ISEA Week 13 – Causal Inference: Difference-in-Differences

Dr. Chris Candelaria, Vanderbilt University

IES Institute of Education Sciences
W AmplifyLearn.AI
UNIVERSITY of WASHINGTON eScience Institute ADVANCING DATA-INTENSIVE DISCOVERY IN ALL FIELDS
UNIVERSITY OF OREGON

Agenda

1. Difference-in-Differences (DiD): The Basics
2. Potential Outcomes in a DiD framework
3. Example of DiD usage in education policy
4. Application: Using DiD approach to assess user learning in an A/B test environment
   - How should we think about causality?
5. Coding

IES Institute of Education Sciences
W AmplifyLearn.AI
UNIVERSITY of WASHINGTON eScience Institute ADVANCING DATA-INTENSIVE DISCOVERY IN ALL FIELDS
O UNIVERSITY OF OREGON

Difference-in-Differences (Graphical Approach)

> Let's draw!

Difference-in-Differences (DiD) Regression Model

$$Y_i = \alpha + \beta T_i + \gamma G_i + \tau(G_i \times T_i) + \varepsilon_i$$

$$E[Y_i | G_i = 1, T_i = 1] =$$

$$E[Y_i | G_i = 1, T_i = 0] =$$

$$E[Y_i | G_i = 0, T_i = 1] =$$

$$E[Y_i | G_i = 0, T_i = 0] =$$

Defining the Difference-in-Difference (DiD) Treatment Effect

$$\tau^{\mathrm{DiD}} = \left[\mathrm{E}[Y_i \mid G_i = 1, T_i = 1] - \mathrm{E}[Y_i \mid G_i = 1, T_i = 0]\right]$$
$$-\left[\mathrm{E}[Y_i \mid G_i = 0, T_i = 1] - \mathrm{E}[Y_i \mid G_i = 0, T_i = 0]\right]$$

Potential Outcomes Revisited: Estimating the Average Treatment Effect (ATE)

Potential Outcomes Framework: (Holland, 1986)

$$Y_i = Y_i(0) + T_i(Y_i(1) - Y_i(0))$$

If $T_i = 1$: $Y_i = Y_i(1)$

$T_i = 0$: $Y_i = Y_i(0)$

Assuming constant treatment effect:

$$Y_i(1) = Y_i(0) + \tau$$

Average Treatment Effect (ATE):

$$E[Y_i(1) - Y_i(0)] = \tau$$

Use difference in averages to estimate the ATE?

$$E[Y_i|T_i = 1] - E[Y_i|T_i = 0]$$

$$= E[Y_i(1)|T_i = 1] - E[Y_i(0)|T_i = 0]$$

$$= E[Y_i(0) + \tau|T_i = 1] - E[Y_i(0)|T_i = 0]$$

$$= \tau + E[Y_i(0)|T_i = 1] - E[Y_i(0)|T_i = 0]$$

ATE

Selection Bias!

IES Institute of Education Sciences · W · AmplifyLearn.AI · University of Washington eScience Institute: Advancing Data-Intensive Discovery in All Fields · University of Oregon

Potential Outcomes in DiD Framework (I)

Potential Outcomes Framework in DiD: (Athey & Imbens, 2006)

$$Y_i = Y_i(0) + I_i(Y_i(1) - Y_i(0))$$

Let:

$$I_i = G_i \times T_i, \quad G_i \in \{0,1\}, \quad T_i \in \{0,1\}$$

If:

$$I_i = 1 : Y_i = Y_i(1)$$

$$I_i = 0 : Y_i = Y_i(0)$$

We need expressions for $Y_i(1)$ and $Y_i(0)$!

Also, let's assume a constant treatment effect:

$$Y_i(1) = Y_i(0) + \tau$$

Potential Outcomes in DiD Framework (II):

Define: $Y_i(0) = \alpha + \beta T_i + \gamma G_i + \varepsilon_i$

$Y_i(1) = Y_i(0) + \tau$ **Because we assume a constant treatment effect**

Observed Outcome:

$$Y_i = \alpha + \beta T_i + \gamma G_i + \tau I_i + \varepsilon_i$$

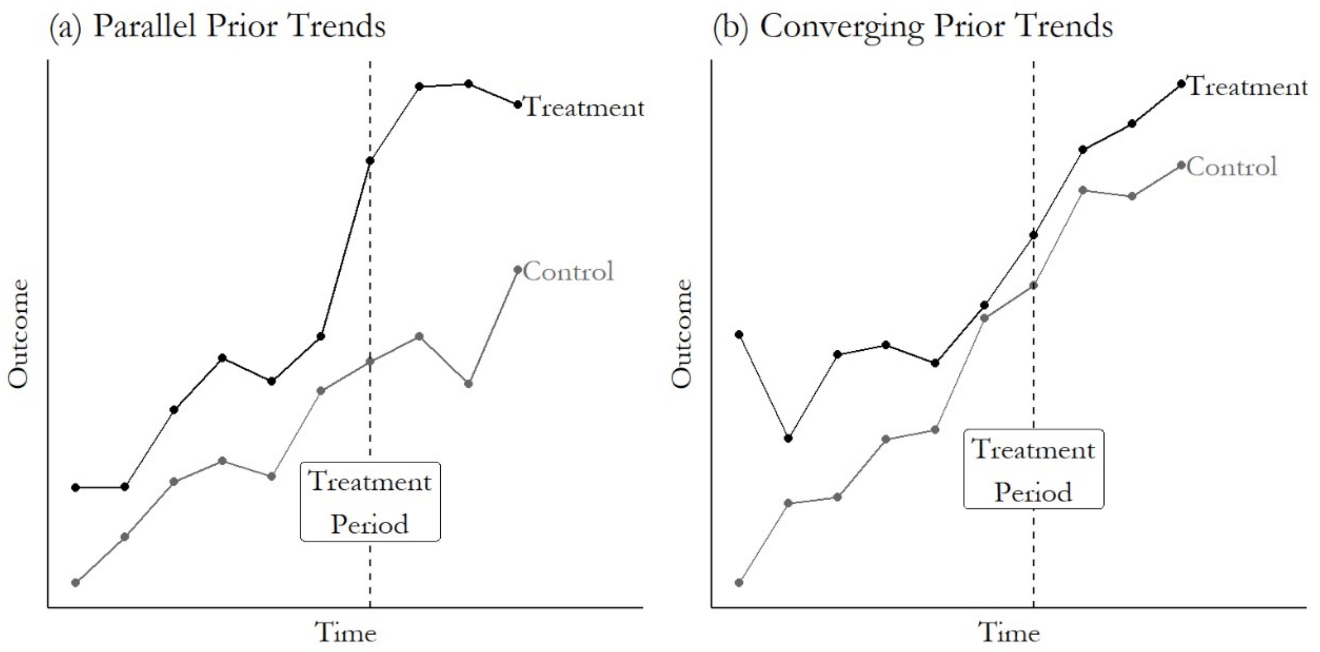
where $I_i = G_i \times T_i$

Interpreting τ as Average Treatment on the Treated (ATT):

$$E[Y_i(1) - Y_i(0) | G_i = 1, T_i = 1] = E[Y_i(1) | G = 1, T = 1] - E[Y_i(0) | G_i = 1, T_i = 1]$$

$$= \tau^{\text{DiD}}$$

Key Assumption: Parallel Trends



Source: Huntington-Klein (2023), Chapter 18

IES Institute of Education Sciences · W AmplifyLearn.AI · UNIVERSITY of WASHINGTON eScience Institute ADVANCING DATA-INTENSIVE DISCOVERY IN ALL FIELDS · UNIVERSITY OF OREGON

Key Assumption: Parallel Trends

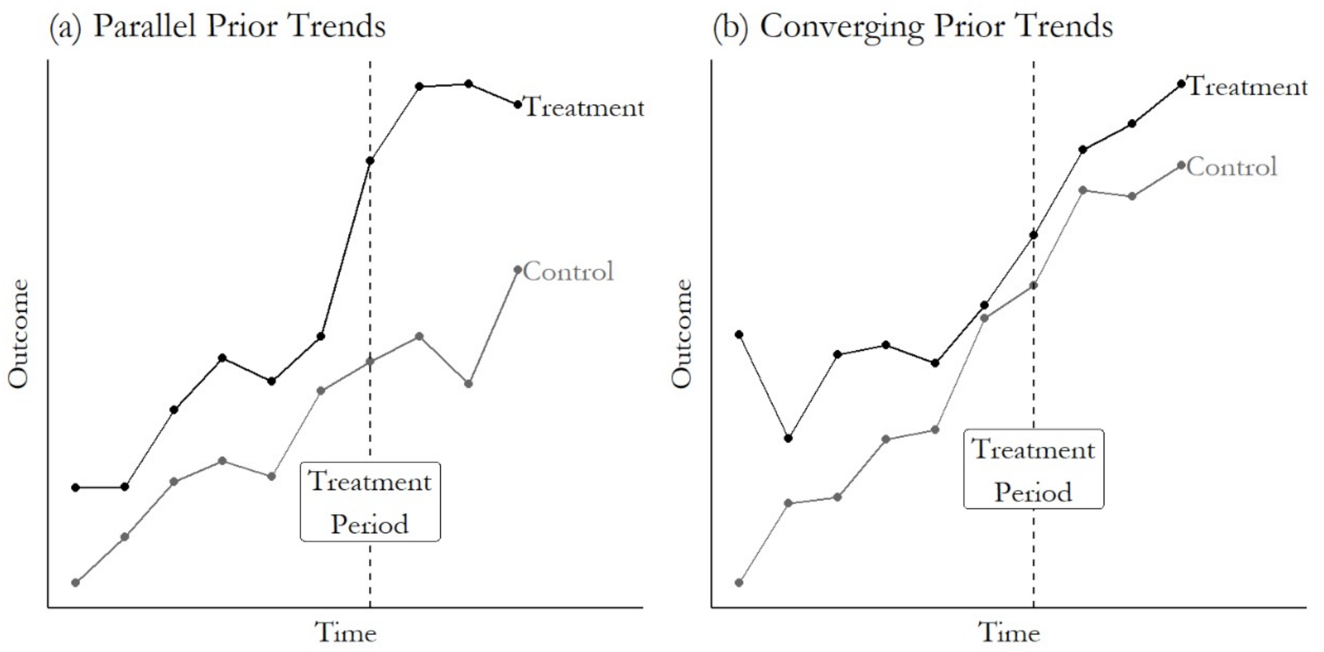
1. Parallel pre-trends (testable):

Before the treatment or intervention is introduced, the outcomes of the treatment group and the control group follow similar trends over time.

2. Common Trends/Shocks (Identifying assumption/untestable):

Any differences observed in the outcomes between the treatment and control groups before the treatment are due to factors other than the treatment itself, and these differences would have persisted in the absence of treatment.

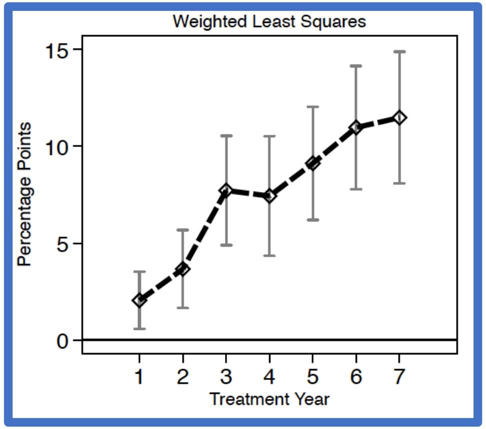
Key Assumption: Parallel Trends



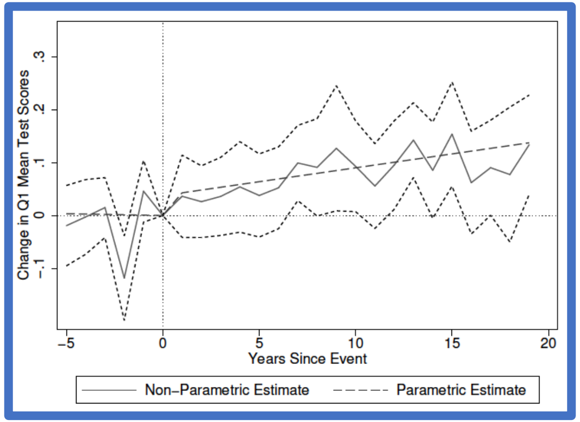
Source: Huntington-Klein (2023), Chapter 18

Institute of Education Sciences
AmplifyLearn.AI
UNIVERSITY of WASHINGTON eScience Institute
ADVANCING DATA-INTENSIVE DISCOVERY IN ALL FIELDS
UNIVERSITY OF OREGON

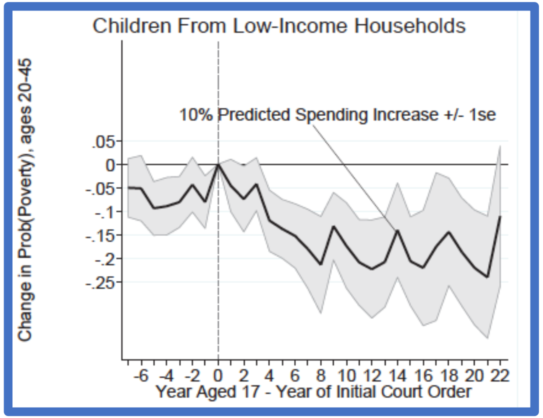
DiD in School Finance Reform Studies: Money Matters for Educational Outcomes



Graduation Rates in High-Poverty Districts ↑
(Candelaria & Shores, 2019)



Test scores for low-SES children ↑
(Lafortune et al., 2018)



Incidence of Poverty ↓
(Jackson et al., 2016)

> Across studies: "On average, a $1000 increase in per-pupil public school spending (for four years) increases test scores by 0.044 standard deviations, high-school graduation by 2.1 percentage points, and college-going by 3.9 percentage points." (Jackson & Mackevicius, 2021)

IES Institute of Education Sciences | W | AmplifyLearn.AI | University of Washington eScience Institute — Advancing Data-Intensive Discovery in All Fields | University of Oregon

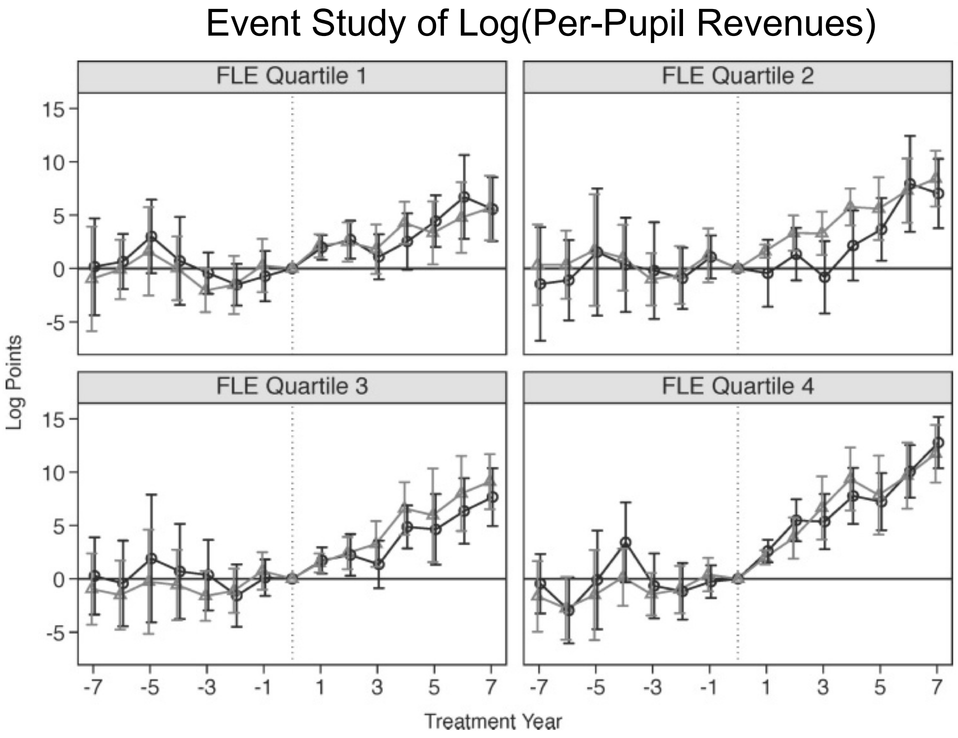
Candelaria & Shores (2019): Funding Effect Example

Let's assess the parallel trends assumption using an event study design:

$$Y_{sqdt} = \theta_d + \delta_{qt}$$

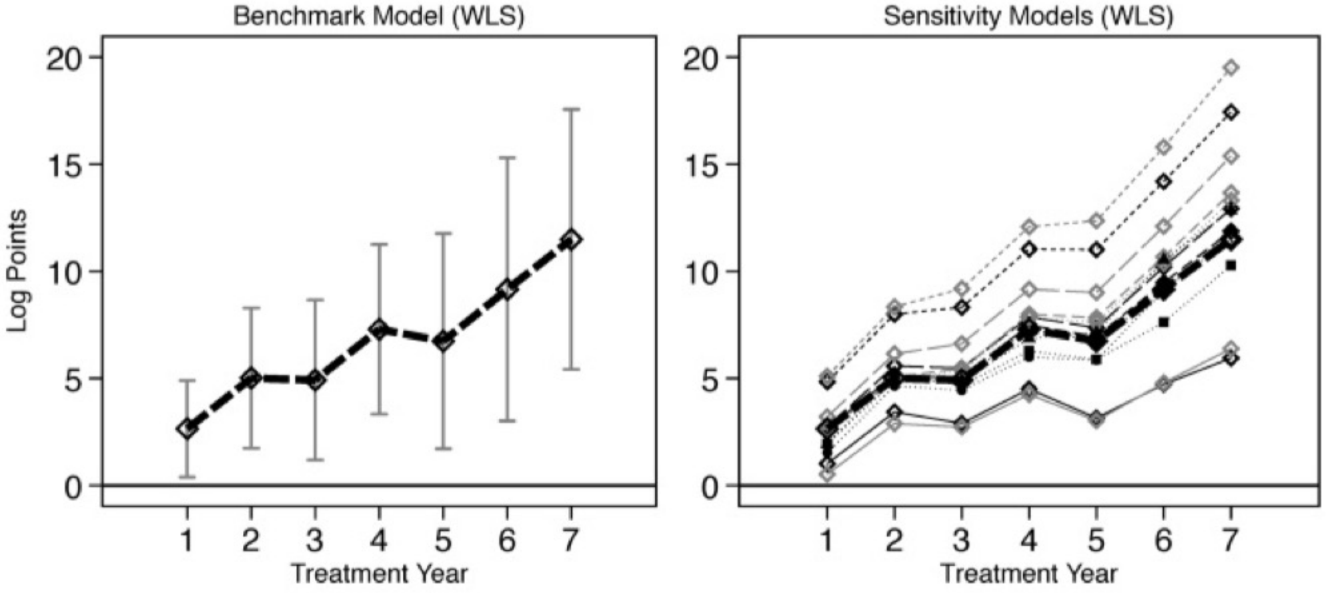
$$+ \sum_{n=1}^{16} \gamma_{(q,-n)}[1(Q_{sd} = q) \times 1(t - t_s^* = -n) \times D_s]$$

$$+ \sum_{n=1}^{19} \gamma_{(q,+n)}[1(Q_{sd} = q) \times 1(t - t_s^* = n) \times D_s] + \varepsilon_{sqdt}$$



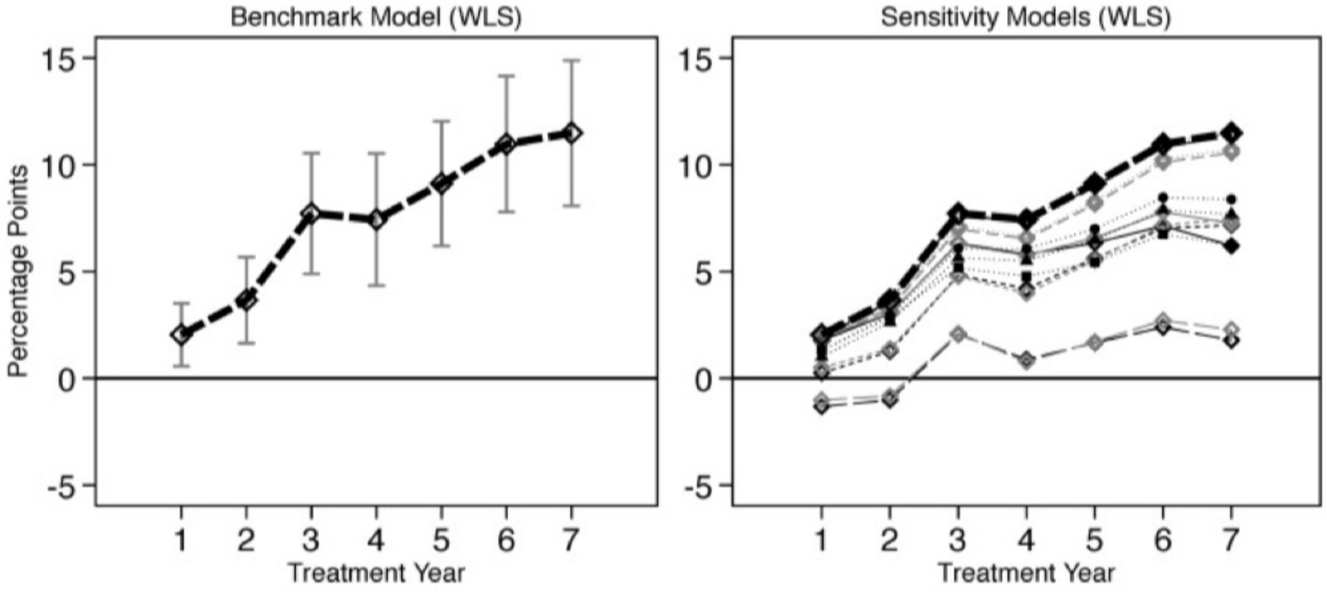
Institute of Education Sciences · AmplifyLearn.AI · University of Washington eScience Institute · University of Oregon

School Finance Reforms (SFRs) Increase Funding Among Lower-Income Districts



Revenues increased by by 11.5% in lower-income districts 7 years after reform

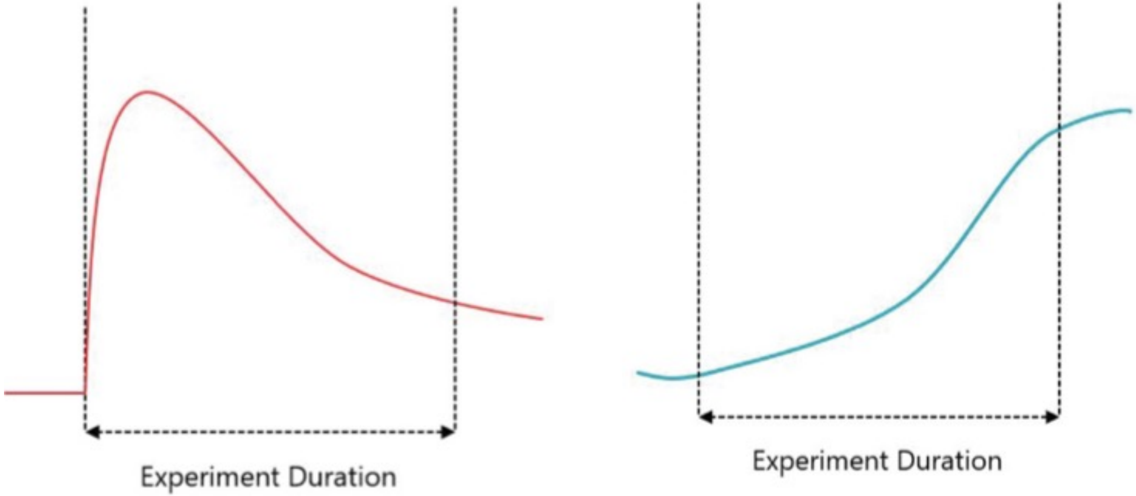
SFRs Increase Graduation Rates Among Lower-Income Districts



Graduation Rates increased by by 12 percentage points in lower-income districts 7 years after reform

Institute of Education Sciences
AmplifyLearn.AI
University of Washington eScience Institute
Advancing Data-Intensive Discovery in All Fields
University of Oregon

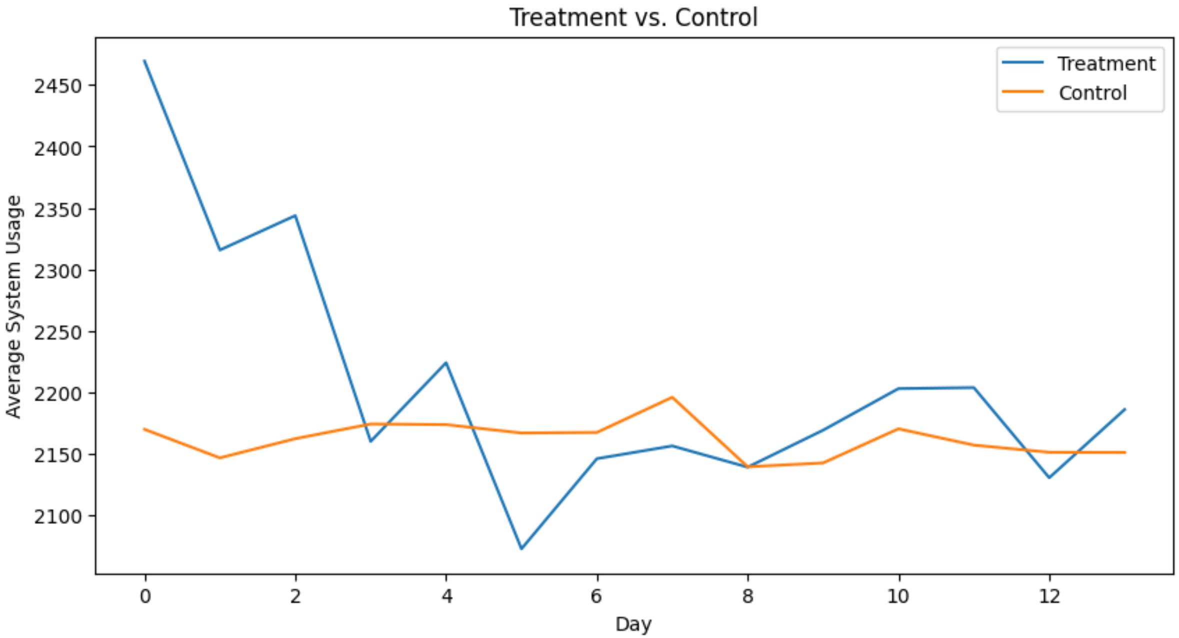
Transition: Using a DiD approach in A/B Testing to Assess User Learning



(a) Novelty effect

(b) Primacy effect

Given A/B test, how should we assess user learning?



Institute of Education Sciences
AmplifyLearn.AI
UNIVERSITY of WASHINGTON eScience Institute
ADVANCING DATA-INTENSIVE DISCOVERY IN ALL FIELDS
UNIVERSITY OF OREGON

Time to code!

- Access the Google Colab site for our coding session