# Gov 50: 8. Observational Studies

Matthew Blackwell

Harvard University

# Roadmap

1. Observational Studies

# 1/ Observational Studies

# Do newspaper endorsements matter?

- Can newspaper endorsements change voters' minds?

# Do newspaper endorsements matter?

- Can newspaper endorsements change voters' minds?
- Why not compare vote choice of readers of different papers?

# Do newspaper endorsements matter?

- Can newspaper endorsements change voters' minds?
- Why not compare vote choice of readers of different papers?
    - Problem: readers choose papers based on their previous beliefs.

# Do newspaper endorsements matter?

- Can newspaper endorsements change voters' minds?
- Why not compare vote choice of readers of different papers?
    - Problem: readers choose papers based on their previous beliefs.
    - Liberals $\rightsquigarrow$ New York Times, conservatives $\rightsquigarrow$ Wall Street Journal.

# Do newspaper endorsements matter?

- Can newspaper endorsements change voters' minds?
- Why not compare vote choice of readers of different papers?
    - Problem: readers choose papers based on their previous beliefs.
    - Liberals $\rightsquigarrow$ New York Times, conservatives $\rightsquigarrow$ Wall Street Journal.
- Study for today: British newspapers switching their endorsements.

# Do newspaper endorsements matter?

- Can newspaper endorsements change voters' minds?
- Why not compare vote choice of readers of different papers?
    - Problem: readers choose papers based on their previous beliefs.
    - Liberals $\rightsquigarrow$ New York Times, conservatives $\rightsquigarrow$ Wall Street Journal.
- Study for today: British newspapers switching their endorsements.
    - Some newspapers endorsing Tories in 1992 switched to Labour in 1997.

# Do newspaper endorsements matter?

- Can newspaper endorsements change voters' minds?
- Why not compare vote choice of readers of different papers?
  - Problem: readers choose papers based on their previous beliefs.
  - Liberals $\rightsquigarrow$ New York Times, conservatives $\rightsquigarrow$ Wall Street Journal.
- Study for today: British newspapers switching their endorsements.
  - Some newspapers endorsing Tories in 1992 switched to Labour in 1997.
  - **Treated group**: readers of Tory $\to$ Labour papers.

# Do newspaper endorsements matter?

- Can newspaper endorsements change voters' minds?
- Why not compare vote choice of readers of different papers?
  - Problem: readers choose papers based on their previous beliefs.
  - Liberals $\rightsquigarrow$ New York Times, conservatives $\rightsquigarrow$ Wall Street Journal.
- Study for today: British newspapers switching their endorsements.
  - Some newspapers endorsing Tories in 1992 switched to Labour in 1997.
  - **Treated group**: readers of Tory $\rightarrow$ Labour papers.
  - **Control group**: readers of papers who didn't switch.

# Data

| Name | Description |
| --- | --- |
| to_labour | Read a newspaper that switched endorsement to Labour between 1992 and 1997 (1=Yes, 0=No)? |
| vote_lab_92 | Did respondent vote for Labour in 1992 election (1=Yes, 0=No)? |
| vote_lab_97 | Did respondent vote for Labour in 1997 election (1=Yes, 0=No)? |
| age | Age of respondent |
| male | Does the respondent identify as Male (1=Yes, 0=No)? |
| parent_labour | Did the respondent's parents vote for Labour (1=Yes, 0=No)? |
| work_class | Does the respondent identify as working class (1=Yes, 0=No)? |

```r
library(tidyverse)
library(gov50data)
newspapers
```

```
## # A tibble: 1,593 x 7
##    to_labour vote_lab_92 vote_lab_97        age  male
##        <dbl>       <dbl>       <dbl> <hvn_lbll> <dbl>
##  1         0           1           1         33     0
##  2         0           1           0         51     0
##  3         0           0           0         46     0
##  4         0           1           1         45     1
##  5         0           1           1         29     0
##  6         0           1           1         47     1
##  7         0           1           1         34     1
##  8         0           1           1         31     0
##  9         0           1           1         24     1
## 10         1           1           1         48     0
## # i 1,583 more rows
## # i 2 more variables: parent_labour <dbl>, work_class <dbl>
```

# Observational studies

- Example of an **observational study**:

# Observational studies

- Example of an **observational study**:
  - We as researchers observe a naturally assigned treatment

# Observational studies

- Example of an **observational study**:
    - We as researchers observe a naturally assigned treatment
    - Very common: often can't randomize for ethical/logistical reasons.

# Observational studies

- Example of an **observational study**:
  - We as researchers observe a naturally assigned treatment
  - Very common: often can't randomize for ethical/logistical reasons.
- **Internal validity**: are the causal assumption satisfied? Can we interpret this as a causal effect?

# Observational studies

- Example of an **observational study**:
    - We as researchers observe a naturally assigned treatment
    - Very common: often can't randomize for ethical/logistical reasons.
- **Internal validity**: are the causal assumption satisfied? Can we interpret this as a causal effect?
    - RCTs usually have higher internal validity.

# Observational studies

- Example of an **observational study**:
    - We as researchers observe a naturally assigned treatment
    - Very common: often can't randomize for ethical/logistical reasons.
- **Internal validity**: are the causal assumption satisfied? Can we interpret this as a causal effect?
    - RCTs usually have higher internal validity.
    - Observational studies less so because treatment and control groups may differ in ways that are hard to measure

# Observational studies

- Example of an **observational study**:
  - We as researchers observe a naturally assigned treatment
  - Very common: often can't randomize for ethical/logistical reasons.
- **Internal validity**: are the causal assumption satisfied? Can we interpret this as a causal effect?
  - RCTs usually have higher internal validity.
  - Observational studies less so because treatment and control groups may differ in ways that are hard to measure
- **External validity**: can the conclusions/estimated effects be generalized beyond this study?

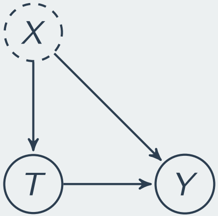
# Observational studies

- Example of an **observational study**:
  - We as researchers observe a naturally assigned treatment
  - Very common: often can't randomize for ethical/logistical reasons.
- **Internal validity**: are the causal assumption satisfied? Can we interpret this as a causal effect?
  - RCTs usually have higher internal validity.
  - Observational studies less so because treatment and control groups may differ in ways that are hard to measure
- **External validity**: can the conclusions/estimated effects be generalized beyond this study?
  - RCTs weaker here because often very expensive to conduct on representative samples.

# Observational studies

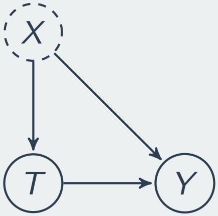
- Example of an **observational study**:
    - We as researchers observe a naturally assigned treatment
    - Very common: often can't randomize for ethical/logistical reasons.
- **Internal validity**: are the causal assumption satisfied? Can we interpret this as a causal effect?
    - RCTs usually have higher internal validity.
    - Observational studies less so because treatment and control groups may differ in ways that are hard to measure
- **External validity**: can the conclusions/estimated effects be generalized beyond this study?
    - RCTs weaker here because often very expensive to conduct on representative samples.
    - Observational studies often have larger/more representative samples that improve external validity.

# Confounding



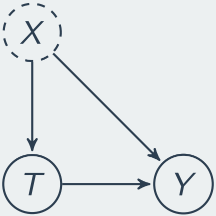

- **Confounder**: pre-treatment variable affecting treatment & the outcome.

# Confounding



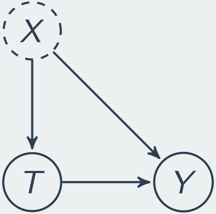

- **Confounder**: pre-treatment variable affecting treatment & the outcome.
    - Leftists ($X$) more likely to read newspapers switching to Labour ($T$).

# Confounding



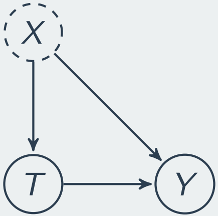

- **Confounder**: pre-treatment variable affecting treatment & the outcome.
    - Leftists ($X$) more likely to read newspapers switching to Labour ($T$).
    - Leftists ($X$) also more likely to vote for Labour ($Y$).

# Confounding



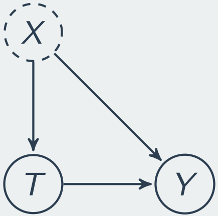

- **Confounder**: pre-treatment variable affecting treatment & the outcome.
  - Leftists ($X$) more likely to read newspapers switching to Labour ($T$).
  - Leftists ($X$) also more likely to vote for Labour ($Y$).
- **Confounding bias** in the estimated SATE due to these differences

# Confounding




- **Confounder**: pre-treatment variable affecting treatment & the outcome.
    - Leftists ($X$) more likely to read newspapers switching to Labour ($T$).
    - Leftists ($X$) also more likely to vote for Labour ($Y$).
- **Confounding bias** in the estimated SATE due to these differences
    - $\overline{Y}_{\text{control}}$ not a good proxy for $\frac{1}{n}\sum_{i=1}^{n} Y_i(0)$ in treated group.

# Confounding




- **Confounder**: pre-treatment variable affecting treatment & the outcome.
  - Leftists ($X$) more likely to read newspapers switching to Labour ($T$).
  - Leftists ($X$) also more likely to vote for Labour ($Y$).
- **Confounding bias** in the estimated SATE due to these differences
  - $\overline{Y}_{\text{control}}$ not a good proxy for $\frac{1}{n}\sum_{i=1}^{n} Y_i(0)$ in treated group.
  - one type: **selection bias** from self-selection into treatment

# Research designs

- How can we find a good comparison group?

# Research designs

- How can we find a good comparison group?
- Depends on the data we have available.

# Research designs

- How can we find a good comparison group?
- Depends on the data we have available.
- Three general types of observational study **reseach designs**:

# Research designs

- How can we find a good comparison group?
- Depends on the data we have available.
- Three general types of observational study **reseach designs**:

    1. **Cross-sectional design**: compare outcomes treated and control units at one point in time.

# Research designs

- How can we find a good comparison group?
- Depends on the data we have available.
- Three general types of observational study **reseach designs**:

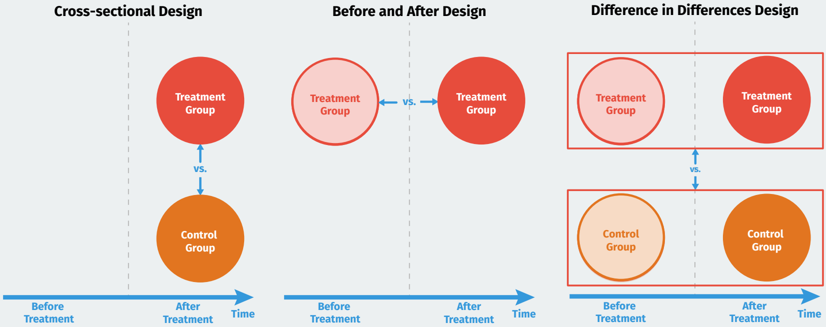
  1. **Cross-sectional design**: compare outcomes treated and control units at one point in time.
  2. **Before-and-after design**: compare outcomes before and after a unit has been treated, but need over-time data on treated group.

# Research designs

- How can we find a good comparison group?
- Depends on the data we have available.
- Three general types of observational study **reseach designs**:

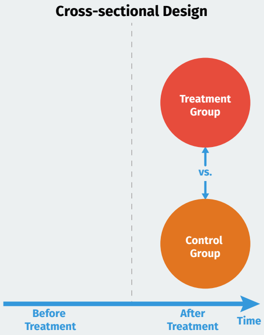
  1. **Cross-sectional design**: compare outcomes treated and control units at one point in time.
  2. **Before-and-after design**: compare outcomes before and after a unit has been treated, but need over-time data on treated group.
  3. **Difference-in-differences design**: use before/after information for the treated and control group; need over-time on treated & control group.

# Research designs

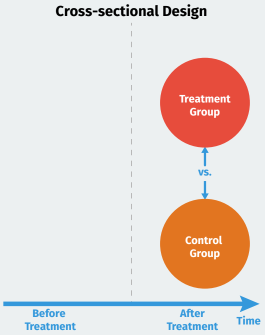


# Cross-sectional design

- Compare treated/control groups after treatment happens.

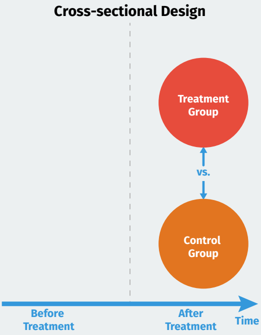


# Cross-sectional design



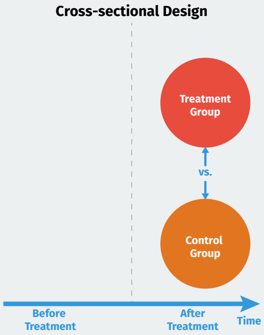

- Compare treated/control groups after treatment happens.
    - Switching readers vs non-switching readers in 1997.

# Cross-sectional design



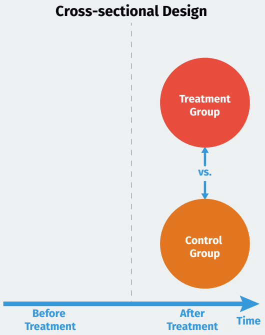

- Compare treated/control groups after treatment happens.
    - Switching readers vs non-switching readers in 1997.
- **Assumption**: groups identical on average (like RCTs)

# Cross-sectional design




- Compare treated/control groups after treatment happens.
    - Switching readers vs non-switching readers in 1997.
- **Assumption**: groups identical on average (like RCTs)
    - Sometimes called **unconfoundedness** or **as-if randomized**.

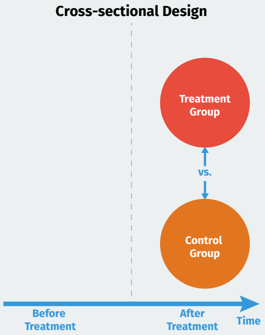
# Cross-sectional design




- Compare treated/control groups after treatment happens.
    - Switching readers vs non-switching readers in 1997.
- **Assumption**: groups identical on average (like RCTs)
    - Sometimes called **unconfoundedness** or **as-if randomized**.
- Cross-section estimate:

$$\overline{Y}^{\text{after}}_{\text{treated}} - \overline{Y}^{\text{after}}_{\text{control}}$$

# Cross-sectional design




- Compare treated/control groups after treatment happens.
  - Switching readers vs non-switching readers in 1997.
- **Assumption**: groups identical on average (like RCTs)
  - Sometimes called **unconfoundedness** or **as-if randomized**.
- Cross-section estimate:

$$\overline{Y}^{\text{after}}_{\text{treated}} - \overline{Y}^{\text{after}}_{\text{control}}$$

- Could there be confounders?

# Cross-sectional design in R

```r
switched <- newspapers |>
  filter(to_labour == 1) |>
  summarize(mean(vote_lab_97))


no_change <- newspapers |>
  filter(to_labour == 0) |>
  summarize(mean(vote_lab_97))

switched - no_change
```

```
##   mean(vote_lab_97)
## 1              0.14
```

# Statistical control

- **Statistical control**: adjust for confounders using statistical procedures.

# Statistical control

- **Statistical control**: adjust for confounders using statistical procedures.
    - Can help to reduce confounding bias.

# Statistical control

- **Statistical control**: adjust for confounders using statistical procedures.
  - Can help to reduce confounding bias.
- One type of statistical control: **subclassification**

# Statistical control

- **Statistical control**: adjust for confounders using statistical procedures.
  - Can help to reduce confounding bias.
- One type of statistical control: **subclassification**
  - Compare treated and control groups within levels of a confounder.

# Statistical control

- **Statistical control**: adjust for confounders using statistical procedures.
  - Can help to reduce confounding bias.
- One type of statistical control: **subclassification**
  - Compare treated and control groups within levels of a confounder.
  - Remaining effect can't be due to the confounder.

# Statistical control

- **Statistical control**: adjust for confounders using statistical procedures.
    - Can help to reduce confounding bias.
- One type of statistical control: **subclassification**
    - Compare treated and control groups within levels of a confounder.
    - Remaining effect can't be due to the confounder.
- Threat to inference: we can only control for observed variables $\rightsquigarrow$ threat of **unmeasured confounding**

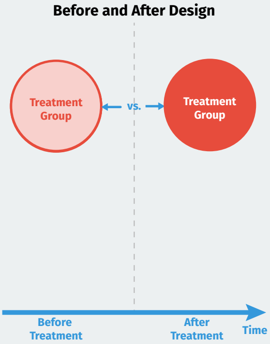
# Statistical control in R

```
newspapers |>
  group_by(parent_labour, to_labour) |>
  summarize(avg_vote = mean(vote_lab_97)) |>
  pivot_wider(
    names_from = to_labour,
    values_from = avg_vote
  ) |>
  mutate(diff_by_parent = `1` - `0`)
```

```
## # A tibble: 2 x 4
## # Groups:   parent_labour [2]
##   parent_labour   `0`   `1` diff_by_parent
##           <dbl> <dbl> <dbl>          <dbl>
## 1             0 0.279 0.434          0.155
## 2             1 0.597 0.698          0.101
```

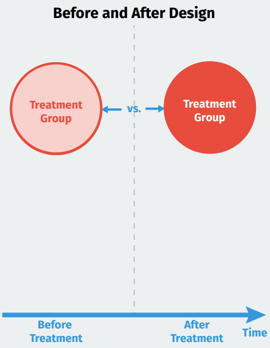
# Before-and-after comparison

- Compare readers of party-switching newspapers before & after switch.

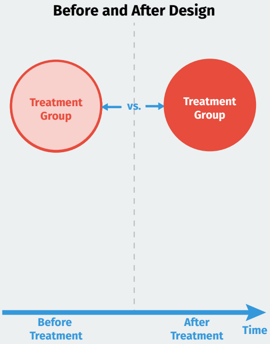


# Before-and-after comparison

- Compare readers of party-switching newspapers before & after switch.
- Advantage: all person-specific features held fixed

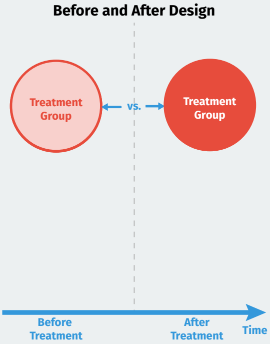


# Before-and-after comparison




- Compare readers of party-switching newspapers before & after switch.
- Advantage: all person-specific features held fixed
    - comparing within a person over time.

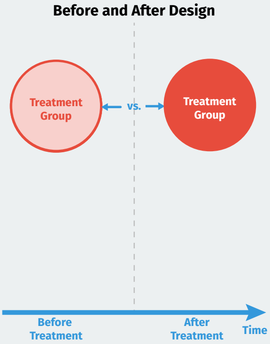
# Before-and-after comparison




- Compare readers of party-switching newspapers before & after switch.
- Advantage: all person-specific features held fixed
  - comparing within a person over time.
- Before-and-after estimate:

$$\overline{Y}^{\text{after}}_{\text{treated}} - \overline{Y}^{\text{before}}_{\text{treated}}$$

# Before-and-after comparison

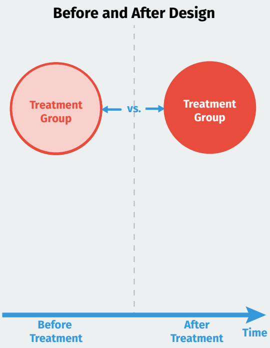



- Compare readers of party-switching newspapers before & after switch.
- Advantage: all person-specific features held fixed
    - comparing within a person over time.
- Before-and-after estimate:

$$\overline{Y}^{\text{after}}_{\text{treated}} - \overline{Y}^{\text{before}}_{\text{treated}}$$

- **Assumption**: no time-varying confounders

# Before-and-after comparison




- Compare readers of party-switching newspapers before & after switch.
- Advantage: all person-specific features held fixed
  - comparing within a person over time.
- Before-and-after estimate:

$$\overline{Y}^{\text{after}}_{\text{treated}} - \overline{Y}^{\text{before}}_{\text{treated}}$$

- **Assumption**: no time-varying confounders
  - Time trend: Labour just did better overall in 1997 compared to 1992.

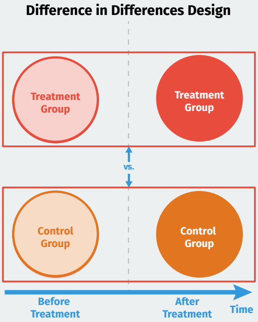
# Before and after in R

```
newspapers |>
  mutate(
    vote_change = vote_lab_97 - vote_lab_92
  ) |>
  summarize(avg_change = mean(vote_change))
```

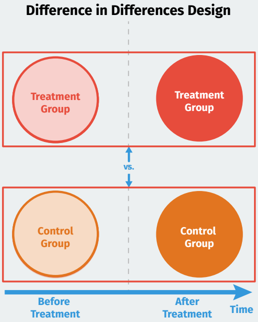
```
## # A tibble: 1 x 1
##   avg_change
##        <dbl>
## 1      0.119
```

# Differences in differences

- Use the before/after difference of **control group** to infer what would have happened to **treatment group** without treatment.



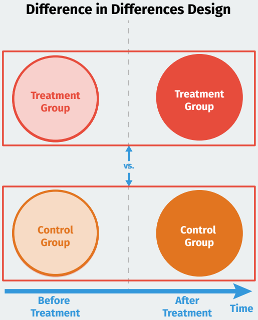
# Differences in differences




- Use the before/after difference of **control group** to infer what would have happened to **treatment group** without treatment.
- DiD estimate:

$$\underbrace{\left(\overline{Y}^{\text{after}}_{\text{treated}} - \overline{Y}^{\text{before}}_{\text{treated}}\right)}_{\text{trend in treated group}} - \underbrace{\left(\overline{Y}^{\text{after}}_{\text{control}} - \overline{Y}^{\text{before}}_{\text{control}}\right)}_{\text{trend in control group}}$$

# Differences in differences

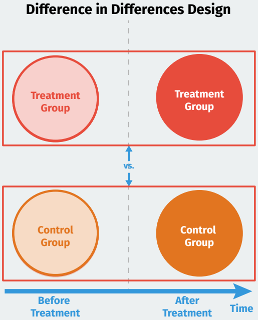



- Use the before/after difference of **control group** to infer what would have happened to **treatment group** without treatment.
- DiD estimate:

$$\underbrace{\left(\overline{Y}^{\text{after}}_{\text{treated}} - \overline{Y}^{\text{before}}_{\text{treated}}\right)}_{\text{trend in treated group}} - \underbrace{\left(\overline{Y}^{\text{after}}_{\text{control}} - \overline{Y}^{\text{before}}_{\text{control}}\right)}_{\text{trend in control group}}$$

- Change in treated group above and beyond the change in control group.

# Differences in differences

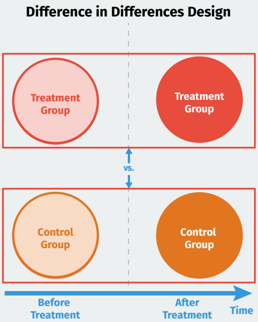



- Use the before/after difference of **control group** to infer what would have happened to **treatment group** without treatment.
- DiD estimate:

$$\underbrace{\left(\overline{Y}^{\text{after}}_{\text{treated}} - \overline{Y}^{\text{before}}_{\text{treated}}\right)}_{\text{trend in treated group}} - \underbrace{\left(\overline{Y}^{\text{after}}_{\text{control}} - \overline{Y}^{\text{before}}_{\text{control}}\right)}_{\text{trend in control group}}$$

- Change in treated group above and beyond the change in control group.
- **Assumption**: parallel trends

# Differences in differences

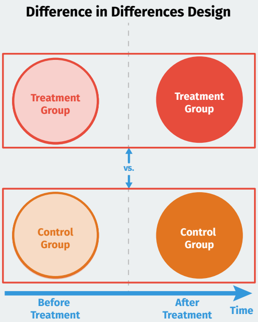



- Use the before/after difference of **control group** to infer what would have happened to **treatment group** without treatment.
- DiD estimate:

$$\underbrace{\left(\overline{Y}^{\text{after}}_{\text{treated}} - \overline{Y}^{\text{before}}_{\text{treated}}\right)}_{\text{trend in treated group}} - \underbrace{\left(\overline{Y}^{\text{after}}_{\text{control}} - \overline{Y}^{\text{before}}_{\text{control}}\right)}_{\text{trend in control group}}$$

- Change in treated group above and beyond the change in control group.
- **Assumption**: parallel trends
  - Changes in vote of readers of non-switching papers roughly the same as changes that readers of switching papers would have been if they read non-switching papers.

# Differences in differences




- Use the before/after difference of **control group** to infer what would have happened to **treatment group** without treatment.
- DiD estimate:

$$\underbrace{\left(\overline{Y}^{\text{after}}_{\text{treated}} - \overline{Y}^{\text{before}}_{\text{treated}}\right)}_{\text{trend in treated group}} - \underbrace{\left(\overline{Y}^{\text{after}}_{\text{control}} - \overline{Y}^{\text{before}}_{\text{control}}\right)}_{\text{trend in control group}}$$

- Change in treated group above and beyond the change in control group.
- **Assumption**: parallel trends
    - Changes in vote of readers of non-switching papers roughly the same as changes that readers of switching papers would have been if they read non-switching papers.
    - Threat to inference: non-parallel trends.

# Difference-in-differences in R

```r
newspapers |>
  mutate(
    vote_change = vote_lab_97 - vote_lab_92,
    to_labour = if_else(to_labour == 1, "switched", "unswitched")
  ) |>
  group_by(to_labour) |>
  summarize(avg_change = mean(vote_change)) |>
  pivot_wider(
    names_from = to_labour,
    values_from = avg_change
  ) |>
  mutate(DID = switched - unswitched)
```

```
## # A tibble: 1 x 3
##   switched unswitched    DID
##      <dbl>      <dbl>  <dbl>
## 1    0.190      0.110 0.0796
```

# Summarizing approaches

1. **Cross-sectional comparison**

# Summarizing approaches

1. **Cross-sectional comparison**
    - Compare treated units with control units after treatment

# Summarizing approaches

1. **Cross-sectional comparison**
    - Compare treated units with control units after treatment
    - Assumption: treated and controls units are comparable

# Summarizing approaches

1. **Cross-sectional comparison**
    - Compare treated units with control units after treatment
    - Assumption: treated and controls units are comparable
    - Possible confounding

# Summarizing approaches

1. **Cross-sectional comparison**
    - Compare treated units with control units after treatment
    - Assumption: treated and controls units are comparable
    - Possible confounding
2. **Before-and-after comparison**

# Summarizing approaches

1. **Cross-sectional comparison**
    - Compare treated units with control units after treatment
    - Assumption: treated and controls units are comparable
    - Possible confounding
2. **Before-and-after comparison**
    - Compare the same units before and after treatment

# Summarizing approaches

1. **Cross-sectional comparison**
    - Compare treated units with control units after treatment
    - Assumption: treated and controls units are comparable
    - Possible confounding
2. **Before-and-after comparison**
    - Compare the same units before and after treatment
    - Assumption: no time-varying confounding

# Summarizing approaches

1. **Cross-sectional comparison**
    - Compare treated units with control units after treatment
    - Assumption: treated and controls units are comparable
    - Possible confounding
2. **Before-and-after comparison**
    - Compare the same units before and after treatment
    - Assumption: no time-varying confounding
3. **Differences-in-differences**

# Summarizing approaches

1. **Cross-sectional comparison**
    - Compare treated units with control units after treatment
    - Assumption: treated and controls units are comparable
    - Possible confounding
2. **Before-and-after comparison**
    - Compare the same units before and after treatment
    - Assumption: no time-varying confounding
3. **Differences-in-differences**
    - Assumption: parallel trends assumptions

# Summarizing approaches

1. **Cross-sectional comparison**
    - Compare treated units with control units after treatment
    - Assumption: treated and controls units are comparable
    - Possible confounding
2. **Before-and-after comparison**
    - Compare the same units before and after treatment
    - Assumption: no time-varying confounding
3. **Differences-in-differences**
    - Assumption: parallel trends assumptions
    - Under this assumption, it accounts for unit-specific and time-varying confounding.

# Summarizing approaches

1. **Cross-sectional comparison**
    - Compare treated units with control units after treatment
    - Assumption: treated and controls units are comparable
    - Possible confounding
2. **Before-and-after comparison**
    - Compare the same units before and after treatment
    - Assumption: no time-varying confounding
3. **Differences-in-differences**
    - Assumption: parallel trends assumptions
    - Under this assumption, it accounts for unit-specific and time-varying confounding.

- All rely on assumptions that can't be verified to handle confounding.

# Summarizing approaches

1. **Cross-sectional comparison**
    - Compare treated units with control units after treatment
    - Assumption: treated and controls units are comparable
    - Possible confounding
2. **Before-and-after comparison**
    - Compare the same units before and after treatment
    - Assumption: no time-varying confounding
3. **Differences-in-differences**
    - Assumption: parallel trends assumptions
    - Under this assumption, it accounts for unit-specific and time-varying confounding.

- All rely on assumptions that can't be verified to handle confounding.
- RCTs handle confounding by design.

# Causality understanding check