# Panel Estimation

## EC 421, Set 12

Edward Rubin

# Prologue

# Schedule

## Last time

Instrumental variables

## Today

**Panel estimation**

- what panel data buy us,
- fixed effects,
- the difference-in-differences estimator,
- *brief extension:* the synthetic-controls estimator.

# Panel estimation

# Panel estimation

## Introduction

Until now, we mostly used one of two data structures:

# Panel estimation

## Introduction

Until now, we mostly used one of two data structures:

1. **Cross section:** many units, one period,

# Panel estimation

## Introduction

Until now, we mostly used one of two data structures:

1. **Cross section:** many units, one period,
2. **Time series:** one unit, many periods.

# Panel estimation

## Introduction

Until now, we mostly used one of two data structures:

1. **Cross section:** many units, one period,
2. **Time series:** one unit, many periods.

A **panel** gives us both:

- units indexed by $i$,
- (time) periods indexed by $t$,
- outcomes and regressors like $y_{it}$ and $x_{it}$.

# Panel estimation

## Example

Consider the following panel of six observations:

```
#>      Unit  Year  y_it  x_it
#>    <char> <num> <num> <num>
#> 1:      A     1    10     2
#> 2:      A     2    11     3
#> 3:      A     3    13     4
#> 4:      B     1     7     1
#> 5:      B     2     8     1
#> 6:      B     3     9     2
```

# Panel estimation

## Example

Consider the following panel of six observations:

```
#>       Unit  Year  y_it  x_it
#>     <char> <num> <num> <num>
#> 1:       A     1    10     2
#> 2:       A     2    11     3
#> 3:       A     3    13     4
#> 4:       B     1     7     1
#> 5:       B     2     8     1
#> 6:       B     3     9     2
```

We observe the same unit repeatedly.

We can compare unit A to itself **over time**, not just A to B (potentially helping with *selection bias*).

# Panel estimation

## Example

Consider the following panel of six observations:

```
#>      Unit  Year  y_it  x_it
#>    <char> <num> <num> <num>
#> 1:      A     1    10     2
#> 2:      A     2    11     3
#> 3:      A     3    13     4
#> 4:      B     1     7     1
#> 5:      B     2     8     1
#> 6:      B     3     9     2
```

Similarly, we observe the period across units.

This feature lets us control for shocks shared across units at a given time.

## The benefits of panels

Panels help solve (some) econometric problems because they allow us to

# Panel estimation

## The benefits of panels

Panels help solve (some) econometric problems because they allow us to

- control for **time-invariant differences** across units
  *removes selection bias from cross-sectional differences across units*

# Panel estimation

## The benefits of panels

Panels help solve (some) econometric problems because they allow us to

- control for **time-invariant differences** across units

  *removes selection bias from cross-sectional differences across units*

- implicitly study **changes**

  *not just levels*

## The benefits of panels

Panels help solve (some) econometric problems because they allow us to

- control for **time-invariant differences** across units
  *removes selection bias from cross-sectional differences across units*

- implicitly study **changes**
  *not just levels*

- *related:* evaluate policies using their **timing**
  *policy-timing is often more plausibly exogeneous than policy existence*

# Panel estimation

## The benefits of panels

Panels help solve (some) econometric problems because they allow us to

- control for **time-invariant differences** across units
  *removes selection bias from cross-sectional differences across units*

- implicitly study **changes**
  *not just levels*

- *related:* evaluate policies using their **timing**
  *policy-timing is often more plausibly exogeneous than policy existence*

Panel methods are workhorses of economics and public policy.

# Panel estimation

## Panels are helpful—not magical

Panel data *can* be helpful for estimating causal effects, but they can still cause problems without careful use.

## Panels are helpful—not magical

Panel data *can* be helpful for estimating causal effects,
but they can still cause problems without careful use.

- We need (sufficient) *within variation* in $x_{it}$.
  *I.e.*, $x_{it}$ must change meaningfully through time for some units.

- Time-varying omitted variables can cause bias.

- A unit's disturbances $u_{it}$ are often correlated over time.

- Attrition and missing periods can matter—as can measurement error.

*(Any tool can be misused.)*

# Panel estimation

## The problem with pooled OLS

Suppose you're interested in a simple linear regression model

$$y_{it} = \beta_0 + \beta_1 x_{it} + u_{it}$$

and have **panel data** (with individual $i$ and time $t$).

## The problem with pooled OLS

Suppose you're interested in a simple linear regression model

$$y_{it} = \beta_0 + \beta_1 x_{it} + u_{it}$$

and have **panel data** (with individual $i$ and time $t$).

Let the disturbance combine (1) unit-specific factor $\alpha_i$ and (2) error $v_{it}$

$$u_{it} = \alpha_i + v_{it}$$

# Panel estimation

## The problem with pooled OLS

Suppose you're interested in a simple linear regression model

$$y_{it} = \beta_0 + \beta_1 x_{it} + u_{it}$$

and have **panel data** (with individual $i$ and time $t$).

Let the disturbance combine (1) unit-specific factor $\alpha_i$ and (2) error $v_{it}$

$$u_{it} = \alpha_i + v_{it}$$

If $\mathbf{Cov}(x_{it}, \alpha_i) \neq \mathbf{0}$, then OLS may be biased. *(Think OVB.)*

## The problem with pooled OLS

Suppose you're interested in a simple linear regression model

$$y_{it} = \beta_0 + \beta_1 x_{it} + u_{it}$$

and have **panel data** (with individual $i$ and time $t$).

Let the disturbance combine (1) unit-specific factor $\alpha_i$ and (2) error $v_{it}$

$$u_{it} = \alpha_i + v_{it}$$

If $\mathrm{Cov}(x_{it}, \alpha_i) \neq \mathbf{0}$, then OLS may be biased. *(Think OVB.)*

Standard OLS estimates of $\beta_1$ (*i.e.*, without FEs) will reflect both

- *between-unit differences* $\alpha_i$,
- *actual changes* in $x_{it}$. *(We want to isolate this last part.)*

## Introducing *fixed effects*

*The solution:* Let each unit have its own intercept $\alpha_i$,
and each time period have its own intercept $\lambda_t$:

$$y_{it} = \beta_1 x_{it} + \alpha_i + \lambda_t + u_{it}$$

These unit- and period-specific intercepts are called **fixed effects**.

## Introducing *fixed effects*

*The solution:* Let each unit have its own intercept $\alpha_i$,
and each time period have its own intercept $\lambda_t$:

$$y_{it} = \beta_1 x_{it} + \alpha_i + \lambda_t + u_{it}$$

These unit- and period-specific intercepts are called **fixed effects**.

```
#>      Unit  Year  y_it  x_it Unit_A Unit_B Year_1 Year_2 Year_3
#>    <char> <num> <num> <num>  <num>  <num>  <num>  <num>  <num>
#> 1:      A     1    10     2      1      0      1      0      0
#> 2:      A     2    11     3      1      0      0      1      0
#> 3:      A     3    13     4      1      0      0      0      1
#> 4:      B     1     7     1      0      1      1      0      0
#> 5:      B     2     8     1      0      1      0      1      0
#> 6:      B     3     9     2      0      1      0      0      1
```

## Fixed effects in action

**Q** So what are these fixed effects (FE) doing?

## Fixed effects in action

**Q** So what are these fixed effects (FE) doing?

Remember the old interpretation of coefficients on binary variables?

## Fixed effects in action

**Q** So what are these fixed effects (FE) doing?

Remember the old interpretation of coefficients on binary variables?

**A** These fixed effects *absorb*

- $\alpha_i$: all time-invariant differences across units,
  *(including selection bias from cross-sectional differences)*
- $\lambda_t$: all common shocks at a given time.

## Fixed effects in action

**Q** So what are these fixed effects (FE) doing?

Remember the old interpretation of coefficients on binary variables?

**A** These fixed effects *absorb*

- $\alpha_i$: all time-invariant differences across units,
  *(including selection bias from cross-sectional differences)*
- $\lambda_t$: all common shocks at a given time.

The **(two-way) fixed-effects estimator** (OLS that includes unit and time FEs) estimates effects using *within-unit variation*—net of common time shocks.

# Panel estimation

## The *within* transformation

For intuition, focus on unit-level fixed effects.

$$y_{it} = \beta_0 + \beta_1 x_{it} + \alpha_i + u_{it}$$

# Panel estimation

## The *within* transformation

For intuition, focus on unit-level fixed effects.

$$y_{it} = \beta_0 + \beta_1 x_{it} + \alpha_i + u_{it}$$

Take the average of this equation for unit $i$ across all periods…

$$\overline{y}_i = \beta_0 + \beta_1 \overline{x}_i + \alpha_i + \overline{u}_i$$

# Panel estimation

## The *within* transformation

For intuition, focus on unit-level fixed effects.

$$y_{it} = \beta_0 + \beta_1 x_{it} + \alpha_i + u_{it}$$

Take the average of this equation for unit $i$ across all periods...

$$\overline{y}_i = \beta_0 + \beta_1 \overline{x}_i + \alpha_i + \overline{u}_i$$

Finally, subtract the unit mean from both sides of the original equation.
(This step is what our unit-level fixed effects are doing.)

$$y_{it} - \overline{y}_i = \beta_1 \left( x_{it} - \overline{x}_i \right) + \left( u_{it} - \overline{u}_i \right)$$

# Panel estimation

## The *within* transformation

For intuition, focus on unit-level fixed effects.

$$y_{it} = \beta_0 + \beta_1 x_{it} + \alpha_i + u_{it}$$

Take the average of this equation for unit $i$ across all periods...

$$\overline{y}_i = \beta_0 + \beta_1 \overline{x}_i + \alpha_i + \overline{u}_i$$

Finally, subtract the unit mean from both sides of the original equation.
(This step is what our unit-level fixed effects are doing.)

$$y_{it} - \overline{y}_i = \beta_1 \left( x_{it} - \overline{x}_i \right) + \left( u_{it} - \overline{u}_i \right)$$

**So what?** Unit-specific FEs identify $\beta_1$ using within-unit variation in $x_{it}$.

*(The same logic applies to time FEs.)*

## Fixed-effects questions

Fixed effects **do not** ask questions like

*Why is unit A different from unit B?*

*In fact, they try to absorb cross-unit differences.*

# Panel estimation

## Fixed-effects questions

Fixed effects **do not** ask questions like

> *Why is unit A different from unit B?*

*In fact, they try to absorb cross-unit differences.*

Fixed effects also **do not** ask questions like

> *When period 1 had a shock, how did all units' outcomes change?*

*They try to absorb common (shared) time shocks.*

## Fixed-effects questions

Fixed effects **do not** ask questions like

*Why is unit A different from unit B?*

*In fact, they try to absorb cross-unit differences.*

Fixed effects also **do not** ask questions like

*When period 1 had a shock, how did all units' outcomes change?*

*They try to absorb common (shared) time shocks.*

Fixed effects **do** ask questions like

*When unit A changed its value of* $x_{it}$*, how did its* $y_{it}$ *change?*

*This is the "within" variation.*

# Panel estimation

## Example: State minimum wage laws

Suppose we want to estimate the effect of min. wage laws on employment

$$\text{Employment}_{it} = \beta_1(\text{Min. Wage})_{it} + \alpha_i + \lambda_t + u_{it}$$

with a state ($i$) by year ($t$) panel of employment and min. wage data.

# Panel estimation

## Example: State minimum wage laws

Suppose we want to estimate the effect of min. wage laws on employment

$$\text{Employment}_{it} = \beta_1 (\text{Min. Wage})_{it} + \alpha_i + \lambda_t + u_{it}$$

with a state ($i$) by year ($t$) panel of employment and min. wage data.

- $\alpha_i$ absorbs time-invariant, cross-state differences;
- $\lambda_t$ absorbs national shocks in year $t$.

# Panel estimation

## Example: State minimum wage laws

Suppose we want to estimate the effect of min. wage laws on employment

$$\textbf{Employment}_{it} = \beta_1 (\textbf{Min. Wage})_{it} + \alpha_i + \lambda_t + u_{it}$$

with a state ($i$) by year ($t$) panel of employment and min. wage data.

- $\alpha_i$ absorbs time-invariant, cross-state differences;
- $\lambda_t$ absorbs national shocks in year $t$.

Q Should we worry that some states have higher min. wage than others?

# Panel estimation

## Example: State minimum wage laws

Suppose we want to estimate the effect of min. wage laws on employment

$$\text{Employment}_{it} = \beta_1 (\text{Min. Wage})_{it} + \alpha_i + \lambda_t + u_{it}$$

with a state ($i$) by year ($t$) panel of employment and min. wage data.

- $\alpha_i$ absorbs time-invariant, cross-state differences;
- $\lambda_t$ absorbs national shocks in year $t$.

Q Should we worry that some states have higher min. wage than others?
A No. $\alpha_i$ absorbs those differences. The FE estimator estimates using within-state variation in min. wage.

# Panel estimation

## Example: State minimum wage laws

Suppose we want to estimate the effect of min. wage laws on employment

$$\text{Employment}_{it} = \beta_1 (\text{Min. Wage})_{it} + \alpha_i + \lambda_t + u_{it}$$

with a state ($i$) by year ($t$) panel of employment and min. wage data.

- $\alpha_i$ absorbs time-invariant, cross-state differences;
- $\lambda_t$ absorbs national shocks in year $t$.

Q What if some *years* have lower emp. than others (recessions)?

# Panel estimation

## Example: State minimum wage laws

Suppose we want to estimate the effect of min. wage laws on employment

$$\text{Employment}_{it} = \beta_1 (\text{Min. Wage})_{it} + \alpha_i + \lambda_t + u_{it}$$

with a state ($i$) by year ($t$) panel of employment and min. wage data.

- $\alpha_i$ absorbs time-invariant, cross-state differences;
- $\lambda_t$ absorbs national shocks in year $t$.

**Q** What if some *years* have lower emp. than others (recessions)?

**A** Still not worried! $\lambda_t$ absorbs *shared* shocks.

# Panel estimation

## Example: State minimum wage laws

Suppose we want to estimate the effect of min. wage laws on employment

$$\text{Employment}_{it} = \beta_1 (\text{Min. Wage})_{it} + \alpha_i + \lambda_t + u_{it}$$

with a state ($i$) by year ($t$) panel of employment and min. wage data.

- $\alpha_i$ absorbs time-invariant, cross-state differences;
- $\lambda_t$ absorbs national shocks in year $t$.

Q So what should we worry about?

# Panel estimation

## Example: State minimum wage laws

Suppose we want to estimate the effect of min. wage laws on employment

$$\text{Employment}_{it} = \beta_1 (\text{Min. Wage})_{it} + \alpha_i + \lambda_t + u_{it}$$

with a state ($i$) by year ($t$) panel of employment and min. wage data.

- $\alpha_i$ absorbs time-invariant, cross-state differences;
- $\lambda_t$ absorbs national shocks in year $t$.

**Q** So what should we worry about?

**A** We should worry about **time-varying differences between states**.

*E.g.*, if some states have faster-growing economies and also raise their minimum wages.

## What FE means in that example

In the minimum-wage panel, fixed effects (FEs) net out factors like

- state geography,
- long-run industry mix,
- baseline labor-market conditions,
- permanent political differences.

## What FE means in that example

In the minimum-wage panel, fixed effects (FEs) net out factors like

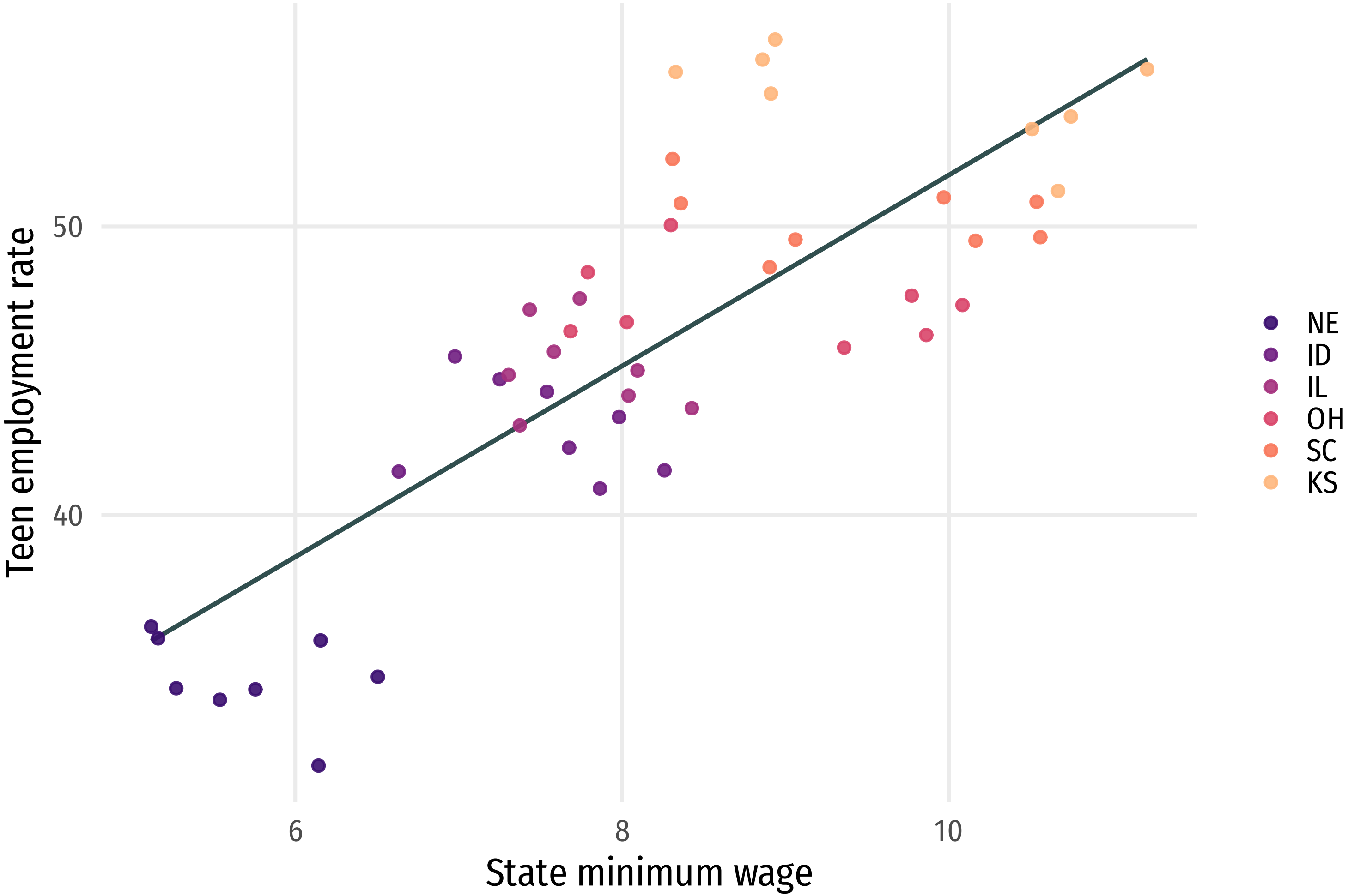
- state geography,
- long-run industry mix,
- baseline labor-market conditions,
- permanent political differences.

$$\mathbf{Employment}_{it} = \beta_1 (\mathbf{Min.\ Wage})_{it} + \alpha_i + \lambda_t + u_{it}$$

With FEs, $\hat{\beta}_1$ comes from **states that change their minimum wage**, relative to their own histories and common national trends.

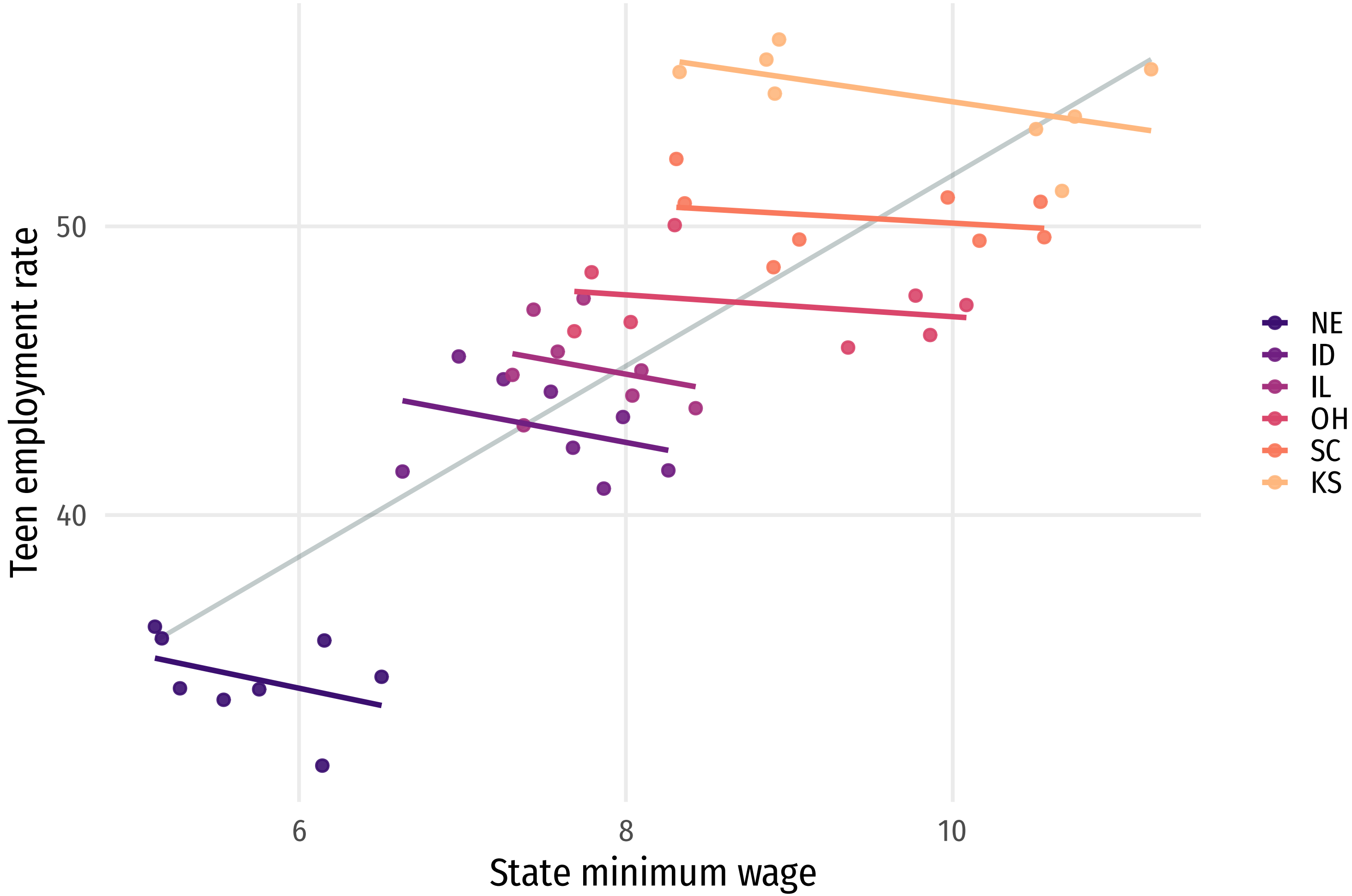
Let's see a graphical example.

(Simulated data but real FE issues.)



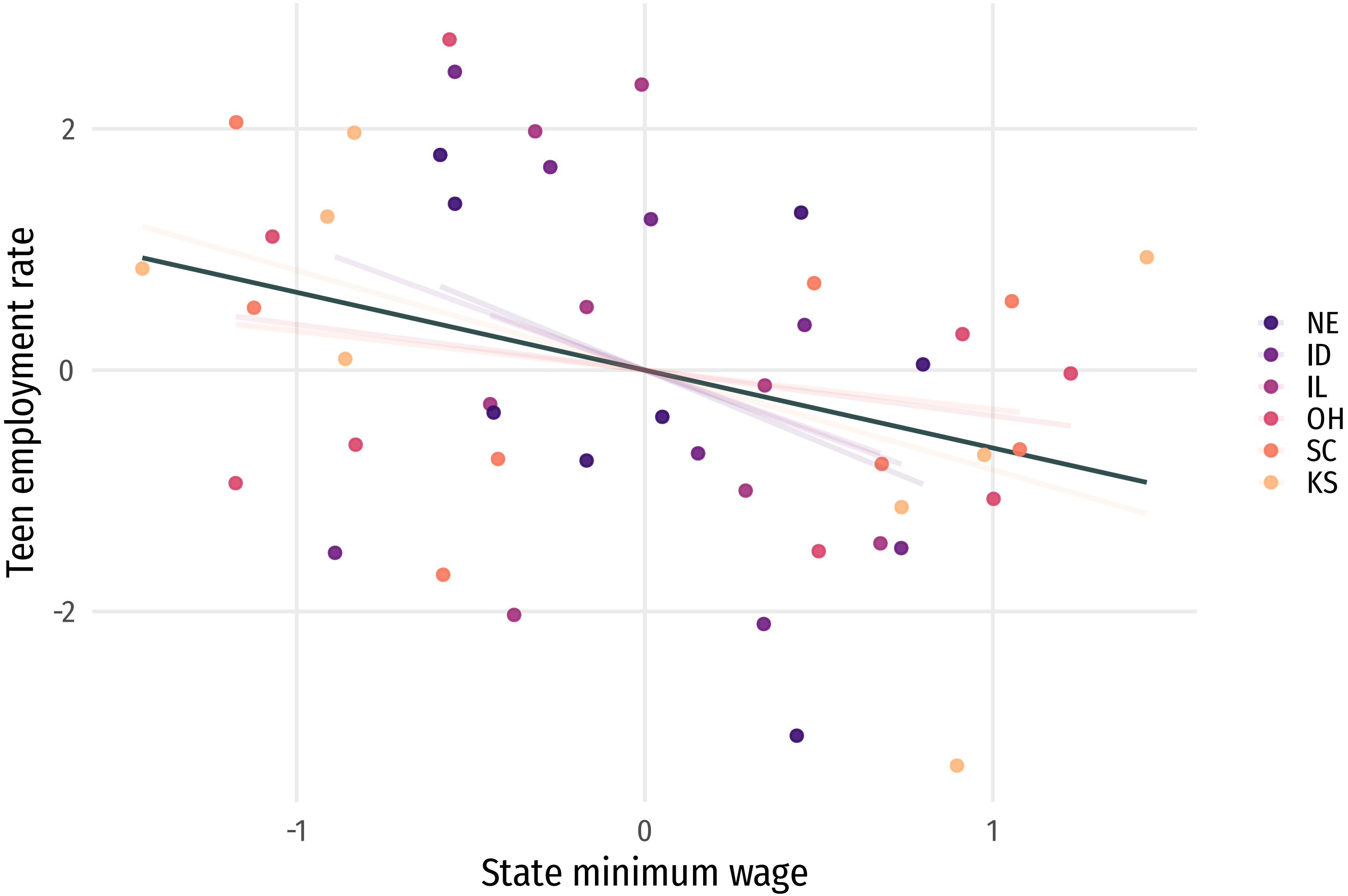

The dark line is **pooled OLS**—a regression without fixed effects.

Here, states with higher min. wage laws have higher employment rates.



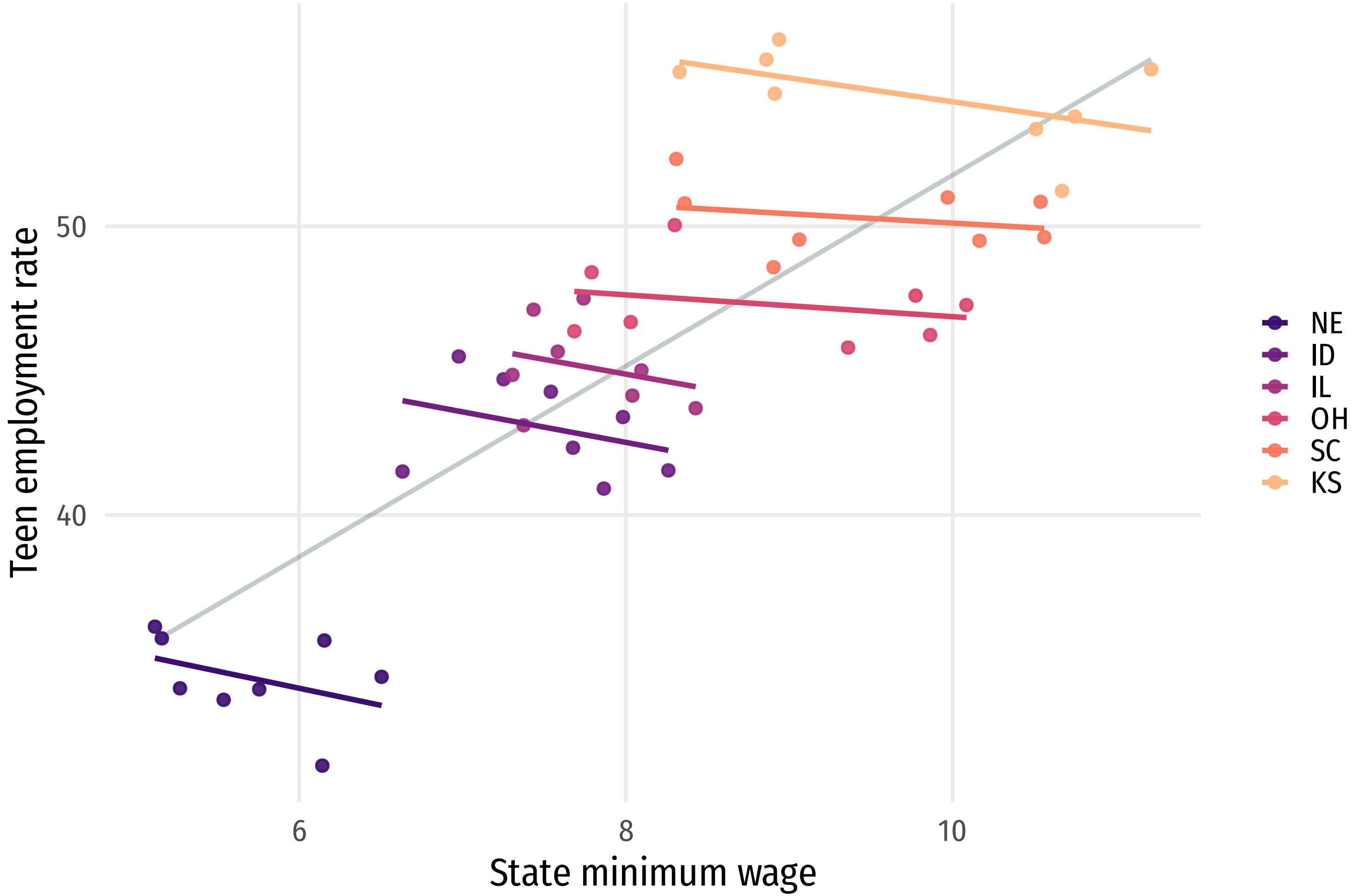

However, the **within-state relationships** tell another story.

The FE estimator focuses on this var.—*absorbing* cross-state differences.




Integrating **fixed effects**, the relationship flips.

States that raise their minimum wage see lower teen employment.




This *flipped* relationship is an example of **Simpson's paradox**—when a relationship reverses after controlling for *"group"*.

# Panel estimation

## `fixest` syntax

Now you know the intution, logic, and mechanics of fixed effects.

You also now understand the `fix` part of the `fixest` package.

To specify FEs: use `lm` syntax along with `|` followed by the FEs.

```
feols(
  teen_emp ~ min_wage | state + year,
  cluster = ~ state,
  data = mw_df
)
```

Here we have fixed effects for state and year.

**Output** `feols` output for the *plain* OLS model *vs.* the FE model.

```
#>                            est_ols             est_fe
#> Dependent Var.:           teen_emp           teen_emp
#>
#> Constant            21.3*** (3.00)
#> min_wage           2.93*** (0.345) -1.44*** (0.196)
#> Fixed-Effects:     --------------- ----------------
#> state                           No                Yes
#> year                            No                Yes
#> _______________    _______________ ________________
#> S.E.: Clustered          by: state          by: state
#> Observations                   320                320
#> R2                         0.48037            0.96360
#> Within R2                       --            0.13134
#> ---
#> Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

**Output** `feols` output for the *plain* OLS model *vs.* the FE model.

```
#>                          est_ols           est_fe
#> Dependent Var.:         teen_emp         teen_emp
#>
#> Constant           21.3*** (3.00)
#> min_wage          2.93*** (0.345) -1.44*** (0.196)
#> Fixed-Effects:   ---------------- ----------------
#> state                         No              Yes
#> year                          No              Yes
#> _______________  ________________ ________________
#> S.E.: Clustered        by: state        by: state
#> Observations                 320              320
#> R2                       0.48037          0.96360
#> Within R2                     --          0.13134
#> ---
#> Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Both $\hat{\beta}_1$ are stat. significant, but they tell very different stories.

**Output** `feols` output for the *plain* OLS model *vs.* the FE model.

```
#>                         est_ols            est_fe
#> Dependent Var.:        teen_emp          teen_emp
#>
#> Constant        21.3*** (3.00)
#> min_wage       2.93*** (0.345) -1.44*** (0.196)
#> Fixed-Effects: --------------- ----------------
#> state                       No               Yes
#> year                        No               Yes
#> _______________ _______________ ________________
#> S.E.: Clustered       by: state        by: state
#> Observations                320              320
#> R2                      0.48037          0.96360
#> Within R2                    --          0.13134
#> ---
#> Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

**Q** Why is there no intercept ($\beta_0$) in the FE model?

**Output** `feols` output for the *plain* OLS model *vs.* the FE model.

```
#>                              est_ols            est_fe
#> Dependent Var.:             teen_emp          teen_emp
#>
#> Constant            21.3*** (3.00)
#> min_wage           2.93*** (0.345) -1.44*** (0.196)
#> Fixed-Effects:     --------------- ----------------
#> state                           No               Yes
#> year                            No               Yes
#> _______________    _______________ ________________
#> S.E.: Clustered          by: state         by: state
#> Observations                   320               320
#> R2                         0.48037           0.96360
#> Within R2                       --           0.13134
#> ---
#> Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

**Q** Why is there no intercept ($\beta_0$) in the FE model?

**A** The fixed effects are collinear with the intercept—they each provide a separate intercept for each unit and time period.

**Output** `feols` output for the *plain* OLS model *vs.* the FE model.

```
#>                               est_ols            est_fe
#> Dependent Var.:              teen_emp          teen_emp
#>
#> Constant           21.3*** (3.00)
#> min_wage          2.93*** (0.345) -1.44*** (0.196)
#> Fixed-Effects:    --------------- ----------------
#> state                          No               Yes
#> year                           No               Yes
#> _______________   _______________ ________________
#> S.E.: Clustered         by: state        by: state
#> Observations                  320              320
#> R2                        0.48037          0.96360
#> Within R2                      --          0.13134
#> ---
#> Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Q What is ***within*** **$R^2$**?

**Output** `feols` output for the *plain* OLS model *vs.* the FE model.

```
#>                          est_ols           est_fe
#> Dependent Var.:         teen_emp         teen_emp
#>
#> Constant             21.3*** (3.00)
#> min_wage           2.93*** (0.345) -1.44*** (0.196)
#> Fixed-Effects:     --------------- ----------------
#> state                           No              Yes
#> year                            No              Yes
#> _______________    _______________ ________________
#> S.E.: Clustered          by: state        by: state
#> Observations                   320              320
#> R2                         0.48037          0.96360
#> Within R2                       --          0.13134
#> ---
#> Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Q What is ***within* R²**?

A It is the $R^2$ of the regression after accouting for the FEs: how much of the outcome's variation regressors explain, *after controlling for FEs*.

**Output** `feols` output for the *plain* OLS model *vs.* the FE model.

```
#>                               est_ols            est_fe
#> Dependent Var.:              teen_emp          teen_emp
#>
#> Constant           21.3*** (3.00)
#> min_wage          2.93*** (0.345) -1.44*** (0.196)
#> Fixed-Effects:    --------------- ----------------
#> state                          No               Yes
#> year                           No               Yes
#> _______________   _______________ ________________
#> S.E.: Clustered         by: state        by: state
#> Observations                  320              320
#> R2                        0.48037          0.96360
#> Within R2                      --          0.13134
#> ---
#> Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The FE model has as high $R^2$ because fixed effects explain a lot of variation (in essence, means of states and years).

# Panel estimation

## Summary

Fixed effects help one key problem: **time-invariant omitted variables**.

# Panel estimation

## Summary

Fixed effects help one key problem: **time-invariant omitted variables**.

But FEs **do not** solve *all* causal problems, *e.g.*,

- time-varying confounders,
- reverse causality
- measurement error,
- poor measurement of treatment timing,
- weak within variation.

As with *all* econometric tools, careful use is required.

# Difference-in-differences

# Difference-in-differences

## Core idea

**Difference-in-differences** (DiD)[†] is a popular panel-data approach for estimating causal effects of policies/events.

DiD compares observations across two dimensions:

1. **Group:** comparing
   - *treated* individuals (potentially affected by the policy)
   - *control* individuals (unaffected by the policy)
2. **Time:** comparing
   - *before* the policy takes effect *(pre-treatment)*
   - *after* the policy takes effect *(post-treatment)*

† Also called *double-differences* or *diff-in-diff*. Increasingly combined with *two-way fixed effects (TWFE)*.

# Difference-in-differences

## Intuition

DiD functions by taking the difference between two differences:

# Difference-in-differences

## Intuition

DiD functions by taking the difference between two differences:

1. $\Delta\bar{\mathbf{Y}}_{\textbf{Pre}} = \bar{\mathbf{Y}}(\text{Trt}; \text{Pre}) - \bar{\mathbf{Y}}(\text{Ctrl}; \text{Pre})$
   the average difference between trt. and control **before** the policy, accounting for pre-existing differences between the groups;

# Difference-in-differences

## Intuition

DiD functions by taking the difference between two differences:

1. $\Delta\bar{\mathbf{Y}}_{\textbf{Pre}} = \bar{\mathbf{Y}}(\text{Trt}; \text{Pre}) - \bar{\mathbf{Y}}(\text{Ctrl}; \text{Pre})$
   the average difference between trt. and control **before** the policy, accounting for pre-existing differences between the groups;

2. $\Delta\bar{\mathbf{Y}}_{\textbf{Post}} = \bar{\mathbf{Y}}(\text{Trt}; \text{Post}) - \bar{\mathbf{Y}}(\text{Ctrl}; \text{Post})$
   the average difference between trt. and control **after** the policy, which includes pre-existing differences *and* the treatment effect.

## Intuition

DiD functions by taking the difference between two differences:

1. $\Delta\bar{\mathbf{Y}}_{\text{Pre}} = \bar{\mathbf{Y}}(\text{Trt; Pre}) - \bar{\mathbf{Y}}(\text{Ctrl; Pre})$
   the average difference between trt. and control **before** the policy, accounting for pre-existing differences between the groups;

2. $\Delta\bar{\mathbf{Y}}_{\text{Post}} = \bar{\mathbf{Y}}(\text{Trt; Post}) - \bar{\mathbf{Y}}(\text{Ctrl; Post})$
   the average difference between trt. and control **after** the policy, which includes pre-existing differences *and* the treatment effect.

The DiD estimate is the difference between these differences

$$\hat{\tau} = \Delta\bar{\mathbf{Y}}_{\text{Post}} - \Delta\bar{\mathbf{Y}}_{\text{Pre}}$$

*I.e.*, how much did the trt-control difference change from pre to post?

**Ex.** Two groups (trt/ctrl) and two time periods (pre/post policy).

```
#>      Group Period  Mean
#>     <char> <char> <num>
#> 1: Treated    Pre 1.965
#> 2: Control    Pre 0.323
#> 3: Treated   Post 6.172
#> 4: Control   Post 1.036
```

**Ex.** Two groups (trt/ctrl) and two time periods (pre/post policy).

```
#>       Group Period  Mean
#>      <char> <char> <num>
#> 1: Treated    Pre 1.965
#> 2: Control    Pre 0.323
#> 3: Treated   Post 6.172
#> 4: Control   Post 1.036
```

1. $\Delta \bar{\mathbf{Y}}_{\text{Pre}} = \bar{\mathbf{Y}}(\text{Trt}; \text{Pre}) - \bar{\mathbf{Y}}(\text{Ctrl}; \text{Pre}) = 1.642.$

**Ex.** Two groups (trt/ctrl) and two time periods (pre/post policy).

```
#>       Group Period  Mean
#>      <char> <char> <num>
#> 1: Treated    Pre 1.965
#> 2: Control    Pre 0.323
#> 3: Treated   Post 6.172
#> 4: Control   Post 1.036
```

1. $\Delta \bar{\mathbf{Y}}_{\text{Pre}} = \bar{\mathbf{Y}}(\text{Trt}; \text{Pre}) - \bar{\mathbf{Y}}(\text{Ctrl}; \text{Pre}) = 1.642.$

2. $\Delta \bar{\mathbf{Y}}_{\text{Post}} = \bar{\mathbf{Y}}(\text{Trt}; \text{Post}) - \bar{\mathbf{Y}}(\text{Ctrl}; \text{Post}) = 5.136$

**Ex.** Two groups (trt/ctrl) and two time periods (pre/post policy).

```
#>       Group Period  Mean
#>      <char> <char> <num>
#> 1: Treated    Pre 1.965
#> 2: Control    Pre 0.323
#> 3: Treated   Post 6.172
#> 4: Control   Post 1.036
```

1. $\Delta\bar{\mathbf{Y}}_{\text{Pre}} = \bar{\mathbf{Y}}(\text{Trt}; \text{Pre}) - \bar{\mathbf{Y}}(\text{Ctrl}; \text{Pre}) = 1.642.$

2. $\Delta\bar{\mathbf{Y}}_{\text{Post}} = \bar{\mathbf{Y}}(\text{Trt}; \text{Post}) - \bar{\mathbf{Y}}(\text{Ctrl}; \text{Post}) = 5.136$

The DiD estimate—the difference between these differences

$$\hat{\tau} = \Delta\bar{\mathbf{Y}}_{\text{Post}} - \Delta\bar{\mathbf{Y}}_{\text{Pre}} = 3.494$$

*Note* You can also write the DiD estimator as two other differences,

$$\hat{\tau} = \Delta\bar{\mathbf{Y}}_{\text{Post}} - \Delta\bar{\mathbf{Y}}_{\text{Pre}}$$

*Note* You can also write the DiD estimator as two other differences,

$$\hat{\tau} = \Delta\bar{\mathbf{Y}}_{\textbf{Post}} - \Delta\bar{\mathbf{Y}}_{\textbf{Pre}}$$

$$= \left[\ \bar{\mathbf{Y}}(\text{Trt}; \text{Post}) - \bar{\mathbf{Y}}(\text{Ctrl}; \text{Post})\ \right] - \left[\ \bar{\mathbf{Y}}(\text{Trt}; \text{Pre}) - \bar{\mathbf{Y}}(\text{Ctrl}; \text{Pre})\ \right]$$

*Note* You can also write the DiD estimator as two other differences,

$$\hat{\tau} = \Delta\bar{\mathbf{Y}}_{\text{Post}} - \Delta\bar{\mathbf{Y}}_{\text{Pre}}$$

$$= \left[\ \bar{\mathbf{Y}}(\text{Trt};\ \text{Post}) - \bar{\mathbf{Y}}(\text{Ctrl};\ \text{Post})\ \right] - \left[\ \bar{\mathbf{Y}}(\text{Trt};\ \text{Pre}) - \bar{\mathbf{Y}}(\text{Ctrl};\ \text{Pre})\ \right]$$

$$= \bar{\mathbf{Y}}(\text{Trt};\ \text{Post}) - \bar{\mathbf{Y}}(\text{Ctrl};\ \text{Post}) - \bar{\mathbf{Y}}(\text{Trt};\ \text{Pre}) + \bar{\mathbf{Y}}(\text{Ctrl};\ \text{Pre})$$

*Note* You can also write the DiD estimator as two other differences,

$$
\begin{aligned}
\hat{\tau} &= \Delta\bar{\mathbf{Y}}_{\textbf{Post}} - \Delta\bar{\mathbf{Y}}_{\textbf{Pre}} \\
&= \left[\, \bar{\mathbf{Y}}(\text{Trt}; \text{Post}) - \bar{\mathbf{Y}}(\text{Ctrl}; \text{Post}) \,\right] - \left[\, \bar{\mathbf{Y}}(\text{Trt}; \text{Pre}) - \bar{\mathbf{Y}}(\text{Ctrl}; \text{Pre}) \,\right] \\
&= \bar{\mathbf{Y}}(\text{Trt}; \text{Post}) - \bar{\mathbf{Y}}(\text{Ctrl}; \text{Post}) - \bar{\mathbf{Y}}(\text{Trt}; \text{Pre}) + \bar{\mathbf{Y}}(\text{Ctrl}; \text{Pre}) \\
&= \left[\, \bar{\mathbf{Y}}(\text{Trt}; \text{Post}) - \bar{\mathbf{Y}}(\text{Trt}; \text{Pre}) \,\right] - \left[\, \bar{\mathbf{Y}}(\text{Ctrl}; \text{Post}) - \bar{\mathbf{Y}}(\text{Ctrl}; \text{Pre}) \,\right]
\end{aligned}
$$

*Note* You can also write the DiD estimator as two other differences,

$$
\begin{aligned}
\hat{\tau} &= \Delta\bar{\mathbf{Y}}_{\textbf{Post}} - \Delta\bar{\mathbf{Y}}_{\textbf{Pre}} \\
&= \left[\ \bar{\mathbf{Y}}(\text{Trt}; \text{Post}) - \bar{\mathbf{Y}}(\text{Ctrl}; \text{Post})\ \right] - \left[\ \bar{\mathbf{Y}}(\text{Trt}; \text{Pre}) - \bar{\mathbf{Y}}(\text{Ctrl}; \text{Pre})\ \right] \\
&= \bar{\mathbf{Y}}(\text{Trt}; \text{Post}) - \bar{\mathbf{Y}}(\text{Ctrl}; \text{Post}) - \bar{\mathbf{Y}}(\text{Trt}; \text{Pre}) + \bar{\mathbf{Y}}(\text{Ctrl}; \text{Pre}) \\
&= \left[\ \bar{\mathbf{Y}}(\text{Trt}; \text{Post}) - \bar{\mathbf{Y}}(\text{Trt}; \text{Pre})\ \right] - \left[\ \bar{\mathbf{Y}}(\text{Ctrl}; \text{Post}) - \bar{\mathbf{Y}}(\text{Ctrl}; \text{Pre})\ \right] \\
&= \Delta\bar{\mathbf{Y}}_{\text{Trt}} - \Delta\bar{\mathbf{Y}}_{\text{Ctrl}}
\end{aligned}
$$

*I.e.,* the change in the treated group (post − pre) relative to the change in the control group.

We're using the control group's **change** as a **counterfactual** for the treated group's **change** had it not received treatment.

# Difference-in-differences

## Classic example: Card and Krueger (1994)

One famous DiD application estimates the **impact of increasing the minimum wage on fast-food employment**.

- *Who:* New Jersey (trt) *vs.* Pennsylvania (ctrl)
- *When:* before 01 April 1992 (pre) *vs.* after 01 April 1992 (post)

DiD estimates *(both questions = same answer)*

1. How did NJ-PA empl. diff. change after the policy, relative to before?
2. How did NJ's empl. change after the policy relative to PA's change?

# Difference-in-differences

## Classic example: Card and Krueger (1994)

One famous DiD application estimates the **impact of increasing the minimum wage on fast-food employment**.

- *Who:* New Jersey (trt) *vs.* Pennsylvania (ctrl)
- *When:* before 01 April 1992 (pre) *vs.* after 01 April 1992 (post)

DiD estimates *(both questions = same answer)*

1. How did NJ-PA empl. diff. change after the policy, relative to before?
2. How did NJ's empl. change after the policy relative to PA's change?

*Key assumption:* PA needs to be a good counterfactual for NJ. In the absence of the policy, NJ and PA would have followed the same empl. trends.

## DiD as a regression

With two groups and two periods

$$y_{it} = \beta_0 + \beta_1 \mathbf{Trt}_i + \beta_2 \mathbf{Post}_t + \delta \left( \mathbf{Trt}_i \times \mathbf{Post}_t \right) + u_{it}$$

where $\mathbf{Trt}_i$ and $\mathbf{Post}_t$ are binary indicators for the treated group and post-treatment period.

The interaction asks if if the post change in the treated group differs from the post change in the control group: the **treatment effect**.

## DiD as a regression

With two groups and two periods

$$y_{it} = \beta_0 + \beta_1 \text{Trt}_i + \beta_2 \text{Post}_t + \delta \left( \text{Trt}_i \times \text{Post}_t \right) + u_{it}$$

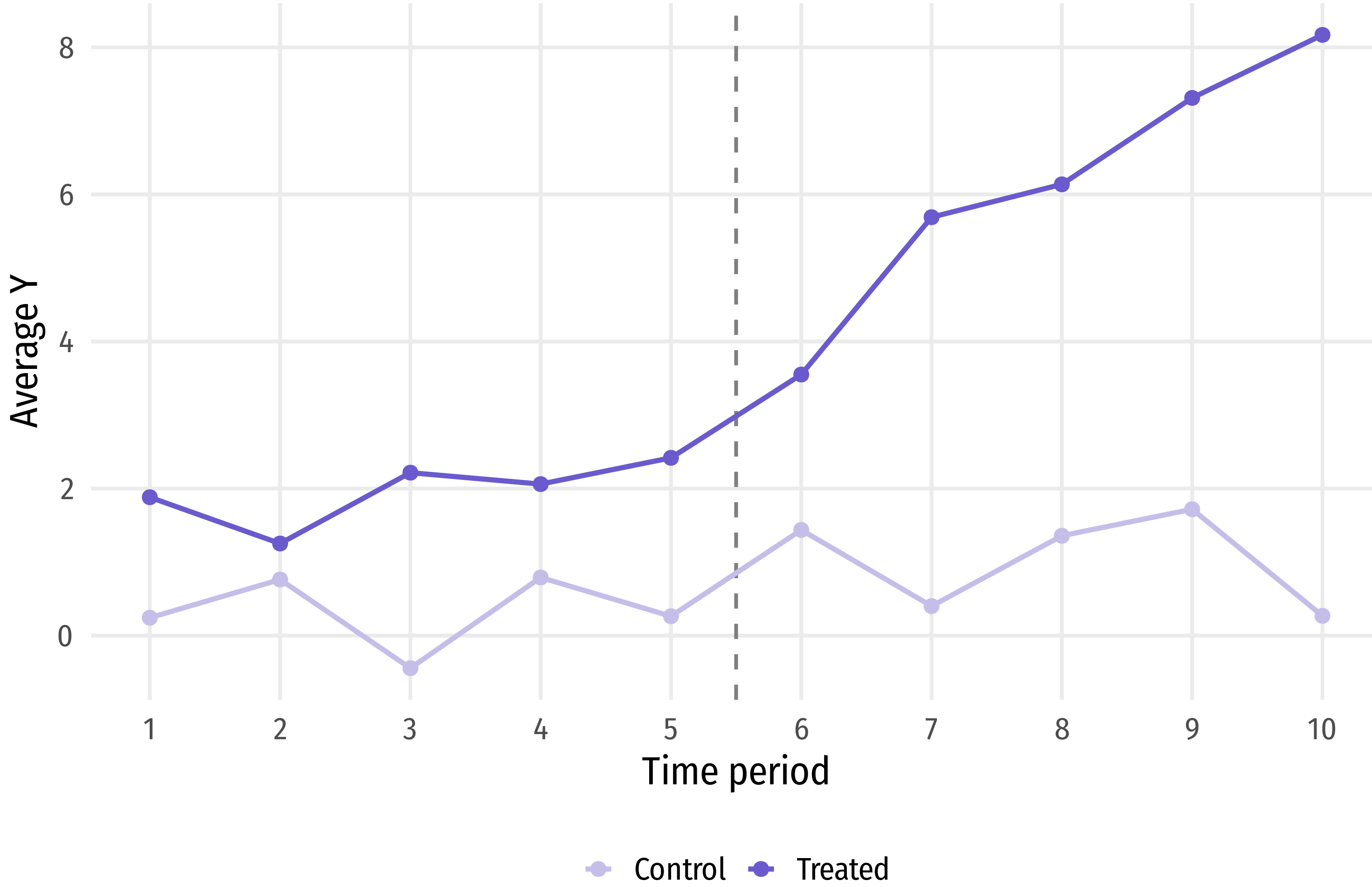
where $\text{Trt}_i$ and $\text{Post}_t$ are binary indicators for the treated group and post-treatment period.

The interaction asks if if the post change in the treated group differs from the post change in the control group: the **treatment effect**.

With panel data, we often write

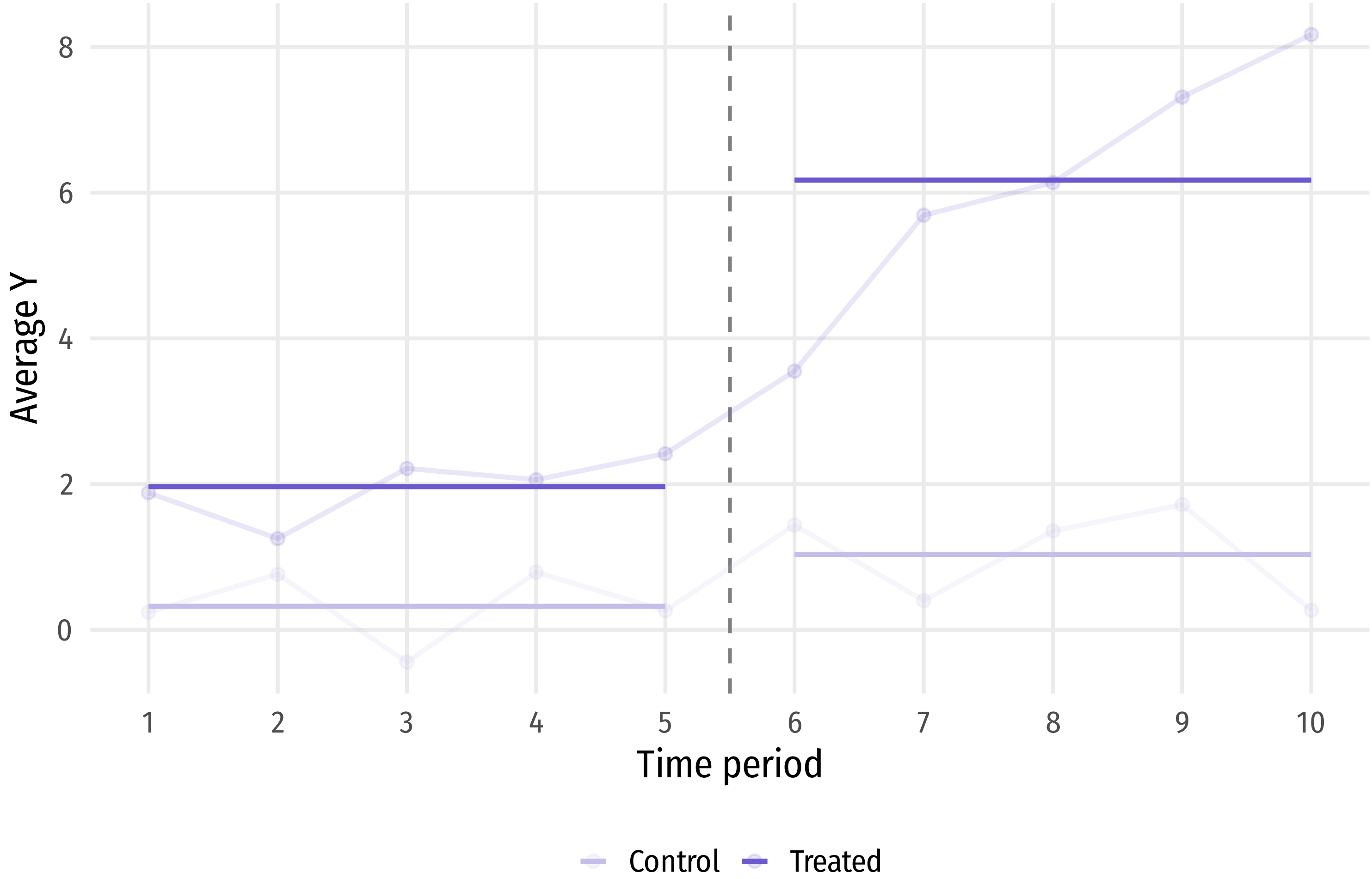
$$y_{it} = \alpha_i + \lambda_t + \delta D_{it} + u_{it}$$

where $D_{it} = 1$ only for treated units after treatment starts.

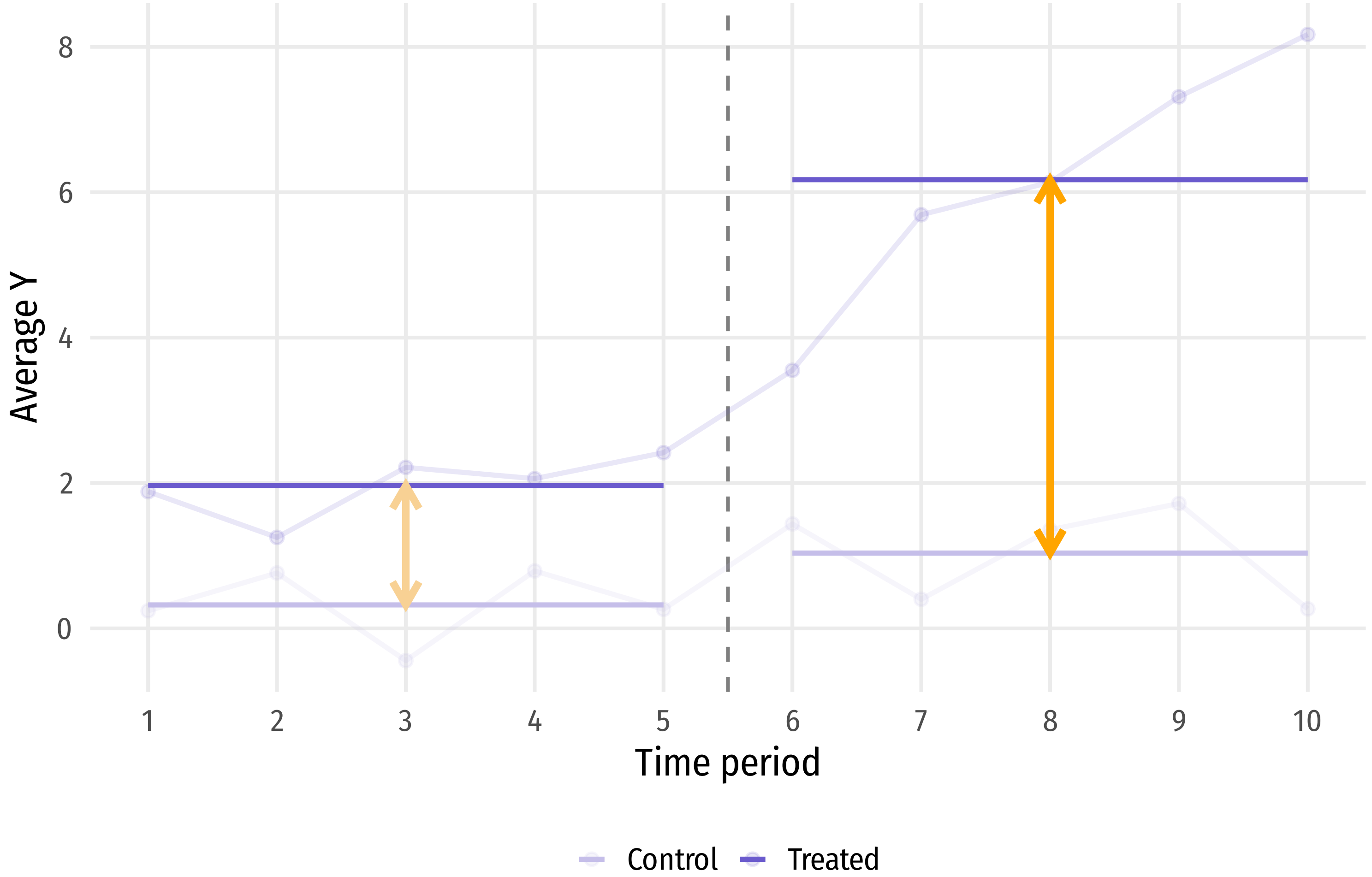



**DiD in pictures** Starting with time series for each unit.

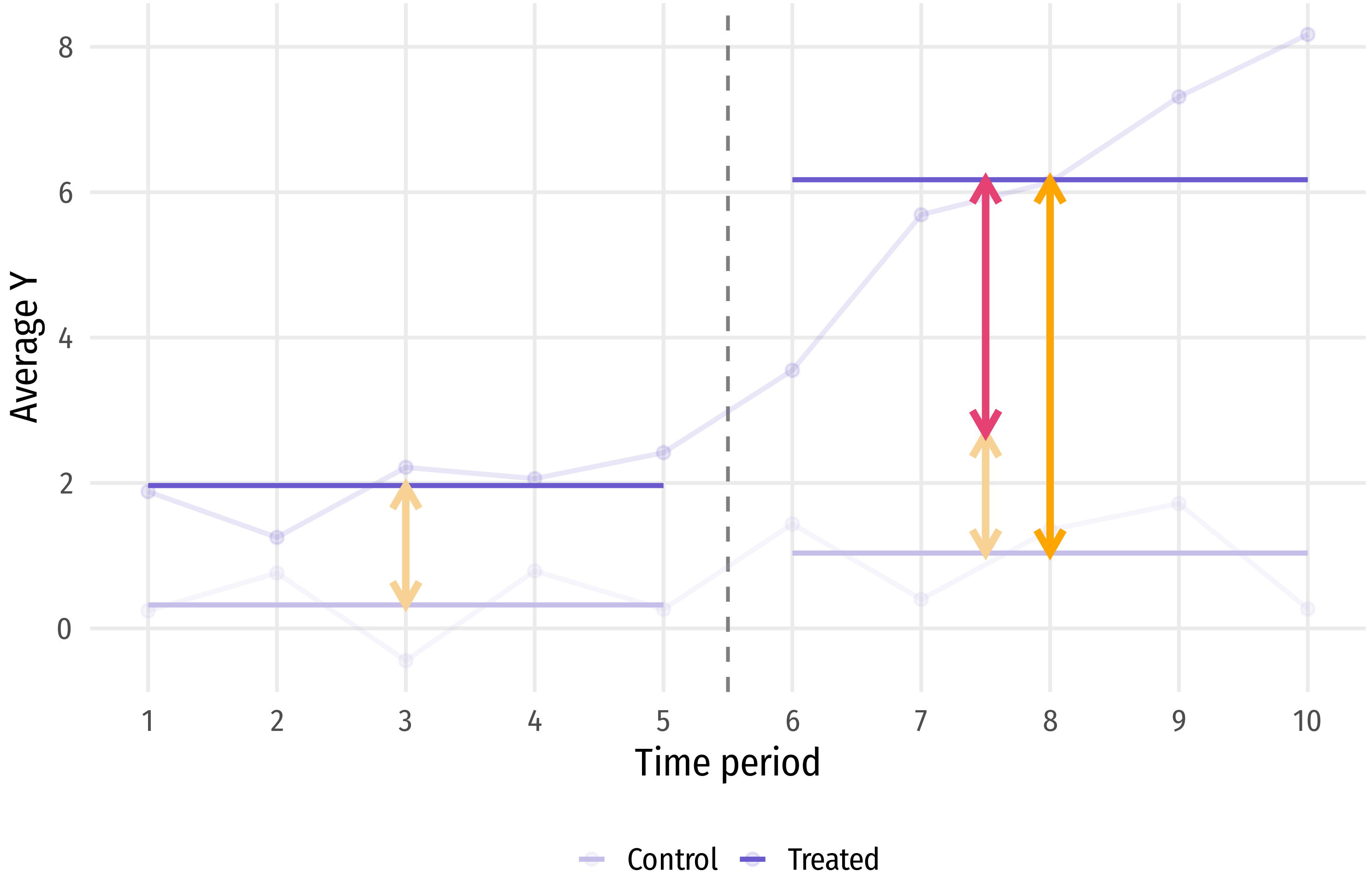
*(Treatment starts between periods 5 and 6.)*




Here we have **group**-by-**time** means.




The DiD estimate compares the **pre-trt** difference to the **post-trt** diff.




The difference in the **pre-trt** and **post-trt** diffs is the DiD estimate.

# DiD in `fixest`

```
did_mod =
  feols(
    y ~ i(post, treat, ref = 0) | id + period,
    cluster = 'id',
    data = base_did
  )
```

## DiD in `fixest`

```r
did_mod =
  feols(
    y ~ i(post, treat, ref = 0) | id + period,
    cluster = 'id',
    data = base_did
  )
```

```
#> OLS estimation, Dep. Var.: y
#> Observations: 1,080
#> Fixed-effects: id: 108,  period: 10
#> Standard-errors: Clustered (id)
#>                  Estimate Std. Error t value   Pr(>|t|)
#> post::1:treat  3.49339   0.521434 6.69959 1.0069e-09  ***
#> ---
#> Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
#> RMSE: 4.74465     Adj. R2: 0.202735
#>                   Within R2: 0.032761
```

The interaction btn post and treat is the **DiD estimate** of the trt effect.

## The key DiD assumption

DiD relies upon a critical assumption often called **parallel trends**:

> In the absence of treatment, treated and control units would have followed the same average trend.

Again, we're assuming that the trend of the control units provide a **good counterfactual** for the treated group.

## What threatens DiD?

So what violates this *parallel trends* assumption?

## What threatens DiD?

So what violates this *parallel trends* assumption?

- treated units were already on a different trend,
- anticipation changes behavior before treatment,
- other shocks hit treated units at the same time as treatment.

# Difference-in-differences

## What threatens DiD?

So what violates this *parallel trends* assumption?

- treated units were already on a different trend,
- anticipation changes behavior before treatment,
- other shocks hit treated units at the same time as treatment.

What else can *mess up* DiD estimates?

- standard errors ignore within-unit dependence,
- measurement error,
- spillovers from trt to control,
- dynamic treatment effects,
- het trt effects *and* het trt timing...

## DiD's place in the toolbox

DiD is especially useful when

- treatment timing is clear,
- treatment definition is clear,
- you have good pre-treatment data
- ctrl group provide a plausible counterfactual for trt group,
- the treatment effect is tied to a discrete policy or event.

## DiD's place in the toolbox

DiD is especially useful when

- treatment timing is clear,
- treatment definition is clear,
- you have good pre-treatment data
- ctrl group provide a plausible counterfactual for trt group,
- the treatment effect is tied to a discrete policy or event.

**Q** What are the other options?

## DiD's place in the toolbox

DiD is especially useful when

- treatment timing is clear,
- treatment definition is clear,
- you have good pre-treatment data
- ctrl group provide a plausible counterfactual for trt group,
- the treatment effect is tied to a discrete policy or event.

**Q** What are the other options?

**A** Synthetic control and friends—build a weighted average of control units to better match the treated unit's pre-treatment path.

# Difference-in-differences

## Takeaways

Key concepts from panel estimation, FEs, and DiD:

- **Panel data** let us compare a unit to itself **over time**.
- **Fixed effects** remove **time-invariant omitted variables**.
- $\implies$ **Identifying variation** is **within variation**.
- **DiD** focuses on how trt-ctrl **differences** change over time.

Panels can certainly help, but they still depend on strong assumptions.