Causal Inference (6.S059/15.C08/17.C08)

Recitation, Week 11.
Topic: Difference-in-Differences and Regression Discontinuity

Licheng Liu

May 3, 2024

MIT

Massachusetts Institute of Technology

Table of contents

1/ DID

Notation for Time Series Data

- Units: $i \in \{1 \dots N\}$
- Time periods (pre/post): $t \in \{0, 1\}$
- Treatment *group* indicator: $D_i \in \{0, 1\}$
- Treatment indicator (whether unit i has been treated at time t): $Z_{it} \in \{0, 1\}^*$

Expected potential outcome for unit i in period t is expressed in the form:

$$\mathbb{E}[Y_{it}(z) | D_i = d]$$

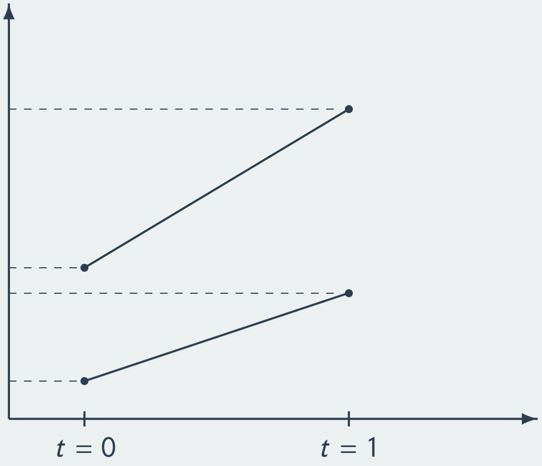
*Note that $Z_{it} = D_i$ in the post period

Notations

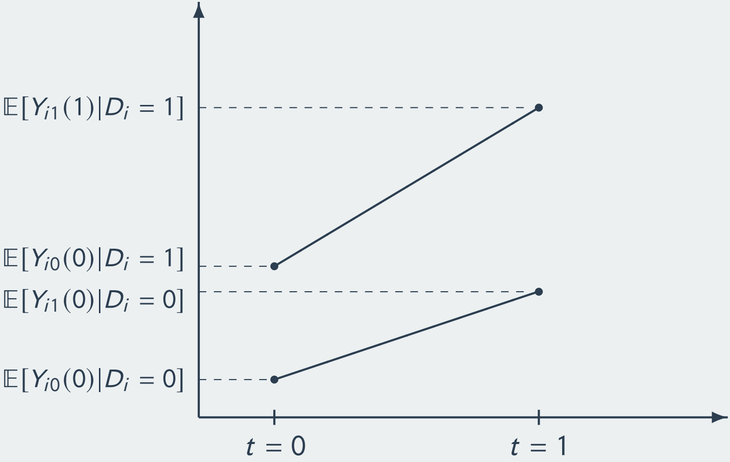
Define the following:

- $\mathbb{E}[Y_{i0}(0)|D_i = 1]$: Expected potential outcome for units in the *treatment* group, in the period *before* treatment occurs
- $\mathbb{E}[Y_{i1}(1)|D_i = 1]$: Expected potential outcome for units in the *treatment* group, in the period *after* treatment occurs
- $\mathbb{E}[Y_{i1}(0)|D_i = 0]$: Expected potential outcome for units in the *control* group, in the period *after* treatment occurs
- $Y_{it}(1) - Y_{it}(0)$: The causal effect of the treatment for unit i at time t
- The ATT in the post-treatment period: $\mathbb{E}[Y_{i1}(1) - Y_{i1}(0)|D_i = 1]$

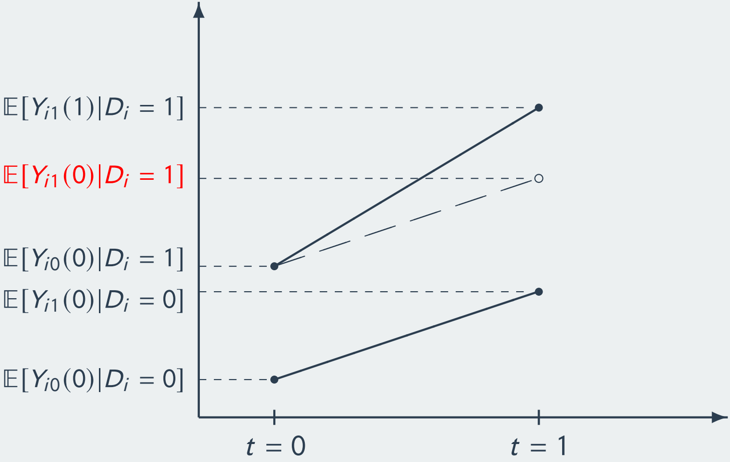
Graphical Version



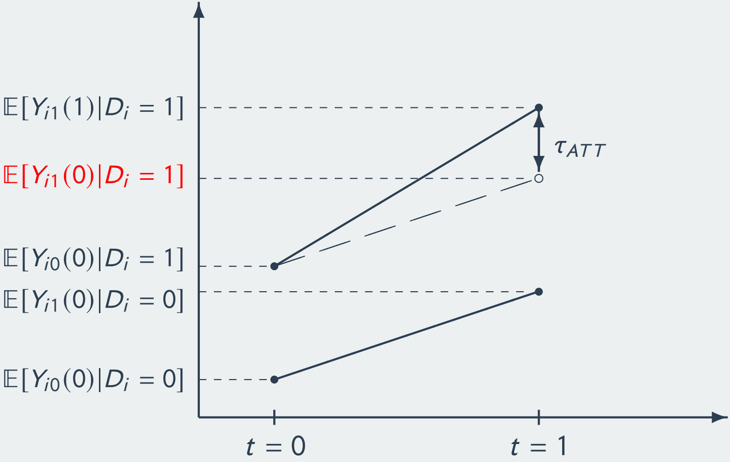
Graphical Version



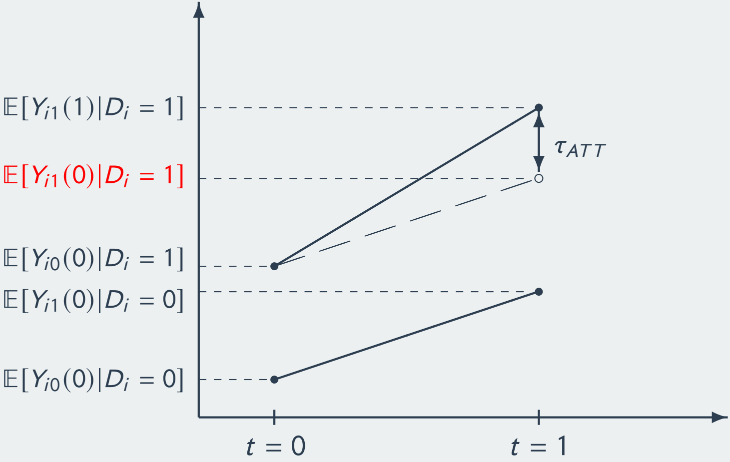
Graphical Version



Graphical Version

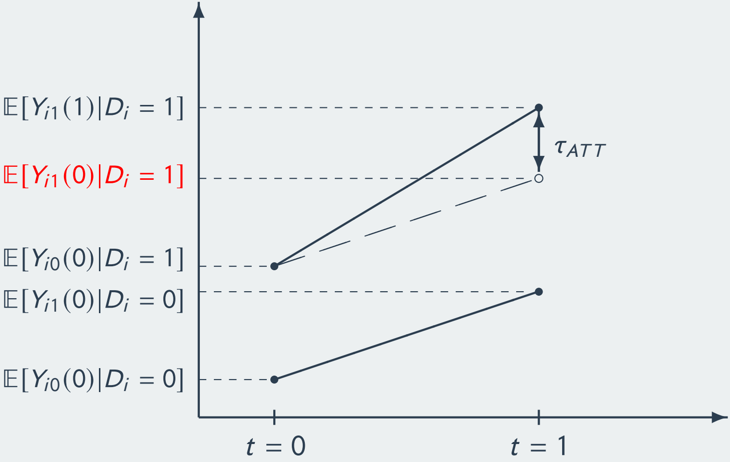


Graphical Version



What assumption do we need to identify the ATT?

Graphical Version



What assumption do we need to identify the ATT? Parallel trends assumption.

$\mathbb{E}[Y_{i1}(0)|D_i = 1] - \mathbb{E}[Y_{i0}(0)|D_i = 1] = \mathbb{E}[Y_{i1}(0)|D_i = 0] - \mathbb{E}[Y_{i0}(0)|D_i = 0]$

Identification Assumptions

Identification with Difference-in-Differences

Under the parallel trends assumption:

$$\mathbb{E}[Y_{i1}(0) - Y_{i0}(0) | D_i = 1] = \mathbb{E}[Y_{i1}(0) - Y_{i0}(0) | D_i = 0]$$

The ATT can be nonparametrically identified as:

$$\begin{aligned} \tau_{ATT} &= \Big\{\mathbb{E}[Y_{i1} | D_i = 1] - \mathbb{E}[Y_{i1} | D_i = 0]\Big\} \\ &\quad - \Big\{\mathbb{E}[Y_{i0} | D_i = 1] - \mathbb{E}[Y_{i0} | D_i = 0]\Big\} \end{aligned}$$

An Example

JAMA 2018: “The April 20 Cannabis Celebration and Fatal Traffic Crashes in the United States”

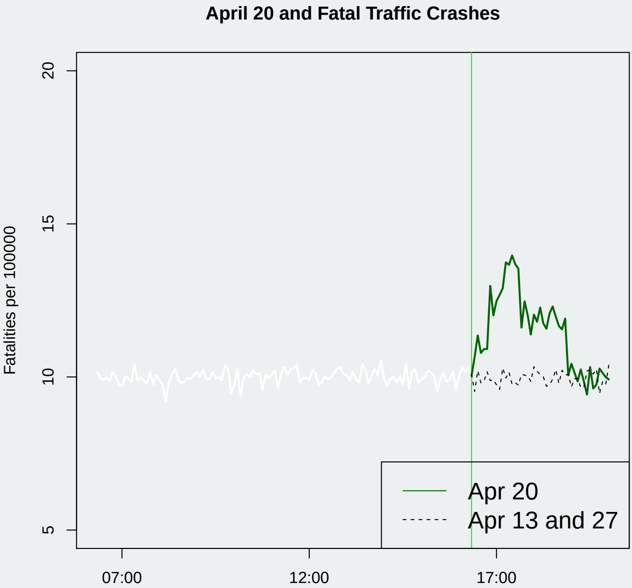
Methods:

- Compare the # of fatal crashes between 4:20 PM and 11:59 PM on April 20 to the number of fatal traffic crashes during the same time intervals on control days 1 week earlier and 1 week later (i.e. April 13 and April 27).

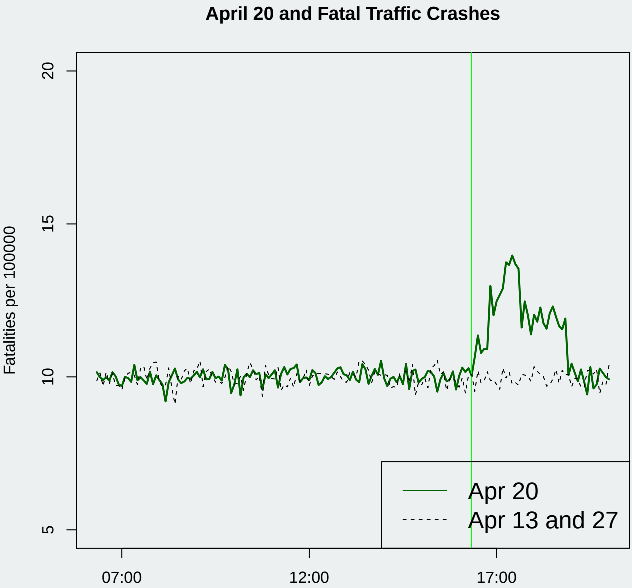
Results:

- Found a 12% increase in the relative risk of a fatal traffic crash after 4:20 PM on April 20 compared with identical time intervals on control days.

Cannabis Celebration and Fatal Traffic Crashes

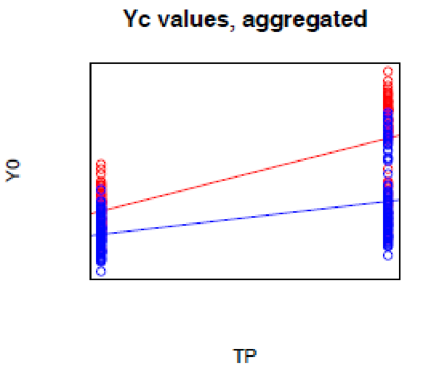


Cannabis Celebration and Fatal Traffic Crashes



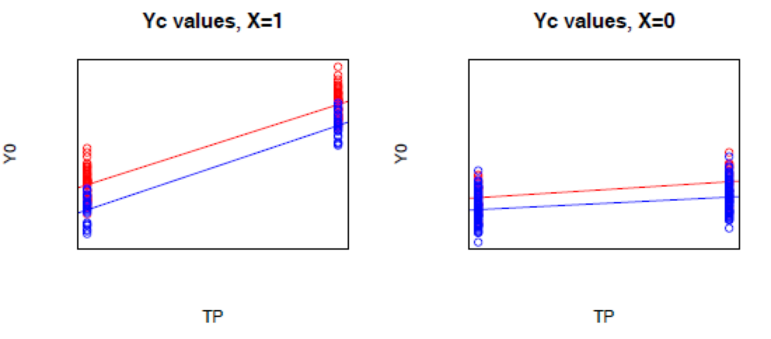
Conditional DID

Parallel trends assumption is violated



Conditional DID

Parallel trends assumption is conditionally satisfied



Estimation of Conditional DID

Estimation under the Parallel trends assumption

$$Y_i = \mu + \gamma D_i + \delta T_i + \tau D_i T_i$$

where μ, γ, δ and τ are OLS regression estimates

Estimation under the Conditional DiD assumption

$$Y_i = \mu + \gamma D_i + \delta T_i + \tau D_i T_i + \tau X_i T_i$$

where x is a time-invariant pre-treatment confounder for unit i.

Estimation of Conditional DID

Estimation under the Parallel trends assumption

$$Y_{it}(z) = \alpha_i + \gamma t + \tau z + \varepsilon_{it}$$

where α_i is a time-invariant unobserved effect for unit i that may be correlated with treatment.

Estimation under the Conditional DiD assumption

$$Y_{it}(z) = \alpha_i + \gamma t + \lambda t \cdot x + \tau z + \varepsilon_{it}$$

where x is a time-invariant pre-treatment confounder for unit i.

2/ Regression Discontinuity

RD Assumptions

What assumption do we need for a RD?

$$\mathbb{E}\left[Y_i(d)|X_i = x\right] \text{ is continuous in } x \text{ around } X_i = c \text{ (for } d = 0, 1)$$

In words? The expected values of potential outcomes are continuous in x, i.e., they do not jump at cut-point (threshold) c.

What is the causal estimand of a RD?

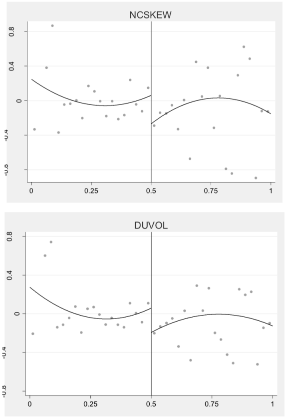
$$\tau_{RD} = \mathbb{E}\left[Y_i(1) - Y_i(0)|X_i = c\right]$$

In words? Local ATE at the threshold.

Choices

- Bandwidth: how far away from the cutoff are your included observations (can outsource this to an algorithm).
- Model for the potential outcomes.
- In both cases we are making a bias-variance trade-off.

Choices: Be Careful!



Notes. This figure presents regression discontinuity plots using a fitted quadratic polynomial estimate with 95% confidence intervals. The x axis is the percentage of votes favoring unionization and the dots depict the average crash risk variables in each of 20 equally spaced bins (with a 5.0% bin width). Union election results are from the NLRB over 1980-2015. Crash risk variables (*NCSKEW* and *DUVOL*) are calculated from CRSP over 1981-2016.