Treatment Effects and Difference-in-Differences

Everything You've Learned, Applied to Policy

Jake Anderson

March 21, 2026

Outline

Outline

A State Raised Its Minimum Wage. Did Employment Fall?

On April 1, 1992, New Jersey raised its minimum wage from $4.25 to $5.05.

Neighboring Pennsylvania kept its minimum wage at $4.25.

Standard prediction: higher minimum wage $\implies$ fewer low-wage jobs.

The question: How do we *measure* the causal effect of this policy?

We cannot observe what *would have happened* in NJ without the increase. This unobserved scenario is the **counterfactual**.

$\implies$ We need a strategy that isolates the policy effect from everything else changing at the same time.

Why Not Just Compare NJ Employment Before vs. After?

Suppose NJ fast-food employment was 20.4 before and 21.0 after the increase.

Can we conclude the minimum wage *increased* employment by 0.6 workers?

No. Other things changed between February and November 1992:

- The national economy was recovering from recession
- Seasonal hiring patterns in fast food
- Changes in consumer demand

A simple before-vs-after comparison confounds the **policy effect** with **time trends**.

$\Longrightarrow$ We need a **control group** that experienced the same time trends but *not* the policy change.

Outline

Treatment and Control Groups

Treatment group: units that receive the intervention (NJ restaurants).
Control group: units that do not receive it (PA restaurants).

Define a treatment indicator:

$$d_i = \begin{cases} 1 & \text{if unit } i \text{ is treated} \\ 0 & \text{if unit } i \text{ is in the control group} \end{cases}$$

The simplest treatment effect model:

$$y_i = \beta_1 + \beta_2 \, d_i + e_i$$

This is just a regression with a dummy variable. You already know how to estimate it.
$\Longrightarrow$ β_2 is the **treatment effect**: the average difference in y between treated and control.

The Difference Estimator

From the model $y_i = \beta_1 + \beta_2 d_i + e_i$:

- Control group ($d_i = 0$): $\text{E}(y_i) = \beta_1$
- Treatment group ($d_i = 1$): $\text{E}(y_i) = \beta_1 + \beta_2$

The OLS estimator of β_2 is simply:

$$b_2 = \bar{y}_1 - \bar{y}_0$$

The **difference estimator**: treatment group mean minus control group mean.

When is this valid?

$b_2 = \beta_2 + (\bar{e}_1 - \bar{e}_0)$

For $\text{E}(b_2) = \beta_2$, we need $\text{E}(\bar{e}_1) = \text{E}(\bar{e}_0)$: all other factors must be equal on average across groups.

Selection Bias

If units *choose* their treatment status, the groups may differ in ways beyond the treatment.

Example: Hospitals and health outcomes.

- Suppose people who visit the hospital have an average health score of 3.21
- People who do not visit have an average score of 3.93

Difference: $3.21 - 3.93 = -0.72$. Does the hospital *make you sicker*?

Of course not. Sick people are more likely to go to the hospital.

$$\underbrace{\bar{e}_1}_{\text{avg. error for hospital-goers}} \neq \underbrace{\bar{e}_0}_{\text{avg. error for non-visitors}}$$

Here $\bar{e}_1$ is the average unobserved health factor for those who visit and $\bar{e}_0$ is the average for those who do not. When $E(\bar{e}_1 - \bar{e}_0) \neq 0$, the simple difference estimator is biased. This is **selection bias.**

$\implies$ Random assignment eliminates selection bias by ensuring $E(\bar{e}_1) = E(\bar{e}_0)$.

Natural Experiments

True randomized experiments (like Project STAR) are rare in economics.

A **natural experiment** (or quasi-experiment) is a real-world situation where a policy change creates treatment and control groups that are plausibly comparable.

Card and Krueger (1994):

- NJ raised its minimum wage; neighboring PA did not
- Both states share similar labor markets, demographics, and fast-food chains
- The policy change was not driven by NJ employment conditions

$\Longrightarrow$ PA serves as a natural control group for NJ.

But we still have the time-trend problem. PA employment also changed between February and November. We need a method that accounts for this.

Outline

The DiD Idea

We observe both groups in **two periods**: before and after the policy.

	Before	**After**
Control (PA)	$\bar{y}_{C,\text{Before}}$	$\bar{y}_{C,\text{After}}$
Treatment (NJ)	$\bar{y}_{T,\text{Before}}$	$\bar{y}_{T,\text{After}}$

First difference (within each group over time):

- Control change: $\bar{y}_{C,\text{After}} - \bar{y}_{C,\text{Before}}$ (time trend only)
- Treatment change: $\bar{y}_{T,\text{After}} - \bar{y}_{T,\text{Before}}$ (time trend + treatment effect)

Second difference (difference of the differences):

$$\hat{\delta} = \underbrace{(\bar{y}_{T,\text{After}} - \bar{y}_{T,\text{Before}})}_{\text{treatment change}} - \underbrace{(\bar{y}_{C,\text{After}} - \bar{y}_{C,\text{Before}})}_{\text{control change}}$$

$\Longrightarrow$ The time trend cancels, leaving only the **treatment effect.**

Card and Krueger: The 2×2 Table

Data: 410 fast-food restaurants (NJ and eastern PA), surveyed in Feb and Nov 1992.

Outcome: FTE = full-time equivalent employees per restaurant.

	Feb 1992	**Nov 1992**	**Change**
PA (Control)	23.33	21.17	-2.17
NJ (Treatment)	20.44	21.03	$+0.59$
Difference	-2.89	-0.14	

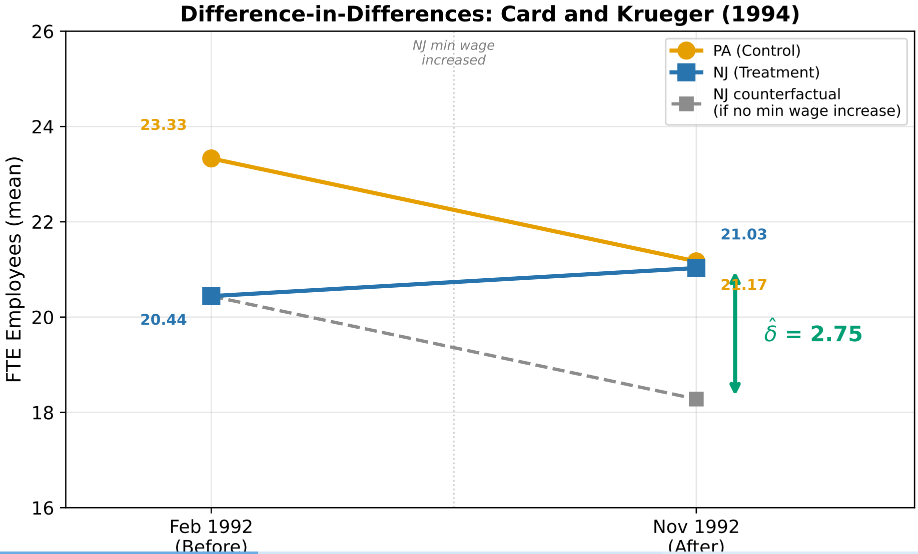
DiD estimate:

$$\hat{\delta} = (+0.59) - (-2.17) = 2.75$$

Or equivalently: $\hat{\delta} = (-0.14) - (-2.89) = 2.75$.

Employment in NJ *increased* by about 2.75 FTE relative to PA.

Visualizing DiD



Outline

From the 2×2 Table to a Regression

The DiD estimate can be obtained from a single regression:

$$y_{it} = \beta_1 + \beta_2\,\text{NJ}_i + \beta_3\,\text{After}_t + \delta\,(\text{NJ}_i \times \text{After}_t) + e_{it}$$

This uses everything you already know:

- NJ_i: an **indicator variable** (Topic 19)
- After_t: another indicator variable
- $\text{NJ}_i \times \text{After}_t$: an **interaction term** (Topic 16)

$\Longrightarrow$ DiD is not a new technique. It is a regression with two dummies and their interaction.

What Each Coefficient Means

$$y_{it} = \beta_1 + \beta_2\,\text{NJ}_i + \beta_3\,\text{After}_t + \delta\,(\text{NJ}_i \times \text{After}_t) + e_{it}$$

Write out the regression function for each cell:

	After $= 0$	After $= 1$
PA (NJ $= 0$)	β_1	$\beta_1 + \beta_3$
NJ (NJ $= 1$)	$\beta_1 + \beta_2$	$\beta_1 + \beta_2 + \beta_3 + \delta$

- $\beta_1 = 23.33$: PA mean before (reference cell)
- $\beta_2 = -2.89$: NJ was 2.89 FTE *below* PA before the policy
- $\beta_3 = -2.17$: PA employment fell by 2.17 (time trend)
- $\delta = 2.75$: the **treatment effect** (DiD estimator)

Verifying: δ Recovers the DiD Formula

From the table of regression functions:

NJ change:

$$(\beta_1 + \beta_2 + \beta_3 + \delta) - (\beta_1 + \beta_2) = \beta_3 + \delta$$

PA change:

$$(\beta_1 + \beta_3) - \beta_1 = \beta_3$$

Difference of differences:

$$(\beta_3 + \delta) - \beta_3 = \delta$$

$\implies$ The coefficient δ on the interaction term is exactly the DiD estimator.

You can test $H_0 : \delta = 0$ with a standard t-test. In Card and Krueger's data: $t \approx 1.63$, $p \approx 0.10$ (two-sided). We fail to reject at the 5% level. A two-sided test is appropriate here because the minimum wage could plausibly raise *or* lower employment.

Adding Controls

Just like any regression, we can add control variables:

$$y_{it} = \beta_1 + \beta_2\,\text{NJ}_i + \beta_3\,\text{After}_t + \delta\,(\text{NJ}_i \times \text{After}_t) + \gamma_1\text{KFC}_i + \gamma_2\text{Roys}_i + \gamma_3\text{Wendys}_i + e_{it}$$

Card and Krueger added:

- Chain dummies (KFC, Roy Rogers, Wendy's; Burger King = reference)
- Regional indicators within NJ

Result: $\hat{\delta}$ barely changes. The treatment effect estimate is robust to these controls.

$\Longrightarrow$ Adding controls can improve precision and test robustness, just as in multiple regression (Topic 13).

Outline

What Makes DiD Valid?

DiD assumes: in the absence of treatment, both groups would have experienced the **same change** over time.

This is the **parallel trends assumption** (also called **common trends**).

Formally:

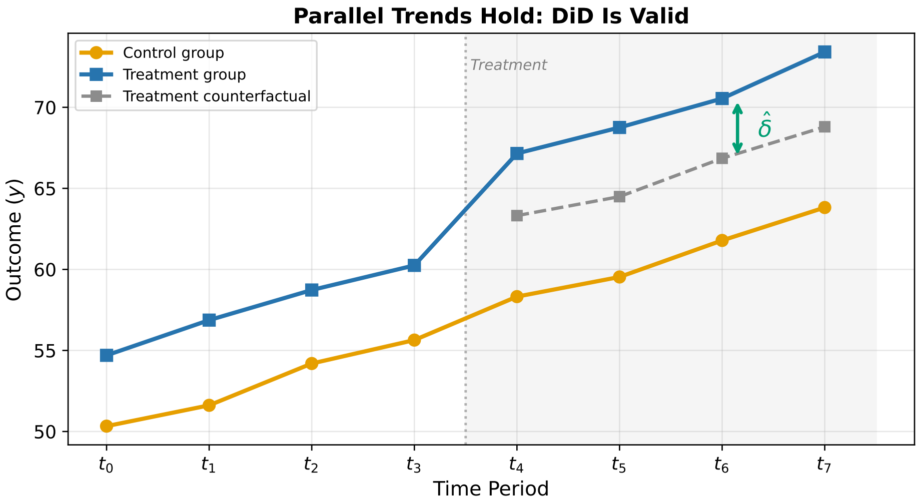
$$\mathrm{E}(y^{(0)}_{T,\text{After}}) - \mathrm{E}(y_{T,\text{Before}}) = \mathrm{E}(y_{C,\text{After}}) - \mathrm{E}(y_{C,\text{Before}})$$

where $y^{(0)}_{T,\text{After}}$ is the **counterfactual** outcome for the treatment group.

In words: the treatment group's trend (absent treatment) would have matched the control group's trend.

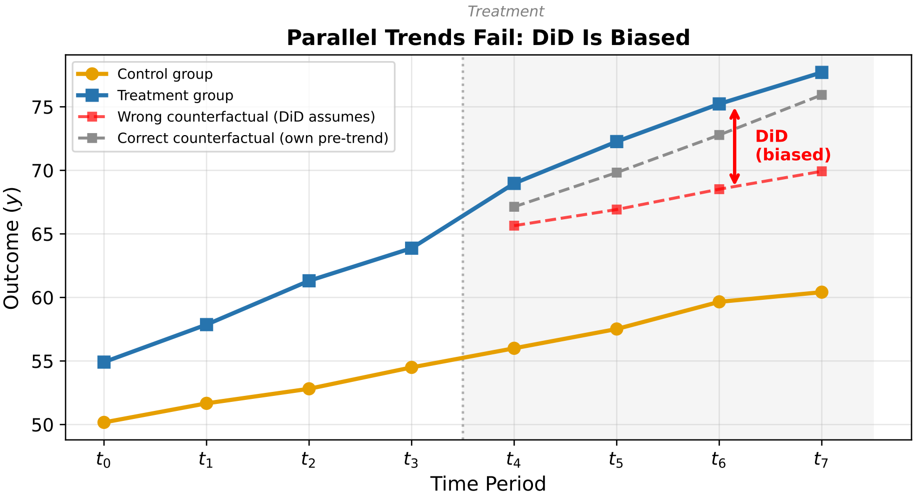
The levels can differ. The trends must be the same.

When Parallel Trends Hold



Before treatment, both groups follow the same trend (the lines are parallel). After treatment, the

When Parallel Trends Fail



The treatment group was already rising faster before the policy. DiD uses the control group's trend

Can We Test Parallel Trends?

Strictly, no. The assumption is about the *counterfactual* (what would have happened), which we never observe.

But we can check plausibility:

1. Plot pre-treatment trends for both groups. If they tracked each other before, it is more plausible they would have continued to.
2. Run a "placebo test": apply DiD to a period *before* the actual treatment. If you find an "effect," the parallel trends assumption is suspect.
3. Add controls for observable differences that might cause divergent trends.

$\Longrightarrow$ Parallel trends is fundamentally **untestable**, but strong pre-treatment evidence makes it more convincing.

In Card and Krueger's case: NJ and PA are neighboring states with similar fast-food markets. The assumption is plausible, though debated.

Outline

Using Panel Data: First Differencing

Card and Krueger observed 384 restaurants in *both* periods. With panel data, we can control for unobserved restaurant-specific characteristics:

$$\text{FTE}_{it} = \beta_1 + \beta_2\,\text{NJ}_i + \beta_3\,D_t + \delta\,(\text{NJ}_i \times D_t) + c_i + e_{it}$$

where c_i captures anything fixed about restaurant i (location quality, manager ability, etc.).

First-difference (Nov $-$ Feb) to eliminate c_i:

$$\Delta\text{FTE}_i = \beta_3 + \delta\,\text{NJ}_i + \Delta e_i$$

The restaurant fixed effect c_i drops out, along with β_1. The $\beta_2\,\text{NJ}_i$ term also disappears: NJ is time-invariant, so $\text{NJ}_{\text{Nov}} - \text{NJ}_{\text{Feb}} = 0$.

Result: $\hat{\delta} = 2.75$, confirming the earlier DiD estimate.

$\implies$ This is the same idea as the fixed effects estimator (previewed here, covered fully in Econ 104).

Outline

DiD Uses Your Entire Toolkit

Tool	Topic	Role in DiD
Indicator variables	19	NJ_i and After_t
Interaction terms	16	$\text{NJ}_i \times \text{After}_t$
Interpreting coefficients	14	What $\beta_1, \beta_2, \beta_3, \delta$ mean
Hypothesis testing	10, 17	t-test on δ; F-test for controls
OVB / controls	13, 18	Adding chain dummies, region indicators
Functional forms	12	Log-linear DiD in other applications

$\Longrightarrow$ Policy evaluation is not a separate topic. It is the application of every regression concept you have learned this quarter.

Summary

1. A **treatment effect** is the causal impact of a policy or intervention. The difference estimator $b_2 = \bar{y}_1 - \bar{y}_0$ works under random assignment.
2. **Selection bias** arises when treated and control groups differ systematically. Natural experiments provide plausible comparison groups.
3. **Difference-in-Differences** uses two groups $\times$ two periods. The DiD estimator $\hat{\delta}$ removes time trends by differencing.
4. In regression form: $y_{it} = \beta_1 + \beta_2\,\text{Treat}_i + \beta_3\,\text{After}_t + \delta(\text{Treat}_i \times \text{After}_t) + e_{it}$. The coefficient δ is the treatment effect.
5. DiD requires **parallel trends**: absent treatment, both groups would have changed the same way. Check plausibility with pre-treatment data.
6. Panel data lets you control for unit-specific fixed effects by first-differencing.

What Comes Next?

This course gave you the complete regression toolkit: estimation, inference, functional forms, indicators, interactions, and policy evaluation.

Econ 104 extends these ideas to settings where standard OLS breaks down:

- **Panel data and fixed effects**: you just saw first-differencing eliminate restaurant fixed effects; panel methods and fixed effects build on exactly this idea
- **Instrumental variables**: what to do when x is endogenous
- **Time series**: forecasting, autocorrelation, stationarity
- **Limited dependent variables**: binary outcomes (logit/probit), counts, censoring

Beyond Econ 104:

- **Regression discontinuity**: treatment assigned by a cutoff rule
- **Synthetic control**: constructing a counterfactual from weighted donors
- **Machine learning**: prediction vs. causal inference

$\Longrightarrow$ You now have the foundation for all of it.

Thank you!

Good luck on the final.

jakeanderson@g.ucla.edu