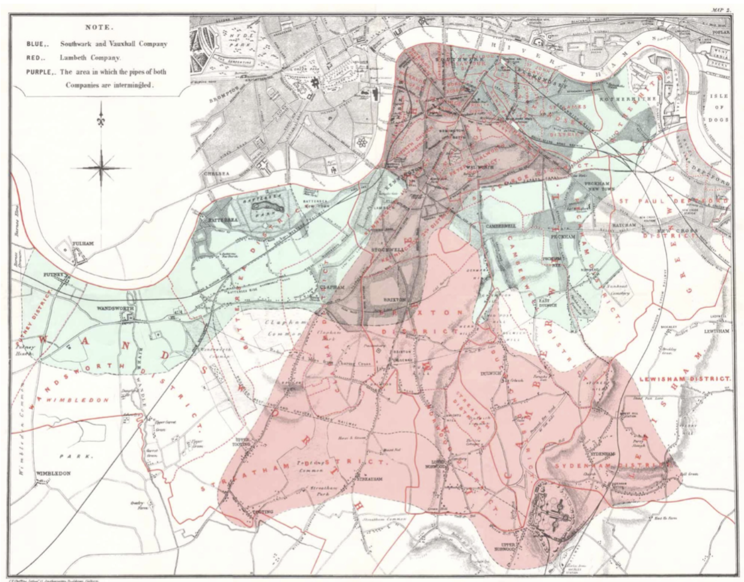



Figure 2: Regions of South London Served by the Southwark & Vauxhall and the Lambeth Companies (Snow [1855])

Causality in the Time of Cholera: John Snow as a Prototype for Causal Inference

Working Paper Version 1.2

Thomas S. Coleman*

March 13, 2019

**Difference-in-Difference in the Time of Cholera: a Gentle Introduction for Epidemiologists**

**Ellen C Caniglia, ScD[1], Eleanor J Murray, ScD, MPH, MS[2]**

# Setting

- Two (or more) units: some treated/exposed, some untreated
- Two time periods: one prior to first treatment, one after

**Example: South London "Grand Experiment" from Coleman 2024**

Untreated: Southwark & Vauxhall Districts (12)

Treated: Joint Southwark & Vauxhall/Lambeth Districts (16)

Time Periods: 1849 (pre-treatment) and 1854 (post-treatment) outbreaks

# Potential Outcomes and Treatment Effect

| Unit | Pre-Treatment | Post-Treatment |
|---|---|---|
| **Exposed** | $Y_{10} = Y_{10}(0)$ | $Y_{11} = Y_{11}(1)$ |
| **Unexposed** | $Y_{00} = Y_{00}(0)$ | $Y_{01} = Y_{01}(0)$ |

Treatment Effect:

$$\theta = E[Y_{11}(1) - Y_{11}(0)]$$

# Change Over Time

Within each unit, we have an interrupted time series:

$$
\begin{aligned}
\Delta_1 &= Y_{11} - Y_{10} \\
\Delta_0 &= Y_{01} - Y_{00}
\end{aligned}
$$

**Key Idea**

Use the observed $\Delta_0$ under control as the potential outcome for the unobserved $\Delta_1$ under treatment.

# Two-by-Two DID

$$\begin{aligned}\hat{Y}_{11}(1) &= Y_{11} \\ \hat{Y}_{11}(0) &= Y_{10} + (Y_{01} - Y_{00}) \\ \hat{\theta} &= (Y_{11} - Y_{10}) - (Y_{01} - Y_{00})\end{aligned}$$

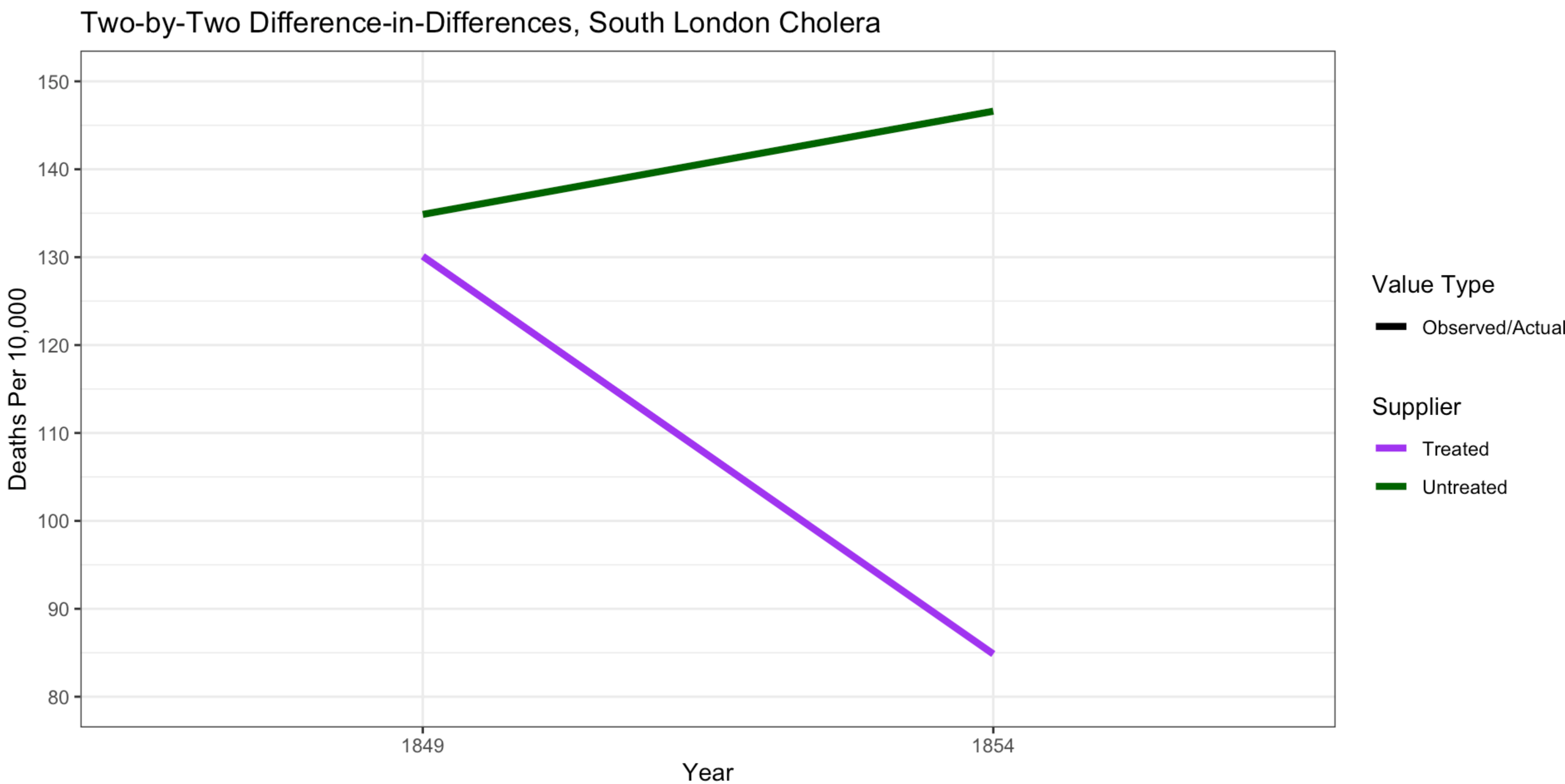
# Two-by-Two DID: Example

| Supplier | Sub-Districts | 1849 Deaths per 10,000 | 1854 Deaths per 10,000 |
|---|---:|---:|---:|
| Joint Southwark & Vauxhall/Lambeth (Treated) | 16 | 130.1 | 84.9 |
| Southwark & Vauxhall Only (Untreated) | 12 | 134.9 | 146.6 |

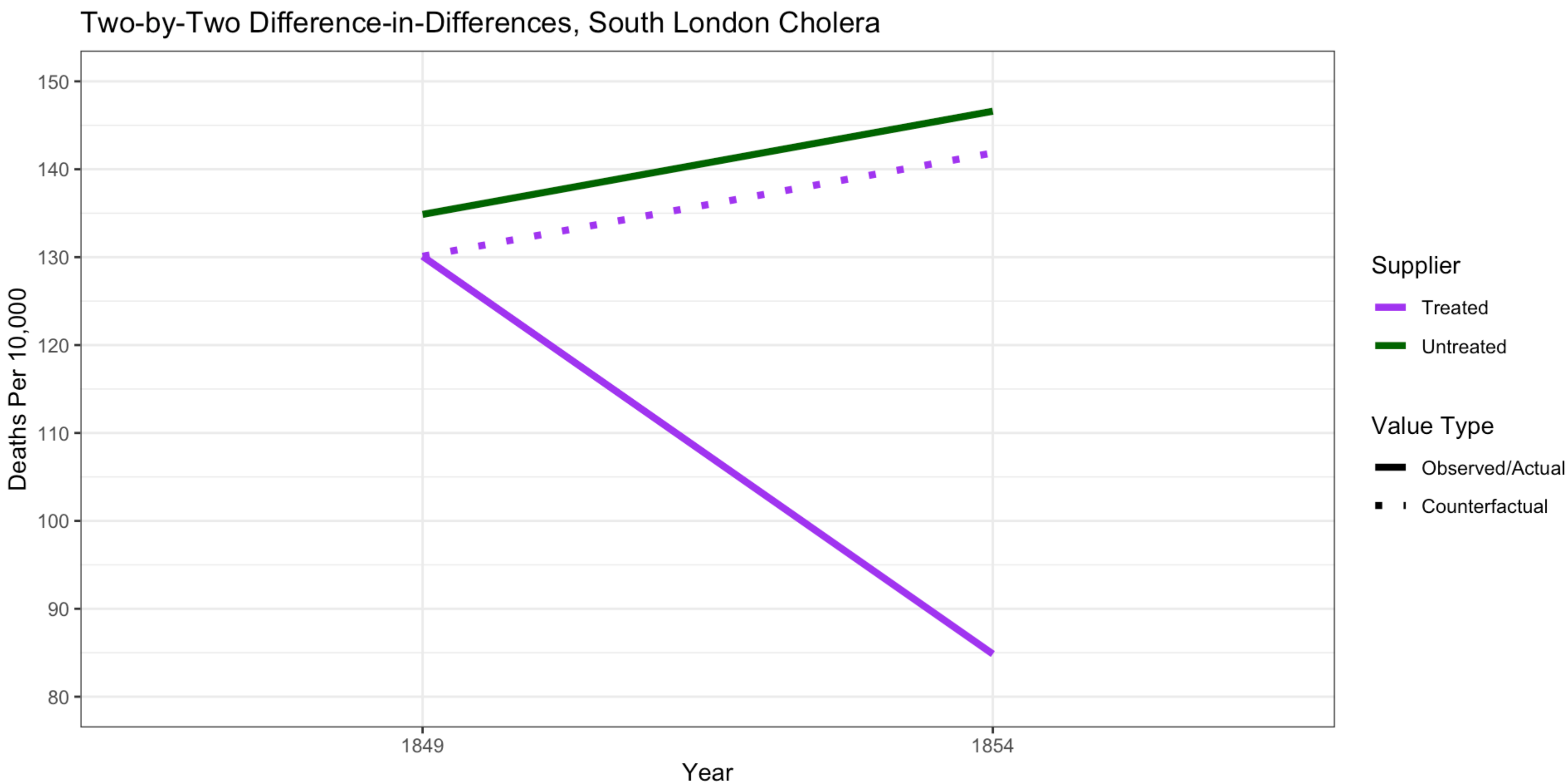
# Two-by-Two DID: Example

| Supplier | 1849 Deaths per 10,000 | 1854 Deaths per 10,000 | Diff, 1854-1849 |
|---|---:|---:|---:|
| Joint Southwark & Vauxhall/Lambeth (Treated) | 130.1 | 84.9 | -45.2 |
| Southwark & Vauxhall Only (Untreated) | 134.9 | 146.6 | 11.8 |
| Diff, Treated-Untreated | -4.8 | -61.8 | -57.0 |

# Two-by-Two DID: Graphically



# Two-by-Two DID: Graphically



# Details and Assumptions

# Regression Form: Two-Way Fixed Effects (TWFE)

$$Y_{it} = \alpha_i + \gamma_t + \theta I(X_{it} = 1) + \epsilon_{it},$$

where:

- $\alpha_i$ is the fixed effect for unit $i$,
- $\gamma_t$ is the fixed effect for time $t$,
- $\epsilon_{it}$ is the error term for unit $i$ in time $t$, and
- $X_{it}$ is the indicator of whether unit $i$ is treated at time $t$.
- $\theta$ is the treatment effect estimand.

# Statistical Inference

Inference can be conducted using the TWFE regression model. This accounts for variability in the outcome if there are multiple treated/untreated units and multiple periods.

Generally, the standard errors are *clustered* by unit to account for correlation. This can also be done with a *block-bootstrap* variance estimation.

**Caution**

This accounts for *statistical uncertainty* but not *causal uncertainty* in the model assumptions. Those cannot be fully assessed statistically.

# Example Analysis Code

See the `analysis/zika-did-handout` file for an example analysis, with visualization and regression-based estimation.

# Key Assumptions

- Parallel trends (in expectation of potential outcomes)
- No spillover
- No anticipation/clear time point for treatment

# Parallel Trends

$$E[Y_{11}(0) - Y_{10}(0)] = E[Y_{01}(0) - Y_{00}(0)]$$

In the absence of treatment, the treated and untreated units would have the same expected outcome trend over time.

# No Spillover

There is no effect of the treatment on any untreated units (similar to a consistency or SUTVA assumption across units).

# No Anticipation

There is no effect of the treatment (or its announcement) prior to the time period assigned as its start (similar to a consistency or SUTVA assumption across periods). A washout period can be incorporated if necessary.

# Assessing Parallel Trends

Placebo/specification tests:

- In-time: conduct the same DID analysis on a time period prior to the actual treatment initiation
- In-space: conduct the same DID analysis as if an untreated unit were the treated one
- Alternative outcome: conduct the same DID analysis on an outcome that should not be affected by the treatment

# Assessing Parallel Trends

These approaches can be used either:

- as a heuristic justification for the assumption,
- to obtain a null distribution for permutation tests, or
- to adjust the estimate for the “null” effect (difference-in-difference-in-differences or triple-differences).

# Approaches to Handle Assumption Violations

# Re-scale the Outcome

Changing the scale of the outcome changes the parallel trends assumption. The most common transformation is to use the **natural log.**

E.g., $\log(Y_{it}) = \alpha_i + \gamma_t + \theta I(X_{it} = 1) + \epsilon_{it}$

Changes parallel trends assumption to:

$$E[\log Y_{11}(0) - \log Y_{10}(0)] = E[\log Y_{01}(0) - \log Y_{00}(0)]$$
$$E\left[\log\left(\frac{Y_{11}(0)}{Y_{10}(0)}\right)\right] = E\left[\log\left(\frac{Y_{01}(0)}{Y_{00}(0)}\right)\right]$$

# Re-scale the Outcome

Caution

- Only one scale can actually have parallel trends
- This changes the estimand (e.g., additive -> multiplicative)

  See Kahn-Lang and Lang (2020) for more considerations and Feng and Bilinski (2024) for examples of different scales/specifications.

# Incorporate Covariates

Incorporating covariates makes the parallel trends assumption *conditional* on those covariates.

E.g., $Y_{it} = \alpha_i + \gamma_t + \theta I(X_{it} = 1) + \beta Z_i + \epsilon_{it}$

Changes parallel trends assumption to:

$$E[Y_{11}(0) - Y_{10}(0) \mid Z_1] = E[Y_{01}(0) - Y_{00}(0) \mid Z_0]$$

# Incorporate Covariates

Caution

- This makes the parallel trends assumption more complex to consider and requires modeling covariates
- This changes the estimand and assumes the effect is homogeneous across covariates

  See Caetano and Callaway (2023) for issues that arise with time-varying covariates.