Gov 51: Observational Studies

Matthew Blackwell

Harvard University

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.
- Could do a lab experiment, but there are concerns over **external validity**

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.
- Could do a lab experiment, but there are concerns over **external validity**
- 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.
- Could do a lab experiment, but there are concerns over **external validity**
- 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.
- Could do a lab experiment, but there are concerns over **external validity**
- 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.

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.
- Could do a lab experiment, but there are concerns over **external validity**
- Study for today: British newspapers switching their endorsements.
    - Some newspapers endorsing Tories in 1992 switched to Labour in 1997.
    - **Treated group**: readers of Tory $\longrightarrow$ Labour papers.
    - **Control group**: readers of papers who didn't switch.

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 pre-treatment variable may differ between treatment and control groups

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 pre-treatment variable may differ between treatment and control groups
- **External validity**: can the conclusions/estimated effects be generalized beyond this study?

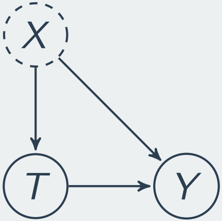
Observational studies

- 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 pre-treatment variable may differ between treatment and control groups
- **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

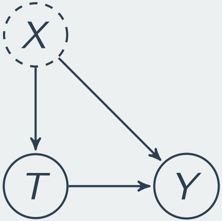
- Example of an **observational study**:
  - We as researchers observe a naturally assigned treatment
  - Very common: often can't randomize for ethical/logistical reasons.
- **Internal validity**: are the causal assumption satisfied? Can we interpret this as a causal effect?
  - RCTs usually have higher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



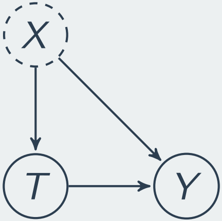
- **Confounder**: pre-treatment variable affecting treatment & the outcome.

Confounding



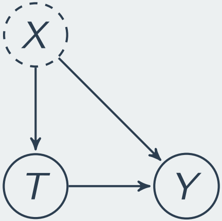
- **Confounder**: pre-treatment variable affecting treatment & the outcome.
  - Leftists (X) more likely to read newspapers switching to Labour (T).

Confounding



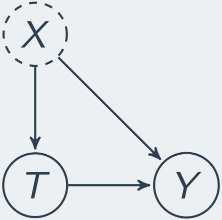
- **Confounder**: pre-treatment variable affecting treatment & the outcome.
  - Leftists (X) more likely to read newspapers switching to Labour (T).
  - Leftists (X) also more likely to vote for Labour (Y).

Confounding



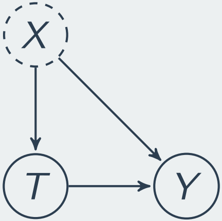
- **Confounder**: pre-treatment variable affecting treatment & the outcome.
    - Leftists (X) more likely to read newspapers switching to Labour (T).
    - Leftists (X) also more likely to vote for Labour (Y).
- **Confounding bias** in the estimated SATE due to these differences

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 $Y_i(0)$ 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 $Y_i(0)$ 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.

Cross-sectional design

- Compare treatment and control groups after treatment happens.

Cross-sectional design

- Compare treatment and control groups after treatment happens.
    - Readers of switching papers vs readers of non-switching papers in 1997.

Cross-sectional design

- Compare treatment and control groups after treatment happens.
    - Readers of switching papers vs readers of non-switching papers in 1997.
- Treatment & control groups assumed identical on average as in RCT.

Cross-sectional design

- Compare treatment and control groups after treatment happens.
  - Readers of switching papers vs readers of non-switching papers in 1997.
- Treatment & control groups assumed identical on average as in RCT.
  - Sometimes called **unconfoundedness** or **as-if randomized**.

Cross-sectional design

- Compare treatment and control groups after treatment happens.
    - Readers of switching papers vs readers of non-switching papers in 1997.
- Treatment & control groups assumed identical on average as in RCT.
    - Sometimes called **unconfoundedness** or **as-if randomized**.
- Cross-section comparison estimate:

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

Cross-sectional design

- Compare treatment and control groups after treatment happens.
  - Readers of switching papers vs readers of non-switching papers in 1997.
- Treatment & control groups assumed identical on average as in RCT.
  - Sometimes called **unconfoundedness** or **as-if randomized**.
- Cross-section comparison estimate:

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

- Could there be confounders?

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**

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:

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

- Threat to inference: **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:

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

- Threat to inference: **time-varying confounders**
  - Time trend: Labour just did better overall in 1997 compared to 1992.

Differences in differences

- Key idea: use the before-and-after difference of **control group** to infer what would have happend to **treatment group** without treatment.

Differences in differences

- Key idea: use the before-and-after difference of **control group** to infer what would have happend 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

- Key idea: use the before-and-after difference of **control group** to infer what would have happend to **treatment group** without treatment.
- DiD estimate:

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

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

Differences in differences

- Key idea: use the before-and-after difference of **control group** to infer what would have happend 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.
- **Parallel time trend assumption**

Differences in differences

- Key idea: use the before-and-after difference of **control group** to infer what would have happend 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.
- **Parallel time trend assumption**
  - 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

- Key idea: use the before-and-after difference of **control group** to infer what would have happend 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.
- **Parallel time trend assumption**
    - 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.

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