UNIVERSITY OF ILLINOIS
AT URBANA-CHAMPAIGN

Difference-in-Differences Estimators

1867

illinois.edu

Outline for the Session

1. Definition of the Average Treatment Effect (ATE)
2. Single difference
3. Constructing the DiD estimator
4. Key assumptions and hurdles to measurement
5. Triple difference

Single Difference

Revisiting the Irrigation Problem

- We know that comparing those with irrigation to those without is a problem due to selection bias.
- What about comparing production before and after irrigation for those who participated in the project?
- Reflexive estimate
- Why is this a problem?

Example—Is the Effect 200?

	Cereal Production Before	Cereal Production After
Household participated in irrigation project	1,500 kg/hectare	1,700 kg/hectare

Problem

- Other unobserved characteristics may have changed also that are unrelated to irrigation.
  - New seed program
  - Good weather

Constructing the DiD Estimator

Difference-in-Differences (DiD)

- Requires data collected before and after the intervention for treated households and control households
- This method controls for unobserved characteristics that do not change over time.

Example—Add a Control Group

	Cereal Production Before	Cereal Production After
Household participated in irrigation project	1,500 kg/hectare	1,700 kg/hectare
Household did not participate in irrigation project	1,500 kg/hectare	1,600 kg/hectare

$$DiD = (1{,}700 - 1{,}500) - (1{,}600 - 1{,}500) = 100$$

- We would do a t-test to see if the differences were statistically significant

Revisiting Millennium Villages Project

- Designed without a proper control group
  - Introduced control group in year 3
- Lancet paper reports child mortality fell 5.9% per year
- Looks great! But….
  - Original paper reported 2.6 % decline in other rural areas
- However, child mortality was falling very rapidly, so corrected decline with new data was 6.4%
  - Mortality fell more slowly in MVP villages.

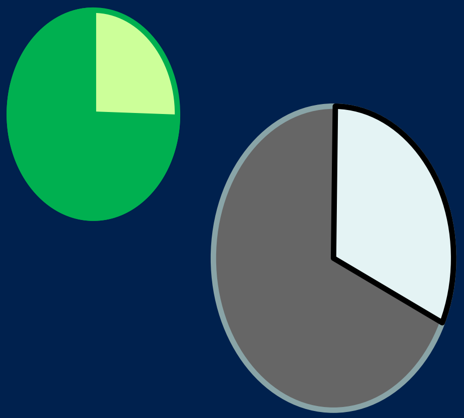
Difference-in-Difference (DiD)

	Treatment (Reserve + Payment)	Control (outside Reserve)
Befor	To	Co
Difference in Difference		ΔT - ΔC

- Control and Treatment groups should be similar
- Choice of Treated cannot be correlated with unobservables that affect outcome
- Control must be 'uncontaminated'. i.e. not affected by treatment
- Works best with random placement. Very rare.

Difference-in-Difference (DiD)

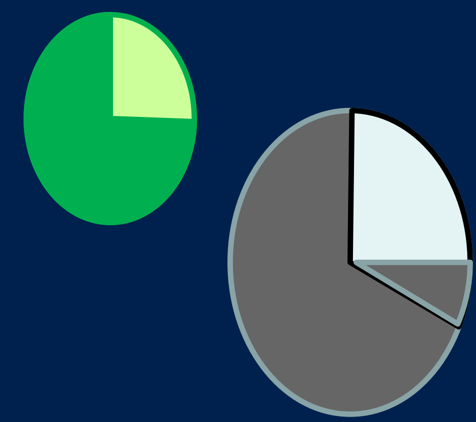
	Treatment (Reserve + Payment)	Control (outside Reserve)
Before	To	Co
Difference in Difference		**ΔT - ΔC**



- Control and Treatment groups should be similar
- Choice of Treated cannot be correlated with unobservables that affect outcome
- Control must be 'uncontaminated'. i.e. not affected by treatment
- Works best with random placement. Very rare.

Difference-in-Difference (DiD)

	Treatment (Reserve + Payment)	Control (outside Reserve)
Before	T_0	C_0
After	T_1	C_1



- Control and Treatment groups should be similar
- Choice of Treated cannot be correlated with unobservables that affect outcome
- Control must be 'uncontaminated'. i.e. not affected by treatment
- Works best with random placement. Very rare.

Difference-in-Difference (DiD)

	Treatment (Reserve + Payment)	**Control (outside Reserve)**
Before	T_0	C_0
After	T_1	C_1
Change	ΔT	ΔC
Difference in Difference		$\Delta T - \Delta C$

- Control and Treatment groups should be similar
- Choice of Treated cannot be correlated with unobservables that affect outcome
- Control must be 'uncontaminated'. i.e. not affected by treatment
- Works best with random placement. Very rare.

What Is the DiD Estimator?

- Non-experimental method
  - Could be used with a natural experiment
  - Could be used with an RCT
- Controls for unobserved characteristics that do not change over time.
- Typically, exploit time dimension. Use data collected before and after the intervention in treatment and control areas.
- Could involve geographic variation as well in only one time period

Treatment and Control

$$DD = E[Y_1^T - Y_0^T | T_1 = 1] - E[Y_1^T - Y_0^T | T_1 = 0]$$

- Change in the control group is the counterfactual for the treatment group
- Regression context

$$Y_{it} = \alpha + \beta T_{i1} t + \rho T_{i1} + \gamma t + \varepsilon_{it}$$

- Treatment effect is coefficient on the interaction between time and treated area

Example

- Alderman (2007)
- Nutrition and Early Child Development Program in Uganda
- Government chose intervention areas
- Program not randomized
- Choose control areas close to treatment areas
- Baseline and follow up

Relate Treatment Effect to Regression

- $E\left[Y_1^T - Y_0^T \middle| T_1 = 1\right] = (\alpha + \beta + \rho + \gamma) - (\alpha + \rho)$
- Get this from regression equation by turning on dummies in period 2.

$$E\left[Y_1^C - Y_0^C \middle| T_1 = 0\right] = (\alpha + \gamma) - \alpha$$

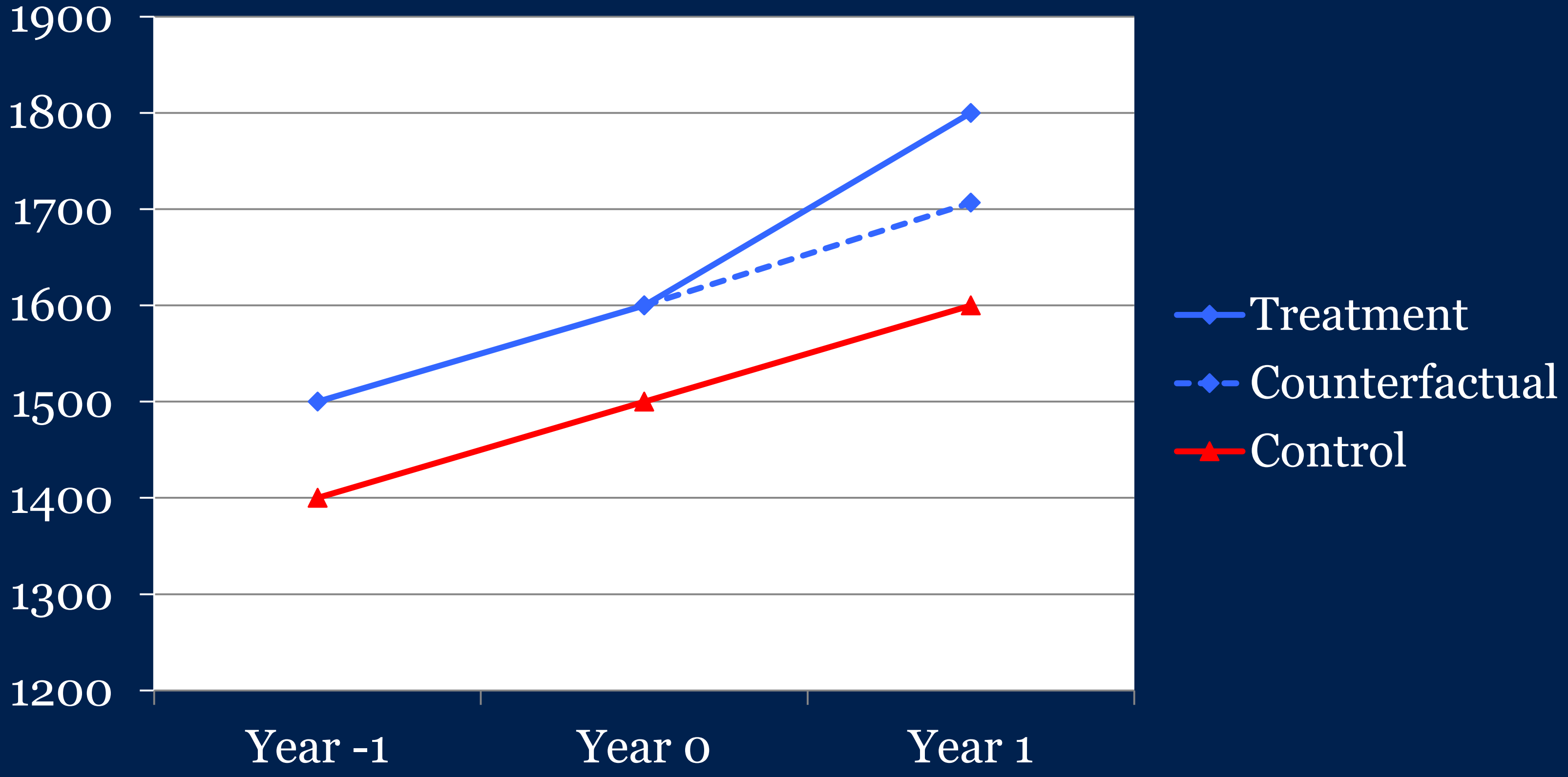
- Subtract to get a DD.
- In reflexive design, estimate equals $\beta + \gamma$, and γ is the bias due to time.

Key Assumptions and Hurdles to Measurement

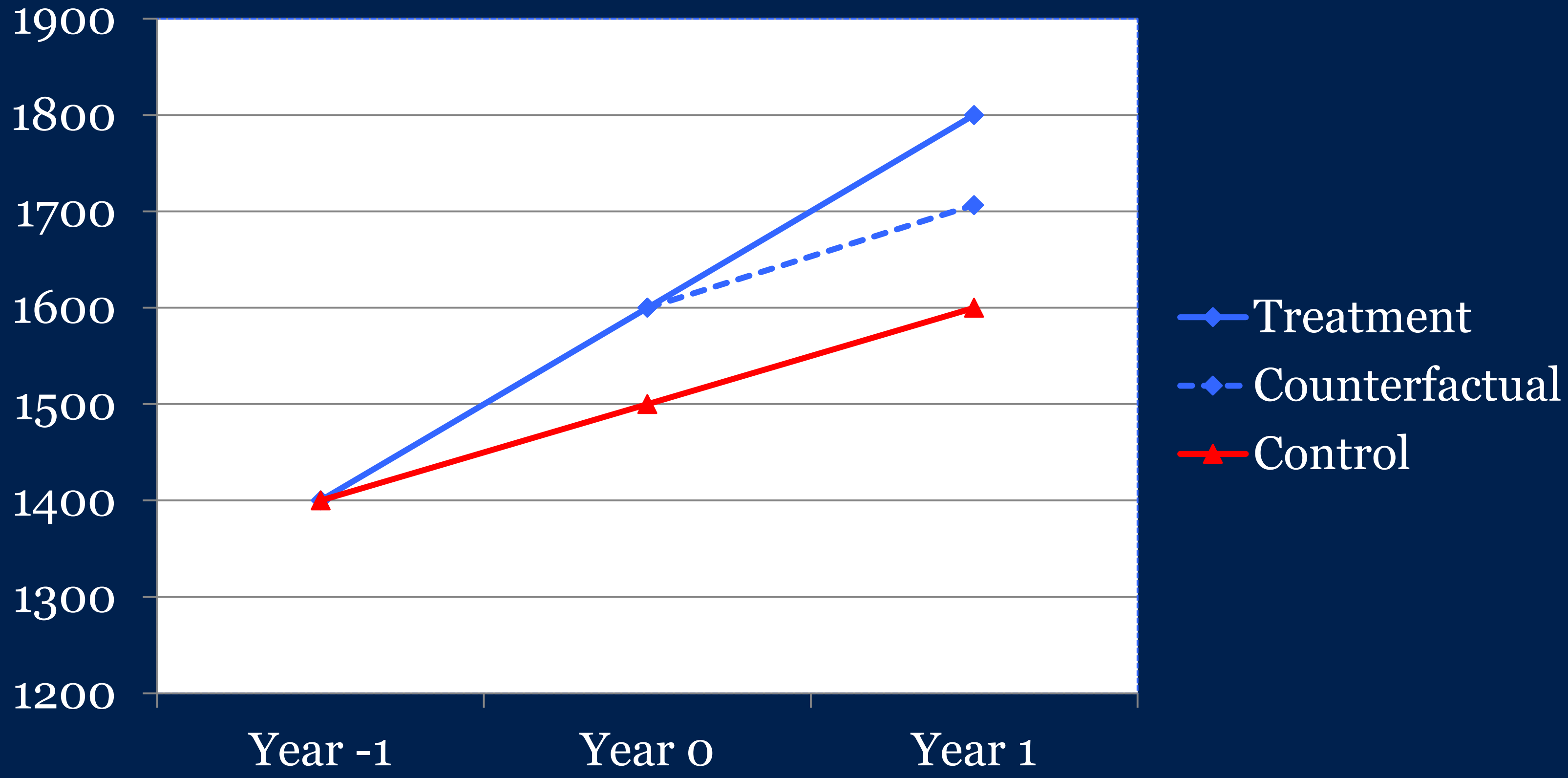
Key Assumptions

- Model is correctly specified with an additive error term
- SUTVA
- OK for selection to be correlated with the time invariant part of the error term
- Correlation between time varying part of the error term and treatment is 0
- Parallel trend assumption

Parallel Trend Assumption (cereal yield per hectare)



What if Trends Were Like This?



Matching and DiD

- What is a good control area for the treatment area in diff in diff?
- Use matching to locate a similar area to the treatment (remember, treatment not random)
- Similarity on observables, outcome at baseline
- Discussion on matching from yesterday is relevant—match in baseline
- Diff in diff solves the problem of selection on time invariant unobservables

Multiple Time Periods

- Panel data

$$Y_{it} = \varphi T_{it} + \delta X_{it} + \eta_i + \varepsilon_{it}$$

- η is the fixed effect. Unobservable that is constant over time
- First differencing or mean differencing causes η to drop out of the equation

Example with VLS data

- Deolalikar and Rose J Popul Econ (1998)
- Looked at impact of the birth of a boy vs. birth of a girl on savings, consumption, income
- Can be viewed as a DiD estimator
- First differences—effect of having a boy vs. no birth, effect of having a girl vs. no birth
- DiD—Difference between having a boy and having a girl
- For medium and large farm households, birth of a son reduces savings relative to birth of girl

Can We Trust DiD Standard Error Calculations?

- Widely cited paper by Bertrand, Duflo, Mullainathan (2004)
- Naïve estimates under assumption that errors are not serially correlated
- But, panel data follows the same unit over time. Errors are serially correlated
- Also, treatments are not turned on and off. They tend to start and remain in effect, which decreases random variation

Correcting the Standard Errors

- Standard errors can be severely underestimated
- Type I errors result
- Solutions
  - Block bootstrap if the number of units (states, villages) is large enough
  - Estimate the variance-covariance matrix directly and model the autocorrelation
  - Collapse time-series data into two periods—pre- and post-treatment

Triple Difference

What is a Triple Difference?

- Typically, we compare the changes in two groups over time, the treatment group and the control group.
- Usually, groups defined based on geography
- Add another difference—by age
- Laws that affect those who are under 18
- Compare 16 to 18 year olds with 19 to 20 year olds within a state
- Compare differences across time and across states

Problems

- Challenge to show parallel trend assumption holds
- Data collected before the baseline
- Missing baseline for control area
- Selection on time-varying unobservables
  - Severe drought affects treatment area in year 1 and pushes people to try irrigation