Differences in Differences

Part II

Fernando Hoces la Guardia
08/04/2022

Housekeeping

- Midterm 2 grades by Friday at the latest.
- PS4 is cancelled. PS1-PS3 will represent 20% of the grade.
- Let's select the chapter for the summary due tomorrow (5pm, gradescope, 300 word limit)

DD and Regression 2/2

- Regression equation (show how $+\delta_{DD}$ is the DD):

$$Y_{dt} = \alpha + \beta TREAT_d + \gamma POST_t + \delta_{DD}(TREAT_d \times POST_t) + e_{dt}$$

DD and Regression 2/2

- Regression equation (show how $+\delta_{DD}$ is the DD):

$$Y_{dt} = \alpha + \beta TREAT_d + \gamma POST_t + \delta_{DD}(TREAT_d \times POST_t) + e_{dt}$$

- Regression estimates:

$$\begin{aligned} Y_{dt} = \mathbf{167} - &\underset{(8.8)}{\mathbf{29}} TREAT_d - \underset{(7.6)}{\mathbf{49}} POST_t + \underset{(10.7)}{\mathbf{20.5}} (TREAT_d \times POST_t) + e_{dt} \end{aligned}$$

- Standard errors of a OLS regression will be to small (overestimate precision) as they assume independent observations.

- Within a unit (district) observations will not be independent, making it less information that with 12 fully independent observations.

DD Estimates Using Real Outputs

- Beyond number of banks what matters most is a measure of economic activity
- Here there is more limited data (back to the world of 4 points) so we inspect the results without regression.
- DD estimate on number of wholesale firms: 181
- DD estimate on net wholesale sales ($ millions): 81

TABLE 5.1
Wholesale firm failures and sales in 1929 and 1933

	1929	1933	Difference (1933–1929)
Panel A. Number of wholesale firms			
Sixth Federal Reserve District (Atlanta)	783	641	−142
Eighth Federal Reserve District (St. Louis)	930	607	−323
Difference (Sixth–Eighth)	−147	34	181
Panel B. Net wholesale sales ($ million)			
Sixth District Federal Reserve (Atlanta)	141	60	−81
Eighth District Federal Reserve (St. Louis)	245	83	−162
Difference (Sixth–Eighth)	−104	−23	81

Notes: This table presents a DD analysis of Federal Reserve liquidity effects on the number of wholesale firms and the dollar value of their sales, paralleling the DD analysis of liquidity effects on bank activity in Figure 5.1.

Back to Minimum Legal Drinking Age (MLDA)

- Wide range of state rules regarding MLDA over time:
  - 1933: After Prohibition Era ended, most states set MLDA at 21.
    - Some exceptions: Kansas, New York, North Carolina.
  - 1971: most states lower MLDA to 18.
    - Some exceptions: Arkansas, California, Pennsylvania.
  - 1984-88: All states transition back to 21. But at different times.
- So much variation at the state level! (makes sense that the DD method was formally developed in the US)

Regression for MLDA using two states

- To illustrate: let's start with a setup equivalent to the Mississippi Study.
- Two states:
    - Alabama (treatment): lower MLDA to 19 in 1975.
    - Arkansas (control): MLDA at 21 since 1933.
- Outcome (Y_{st}): death rates per state (s) for 18-20-year-olds from 1970 to 1983 (t).

$$Y_{st} = \alpha + \beta TREAT_s + \gamma POST_t + \delta_{DD}(TREAT_s \times POST_t) + e_{st}$$

Regression for MLDA using two states

- To illustrate: let's start with a setup equivalent to the Mississippi Study.
- Two states:
    - Alabama (treatment): lower MLDA to 19 in 1975.
    - Arkansas (control): MLDA at 21 since 1933.
- Outcome (Y_{st}): death rates per state (s) for 18-20-year-olds from 1970 to 1983 (t).

$$Y_{st} = \alpha + \beta TREAT_s + \gamma POST_t + \delta_{DD}(TREAT_s \times POST_t) + e_{st}$$

- Where $TREAT_s$ is a binary variable that takes the value 1 for Alabana and 0 for Arkansas. And $POST_t$ is a binary variable that takes the value 1 from the year 1975 onwards and 0 otherwise.

Regression Using All States 1/3

- But why stop there? There are other "experiments" in other states (e.g. Tennessee's MLDA drop to 18 in 1971, then up to 19 in 1979)

- Two state regression requires some changes:

  - There are many post treatment periods, so instead of $POST_t$, we control for each year by including a binary per year $YEAR_{jt}$ (leaving out one year as the category of reference).
    - E.g., $YEAR_{1972,t}$ is a binary variable that takes the value of 1 when the observation, indexed by t, is in the year 1972 and 0 otherwise.
    - This variables that capture the effects that are fixed within a year, are called year fixed effects.

Regression Using All States 2/3

- More changes to the two state regression:
  - Before the variable $TREAT_s$ effectively was controlling for the differences between the two states in the regression.
  - Now there are many states, and each vary in treatment type, but we still want to control for the effect of each state. What should we do?

Regression Using All States 2/3

- More changes to the two state regression:
  - Before the variable $TREAT_s$ effectively was controlling for the differences between the two states in the regression.
  - Now there are many states, and each vary in treatment type, but we still want to control for the effect of each state. What should we do?
  - Instead of $TREAT_s$ we control for each state by incluiding a binary per state $STATE_{ks}$ (leaving out one state as the category of reference).
    - E.g., $STATE_{CA,s}$ is a binary variable that takes the value of 1 when the observation, indexed by s, is in the state of California and 0 otherwise.

Regression Using All States 3/3

- More changes to the two state regression:
  - Finally, there are two variations required regarding the measurement of treatment (captured before by the interaction $TREAT_s \times POST_t$):
    - Time and location of treatment application cannot be pinned down with one single interaction
    - Treatment intensity varies across states and time:
      - Some states went form 21 to 18 (similar to $TREAT_s \times POST_t = 1$ before)
      - Other states went, for example, from 18 to 19.
      - To capture this new treatment we defined $LEGAL_{st}$ as the fraction of the population with ages between 18 - 20 that were legaly allowed to drink in state s at time t.

Regression Equation

- Given the definitions for $LEGAL_{st}, STATE_{ks}, YEAR_{j,t}$, and of an outcome Y_{st} that measures the death rates for 18 - 20 years-olds in state s at time t our regression equations for the period 1970 to 1983 is:

Regression Equation

- Given the definitions for $LEGAL_{st}, STATE_{ks}, YEAR_{j,t}$, and of an outcome Y_{st} that measures the death rates for 18 - 20 years-olds in state s at time t our regression equations for the period 1970 to 1983 is:

$$Y_{st} = \alpha + \delta_{DD} LEGAL_{st} + \ldots$$

Regression Equation

- Given the definitions for $LEGAL_{st}, STATE_{ks}, YEAR_{j,t}$, and of an outcome Y_{st} that measures the death rates for 18 - 20 years-olds in state s at time t our regression equations for the period 1970 to 1983 is:

$$Y_{st} = \alpha + \delta_{DD} LEGAL_{st} + \sum_{k=Alaska}^{Wyoming} \beta_k STATE_{ks} + \ldots$$

Regression Equation

- Given the definitions for $LEGAL_{st}, STATE_{ks}, YEAR_{j,t}$, and of an outcome Y_{st} that measures the death rates for 18 - 20 years-olds in state s at time t our regression equations for the period 1970 to 1983 is:

$$Y_{st} = \alpha + \delta_{DD} LEGAL_{st} + \sum_{k=Alaska}^{Wyoming} \beta_k STATE_{ks} + \sum_{j=1971}^{1983} \gamma_j YEAR_{jt} + e_{st}$$

Two-Way Fixed Effect = Generalized DD

$$Y_{st} = \alpha + \delta_{DD} LEGAL_{st} + \sum_{k=Alaska}^{Wyoming} \beta_k STATE_{ks} + \sum_{j=1971}^{1983} \gamma_j YEAR_{jt} + e_{st}$$

- The variables $STATE_{ks}, YEAR_{j,t}$ are known as state and year fixed effects. Combined in one regression equation are sometimes called two-way fixed effect model.

Two-Way Fixed Effect = Generalized DD

$$Y_{st} = \alpha + \delta_{DD} LEGAL_{st} + \sum_{k=Alaska}^{Wyoming} \beta_k STATE_{ks} + \sum_{j=1971}^{1983} \gamma_j YEAR_{jt} + e_{st}$$

- The variables $STATE_{ks}, YEAR_{j,t}$ are known as state and year fixed effects. Combined in one regression equation are sometimes called two-way fixed effect model.
- This data structure where there are observations across an entity dimension (state) and another dimension (typically time), is called a **panel data**.

Two-Way Fixed Effect = Generalized DD

$$Y_{st} = \alpha + \delta_{DD} LEGAL_{st} + \sum_{k=Alaska}^{Wyoming} \beta_k STATE_{ks} + \sum_{j=1971}^{1983} \gamma_j YEAR_{jt} + e_{st}$$

- The variables $STATE_{ks}, YEAR_{j,t}$ are known as state and year fixed effects. Combined in one regression equation are sometimes called two-way fixed effect model.
- This data structure where there are observations across an entity dimension (state) and another dimension (typically time), is called a **panel data**.
- We have just seen how panel data estimation with fixed effects for its two dimensions, is a generalized version of the DD estimation method!
- The books makes this connection but it does not emphasize it enough (given the widespread use of "FE" terminology in economics these days).

Results

- Focus on column 1 for now.
- Qualitatively similar effect to the RDD study (7.7-9.6) for all deaths.
- Slightly larger effects on MVA deaths than RDD study (4.5 - 5.9)
- Smaller effects on suicide deaths
- Similar effects on internal deaths (non alcohol related)

TABLE 5.2
Regression DD estimates of MLDA effects on death rates

Dependent variable	(1)	(2)	(3)	(4)
All deaths	10.80	8.47	12.41	9.65
	(4.59)	(5.10)	(4.60)	(4.64)
Motor vehicle accidents	7.59	6.64	7.50	6.46
	(2.50)	(2.66)	(2.27)	(2.24)
Suicide	.59	.47	1.49	1.26
	(.59)	(.79)	(.88)	(.89)
All internal causes	1.33	.08	1.89	1.28
	(1.59)	(1.93)	(1.78)	(1.45)
State trends	No	Yes	No	Yes
Weights	No	No	Yes	Yes

Notes: This table reports regression DD estimates of minimum legal drinking age (MLDA) effects on the death rates (per 100,000) of 18–20-year-olds. The table shows coefficients on the proportion of legal drinkers by state and year from models controlling for state and year effects. The models used to construct the estimates in columns (2) and (4) include state-specific linear time trends. Columns (3) and (4) show weighted least squares estimates, weighting by state population. The sample size is 714. Standard errors are reported in parentheses.

Relaxing the parallel trends assumption

- Whenever there is more data on previous trends (before the treatment), the parallel trends assumption can be relaxed by controlling for a different slope for each state over time.
- When relaxing this assumption DD will only be able to identify large and sharp effects. If the effects are small and/or appear in the outcomes slowly over time, this modification will not find it.

$$Y_{st} = \alpha + \delta_{DD} LEGAL_{st} + \sum_{k=Alaska}^{Wyoming} \beta_k STATE_{ks} + \sum_{j=1971}^{1983} \gamma_j YEAR_{jt} + \sum_{k=Alaska}^{Wyoming} \theta_k (STATE_{ks} \times t) + e_{st}$$

Illustration of Parallel Trends

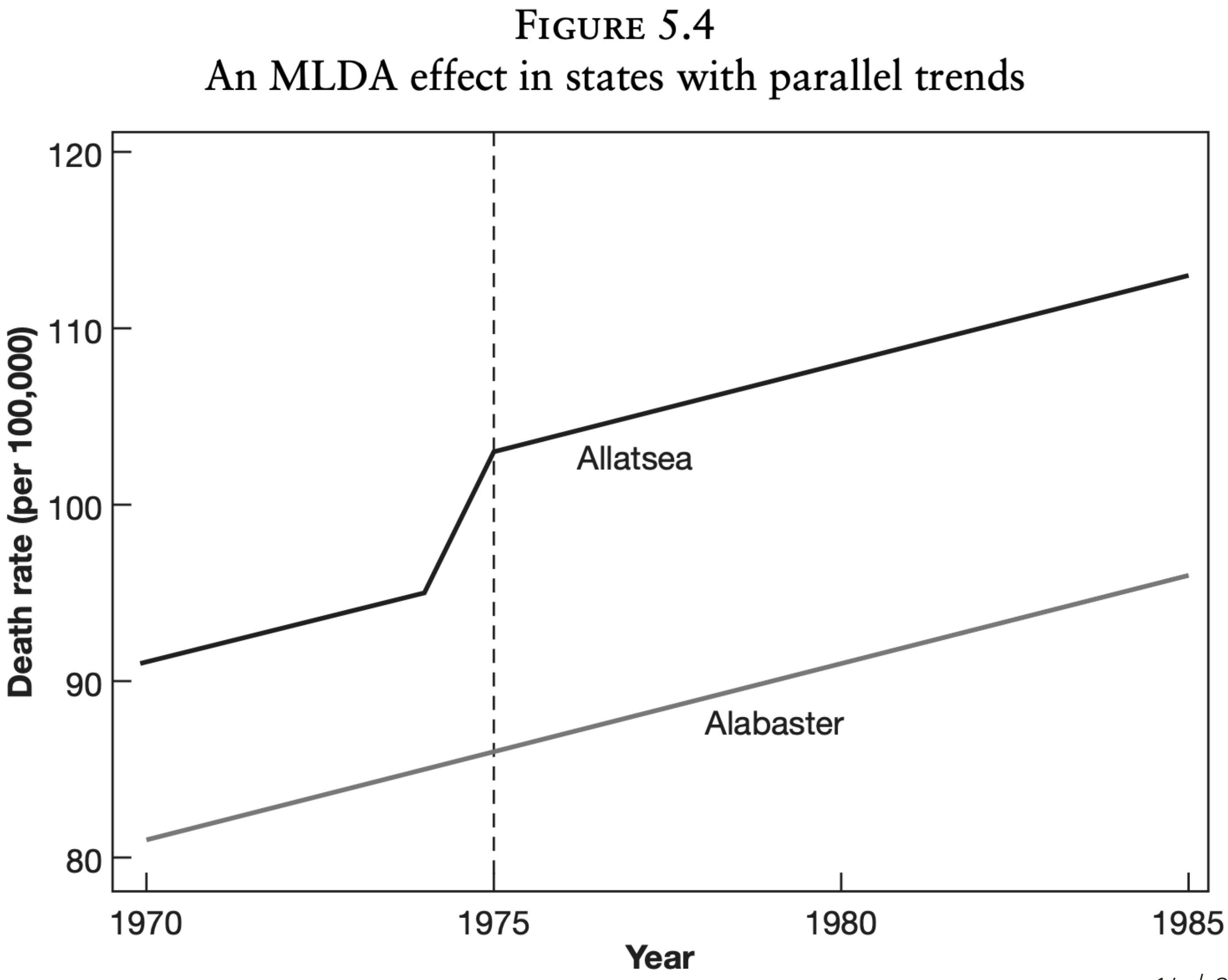


Illustration of No Parallel Trends: No Effect

- Here, the DD estimation without trends would find an effect where there is none.
- There DD estimation with the trends will find no effect.

FIGURE 5.5
A spurious MLDA effect in states where trends are not parallel

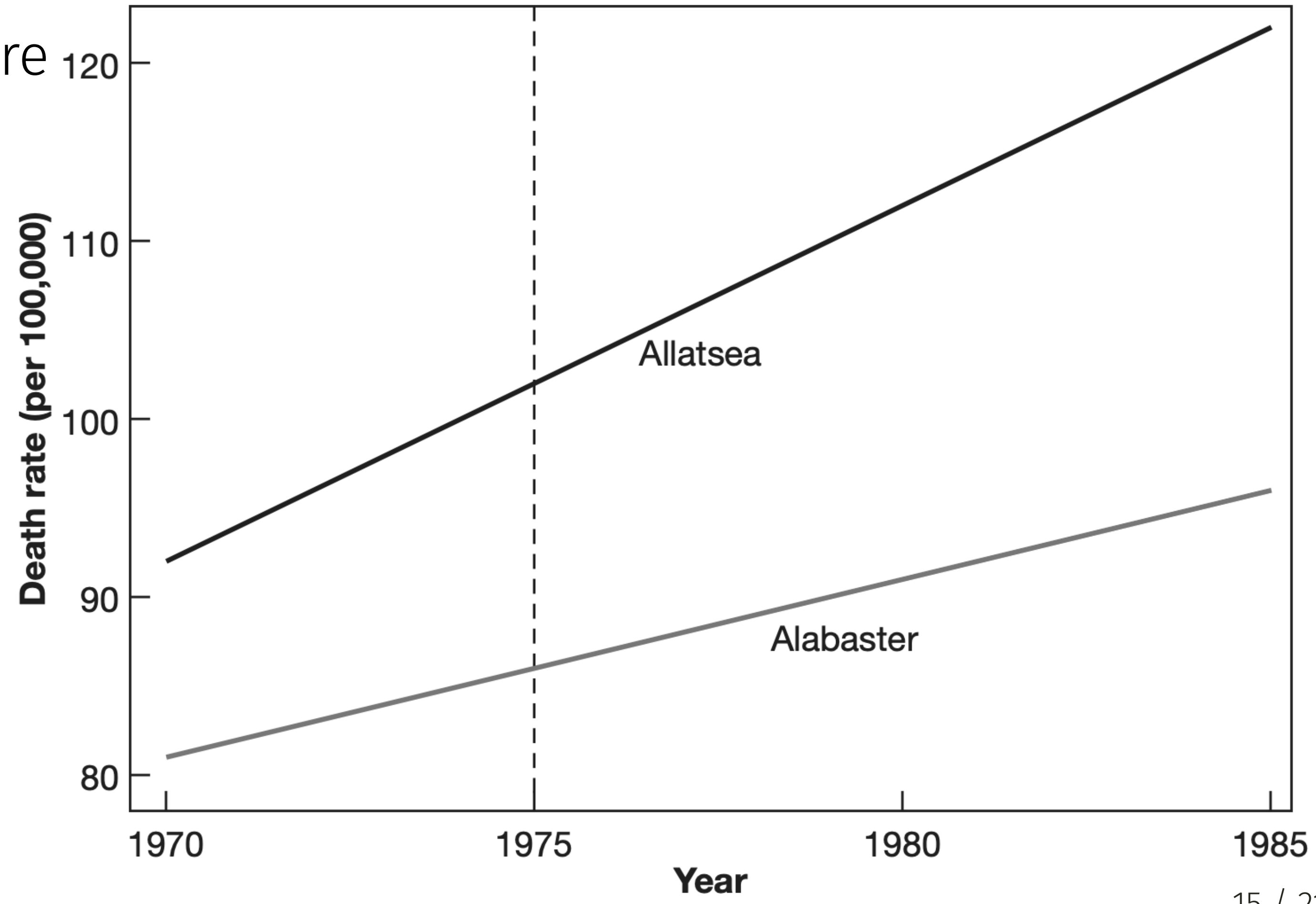
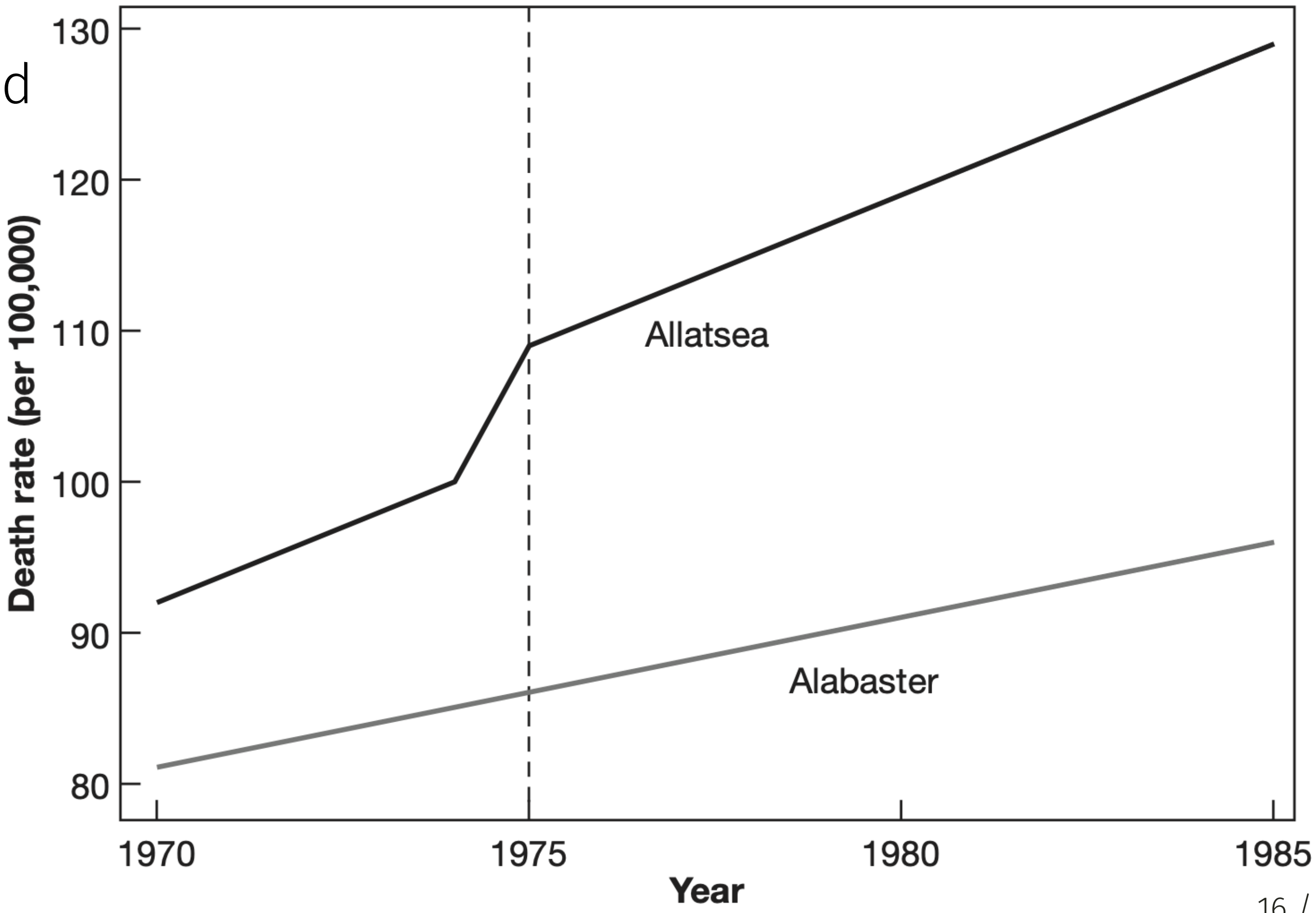


Illustration of No Parallel Trends: Positive Effect

- Here, both the DD estimation with and without trends would find an effect.
- The effect with trend would more smaller and more accurate.

Figure 5.6
A real MLDA effect, visible even though trends are not parallel



Snow example

FIGURE 5.7
John Snow's DD recipe

TABLE XII.

Sub-Districts.	Deaths from Cholera in 1849.	Deaths from Cholera in 1854.	Water Supply.
St. Saviour, Southwark	283	371	Southwark & Vauxhall Company only.
St. Olave	157	161	
St. John, Horsleydown	192	148	
St. James, Bermondsey	249	362	
St. Mary Magdalen	259	244	
Leather Market	226	237	
Rotherhithe*	352	282	
Wandsworth	97	59	
Battersea	111	171	
Putney	8	9	
Camberwell	235	240	
Peckham	92	174	
Christchurch, Southwark	256	113	Lambeth Company, and Southwark and Vauxhall Compy.
Kent Road	267	174	
Borough Road	312	270	
London Road	257	93	
Trinity, Newington	318	210	
St. Peter, Walworth	446	388	
St. Mary, Newington	143	92	
Waterloo Road (1st)	193	58	
Waterloo Road (2nd)	243	117	
Lambeth Church (1st)	215	49	
Lambeth Church (2nd)	544	193	
Kennington (1st)	187	303	
Kennington (2nd)	153	142	
Brixton	81	48	
Clapham	114	165	
St. George, Camberwell	176	132	
Norwood	2	10	Lambeth Company only.
Streatham	154	15	
Dulwich	1	—	
Sydenham	5	12	
First 12 sub-districts	2261	2458	Southwk. & Vauxhall.
Next 16 sub-districts	3905	2547	Both Companies.
Last 4 sub-districts	162	37	Lambeth Company.

* A small part of Rotherhithe is now supplied by the Kent Water Company.

Minimum Wage Example

- Paper here
- Slides from another course here

Mariel Boatlift Example

- Paper here
- Slides from another course here or here

Final Condideration of DD: The Key Requirement Variation Over Time

- Remember the short description of MM about DD: “The DD tool amounts to a comparison of trends over time”
- Implicit in this statement is that DD depends on variation in the changes of a variable over time (in addition to betwen treatment and control).
- This approach has the big benefit of removing any OVB that is constant over time. But it comes at the costs of loosing all the variation within a specific time period.
- Less variation in the data will imply larger SEs, hence it will be harder to detect significance (or easier to not reject the null).

Acknowledgments

- MM