Causal Analysis

Chapter 19: A Framework for Causal Analysis

Álmos Telegdy

Corvinus University of Budapest

The setup: Intervention, treatment, subjects, outcomes

- **Intervention** describes a decision that aims changing the behavior or situation of people, firms. Also called **Treatment**
- **Subjects** of an intervention are those that may be affected. Treated or untreated
- **Outcome variables**, or outcomes, are variables that may be affected by the intervention
- **Causal variables, or treatment variables** are the variables that indicate the intervention
- **Mechanism** by which an intervention may affect an outcome variable

The causal question

Most important elements of a precise causal question are

- What's the outcome (Y) variable?
- What's the causal (X) variable?
  - The causal variable may be a binary variable (intervention takes place or not) or a quantitative variable (amount of intervention)
- What are the subjects (the outcome for whom?)
- What is the specific intervention (who, and how, would manipulate the cause to alter the outcome?)
- What is or could be the mechanism (why should one expect an effect of the intervention on the subject?)

Potential outcomes framework

Potential outcomes: a structure to study causal questions

- We study binary treatment (treated/not treated)
- The outcome variable may be binary or continuous

Potential outcomes framework

Imagine **two potential outcomes for each subject**:

1. **Treated outcome** (y_i^1): what would the outcome be if subjects were treated
2. **Untreated outcome** (y_i^0): what would the outcome be if subjects were untreated

Only one of these two outcomes materializes for each subject (unless you have a time machine)

1. Treated subject: observed outcome is the treated outcome (y_i^1)
2. Untreated subject: observed outcome is the untreated outcome (y_i^0)

- The other, **unobserved**, potential outcome is the **counterfactual outcome** $\rightarrow$ what would the outcome be if the subject experienced what did not happen

Individual treatment effect

The **individual treatment effect (ITE)** of the intervention for subject i

- equals to the difference of the two potential outcomes
- $te_i = y_i^1 - y_i^0$

Individual treatment effect

The **individual treatment effect (ITE)** of the intervention for subject i

- equals to the difference of the two potential outcomes
- $te_i = y_i^1 - y_i^0$

ITE is not observable $\rightarrow$ In the data we observe y_i^1 OR y_i^0

- Treated subject: y_i^0 not observed
- Untreated subject: y_i^1 not observed

Average treatment effect

We do not observe ITE, but we can compute its expectation (or average)

The **average treatment effect (ATE)** equals to the average of ITE across all subjects

$$ATE = E[te_i] = E[y_i^1 - y_i^0] = E[y_i^1] - E[y_i^0] \tag{1}$$

(The average of the differences equals the difference of the averages.)

When is ATE a good proxy of ITE?

Two questions

1. Is the average outcome of the actually treated subjects a good approximation of the average outcome of all potentially treated subjects?
2. Is the average outcome of the actually untreated subjects a good approximation of the average outcome of all potentially untreated subjects?

$$E[y_i | \, i \text{ treated} \,] \stackrel{?}{\approx} E[y_i^1] \tag{2}$$

$$E[y_i | \, i \text{ untreated} \,] \stackrel{?}{\approx} E[y_i^0] \tag{3}$$

Example: Effect of individual action on global warming

What is the effect of vegetarianism on global worming?

- $x =$ vegetarian; $y = CO_2$ emission

- $E[y|\, i\ \textit{treated}\,] =$ emission of actual vegetarians – observable
- $E[y^1] =$ emission of vegetarians – not observable
- $E[y|\, i\ \textit{untreated}\,] =$ emission of actual non-vegetarians – observable
- $E[y^0] =$ emission of non-vegetarians – not observable

Example: Effect of individual action on global warming

What is the effect of vegetarianism on global worming?

- $x =$ vegetarian; $y = CO_2$ emission

- $E[y| \, i \textit{ treated} \,]$ = emission of actual vegetarians – observable
- $E[y^1]$ = emission of vegetarians – not observable
- $E[y| \, i \textit{ untreated} \,]$ = emission of actual non-vegetarians – observable
- $E[y^0]$ = emission of non-vegetarians – not observable

$E[y_i| \, i \textit{ treated} \,] = E[y_i^1] \rightarrow$ probably true: they do not eat meat
$E[y_i| \, i \textit{ untreated} \,] = E[y_i^0] \rightarrow$ probably not true: those who do not like meat, become vegetarian

ATE is the treatment effect

- We usually think of ATE, when we talk of the effect of the intervention $\Rightarrow$ the difference in averages is a natural measure of the treatement effect
- If we multiply ATE with the size of the population, we get the total effect of the intervention (policy makers are often interested in it)
  - Example: multiply the treatment effect of firm subsidies with the employment of subsidized firms and you get the total jobs created by the subsidies

ATE on the treated

There is a special group: those which received the treatment

- **Average Treatment Effect on the Treated (ATET)**

$$ATET = E[y_i^1 - y_i^0 | D = 1]$$

$$ATET = E[y_i^1 | D = 1] - E[y_i^0 | D = 1]$$

- ATET highlights the essence of the potential outcomes framework: **we compare the outcome with the counterfactual**
- ATET cannot be observed directly (same reason as for ITE), but sometimes can be estimated
- ATET does not have to be equal to ATE (but they may be equal)

Example: book ratings

Question: What is the rating of Harry Potter?

Are book ratings ATE or ATET?

Example: book ratings

Question: What is the rating of Harry Potter?

Are book ratings ATE or ATET?

- It depends on our assumptions
- Only those readers fill it out who were interested in the genre $\rightarrow$ young people and those who like fantasy literature
- This is the focus of the editors $\rightarrow$ ATET

Heterogeneous ATE

ATE can be measured in subgroups of the data

- Can be computed if we look only at a certain group
  - e.g., men-women, SMEs-large firms, cities-villages

Quantitative causal variables

Treatment can be continuous

- value of subsidy the firm received
- hours of education received by unemployed
- amount of subsidy poor families receive

The potential outcomes framework was developed to deal with binary treatment

- Is also works for continuous treatment, but it is more complicated

ATE and Quantitative Causal Variables

- Quantitative causal variables lead to a series of individual treatment effect (instead of one)
- Difficult to think about average effects of a quantitative causal variables
- But the idea is fundamentally the same
- Often transform a quantitative into binary: low vs high
  - even if you don't do this, it's easier to think this way

Ceteris paribus

Other things being the same

- Our goal: the difference between treated and untreated outcomes is the intervention **and only the intervention**
- All things that affect the outcome variable are the same in the treated and untreated groups
- The best is ITE: the outcome is measured for the same subjects in two states of the world (treated–untreated)
- In reality this is not possible and so ATE compares different subjects

Example

Does the training of unemployed affect the probability of getting a job?

- How similar are the treated and untreated groups?

Example

Does the training of unemployed affect the probability of getting a job?

- How similar are the treated and untreated groups?
  - Those who are more able to use the knowledge taught, are more likely to take the training
  - Those who are more able to use the knowledge taught, find jobs more easily (regardless of the training)
- Only those characteristics matter, which **affect both** the causal and the outcome variables

Ceteris paribus vs. multivariate regression

$$y^E = \beta_0 + \beta_1 x + \beta_2 z \tag{4}$$

- In regression we **condition** on the vector $z \rightarrow$ compare two observations that have the same z but are different in x by one unit.

- Can we condition on **all** potential confounders in regression? $\Rightarrow$ that would be ceteris paribus analysis
- Probably not
  - We can include only what we observe in data
  - We cannot be sure that there are no confounders among what's not observed in data
  - How do we know that we controlled for everything relevant?

ATE is our main interest

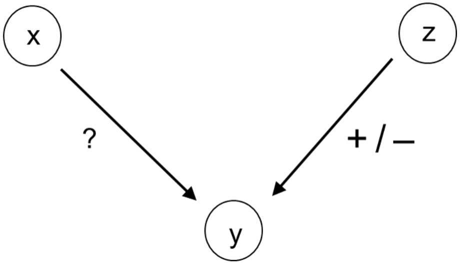
- How to calculate ATE - main issue for this course
    - Because te_i cannot be calculated and averaged
    - Because we need to work hard to get close to ceteris paribus

Causal maps to uncover causal structure

- Causal maps: key tool to think about causality
- A causal map (diagram, graph) = graph that connects variables (nodes) with arrows (directed edges).
- The arrows represent effects.
- Another name for causal map is **directed acyclic graphs, DAG** - graph of nodes and arrows.

DAG: simplest case

- An example with x causing y, but also a variable z causing y
- When an outcome variable is caused by the intervention of interest (x) but also other variables like z
- On this graph x and z are unrelated

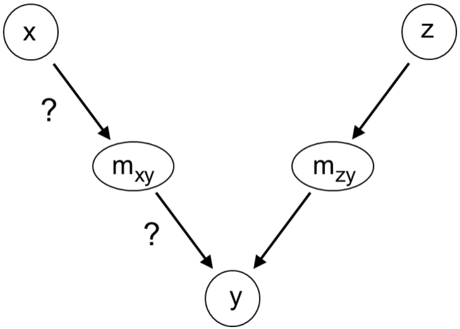


Causal maps to uncover causal structure

- Our aim: summarizing our assumptions about how variables affect each other
- A causal map is a graph that connects variables (nodes) with arrows (directed edges)
- The arrows represent effects
- Causal maps help understand **whether** and **how** we can uncover the effect we are after

DAG: mechanisms

- Add variables that measure the mechanisms (m) through which x and z affect y
- m_{zx} = through which x affects y
- m_{zy} = through which z affects y



Example: TV ads and product purchase

- Potential outcomes = 0 (no purchase) or 1 (purchase)
- $te_i = \{-1, 0, 1\}$
- ATE = the change in the likelihood of purchasing the product due to seeing the ad
  - That is the combination of the three possible treatment effects $(1, 0, -1)$
- The higher the proportion of people with treatment effect 1, the more positive the average effect

Example: TV ads and product purchase

- Without being presented the ad, 10% of the subjects would buy the product
  - Untreated PO = 1 for 10% of the subjects and 0 for 90%
- If presented the ad, 11% of the subjects would buy the product
  - Treated PO = 1 for 11% of the subjects and 0 for 89%
- The average treatment effect here is 1 percentage point: $ATE = 0.01$.

Random assignment

How can we make ATE to be a good approximation of ITE?
Remember:

$$E[y_i|\ i\ treated\ ] \stackrel{?}{\approx} E[y_i^1] \tag{5}$$

$$E[y_i|\ i\ untreated\ ] \stackrel{?}{\approx} E[y_i^0] \tag{6}$$

Solution

- Find a rule which assigns x to subjects such that it does not affect $y \Rightarrow$ **treatment will be independent of outcomes**

Random assignment $\rightarrow ATE$

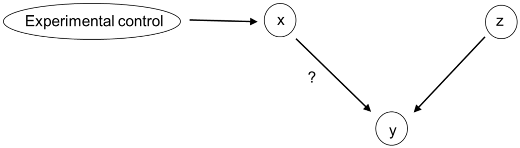
If assignment to x is randomized, the **average y in treated and untreated groups would be the same in absence of the treatment**

- The potential average treated outcome will be the same in the treated and untreated groups
- The potential average untreated outcome be the same in the treated and untreated groups
- $ATE = E[ITE_i]$

This is the most straightforward method to compute ATE (at least in theory)

DAG of random assignment

- Question: how does x affect y?
- There is nothing that affects x, except the randomized assignment
- All other z variables may have an effect on y but not on x



Natural experiments

Sometimes nature randomized the allocation of subjects to treatment status

- ... as if it came from a controlled experiment

Natural experiments

Sometimes nature randomized the allocation of subjects to treatment status

- ... as if it came from a controlled experiment

Example: consequences of slave trade in Africa

- Are countries with lots of slave trade poorer today?

Natural experiments

Sometimes nature randomized the allocation of subjects to treatment status

- ... as if it came from a controlled experiment

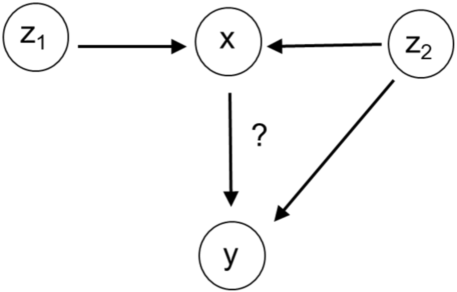
Example: consequences of slave trade in Africa

- Are countries with lots of slave trade poorer today?
  - The reasons slaves were taken 150 years ago do not correlate with development (= the grade of development 150 years ago does not differ between the two groups)
  - Natural experiment: slave trade (causal variable) is not correlated with current development (outcome variable)
  - The gap between Africa and the rest of the world would be 12-47% smaller
  - Mechanism: fewer people (especially young), less trust

Observational data

Most of the time, we use nonexperimental – "normal" – data

- Treatment variable is affected by other variables

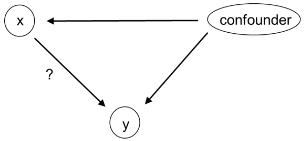


Confounders in observational data

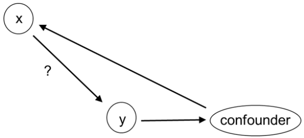
Types of endogeneity

1. **Common cause confounder**: $z \rightarrow x, y$
2. **Reverse causality**: $y \rightarrow x$
3. **Unwanted mechanism**: $x \rightarrow y$, but not through the desired mechanism

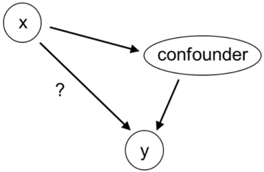
Types of endogeneity



(a) Common cause



(b) Reverse causality



(c) Unwanted mechanism

Example: common cause confounder

Question: what is the effect of higher education on wages?

- Clever people study with ease and so they often go to university
- Clever people usually earn well

$\rightarrow$ Even if higher education does increase earning capacity, the estimated coefficient of higher education on earnings will be positive $\Rightarrow$ **self-selection**

Example: reverse causality

Question: What is the effect of advertising on sales?

- If sales decline, the management may start an aggressive ad campaign – **self-selection**
- sales may decline for a structural reason (e.g., the product is obsolete)

$\rightarrow$ Even if advertising raises sales, the estimated coefficient of ads on sales will be negative

Example: unwanted mechanism

Question: what is the effect of an abortion ban on the education of children born after the ban?

- Educated women overrepresented in abortions – **self-selection**
  - Kids from educated families tend to go to university

$\rightarrow$ Even if abortion ban decreases the level of education of children, the estimated coefficient of abortion ban on education will be positive

Control variables

Data problems: missing variables, proxy variables

1. We do not always have all the variables we need → **missing variables**
2. We usually do not measure well what we want to control for → **proxy variables**
   - E.g., the number of workers is only an approximate measure for the labor input of firms (hours worked, abilities of workers, motivation of workers, etc.)

What to control for?

To be on the safe side, should we control for all variables in the data?

What to control for?

To be on the safe side, should we control for all variables in the data?

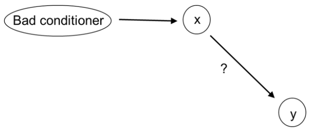
No

The three types of bad conditioning variables

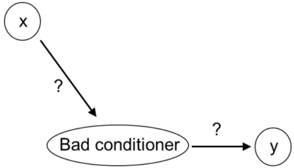
- **Exogenous source of variation** in the causal variable x
- **Part of the mechanism** by which x affects y – that is of course if we want to include that mechanism in the effect we want to uncover
- **Collider variable**: a common effect, or common consequence, of both x and y

- How to know if we should condition on a variable or not?
  - Use your common sense
  - Causal map helps

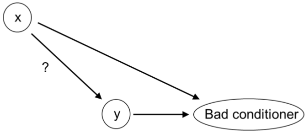
Causal maps of bad conditioners



(a) Exogenous source of variation



(b) Part of mechanism



(c) Common consequence

Examples: bad conditioners

Question: do ads increase sales?

- **Exogenous source of variation:** the ad is on internet only on given days of the week – do not control for it

Question: Do investment subsidies increase firm exports?

- **Mechanism:** the subsidy increases firm size – do not control for it

Question: does arthritis cause heart conditions?

- **Common consequence:** both illnesses require hospitalization – do not control for it
  - $\rightarrow$ You "control" for hospitalization even if you use a sample hospitalized people

Case study: food and health

- Causal statement: some kinds of food make you healthier than other kinds of food.
- Question: does eating fruit and vegetables help us avoid high blood pressure?

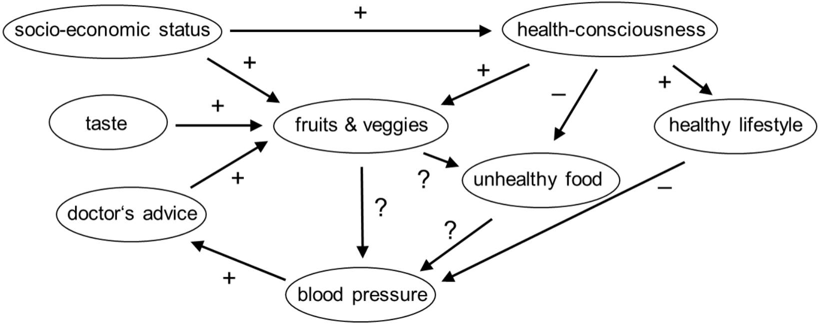
- Data: National Health and Nutrition Examination Survey (NHANES), USA
- Fruit and vegetables consumed per day and blood pressure measured by an interview that asks respondents to recall everything they ate in two days
- Blood pressure is sum of systolic and diastolic measures
- Fruit and vegetables is the amount consumed per day (g)
- Ages 30–59, 2009–2013. N=7358

Descriptive statistics

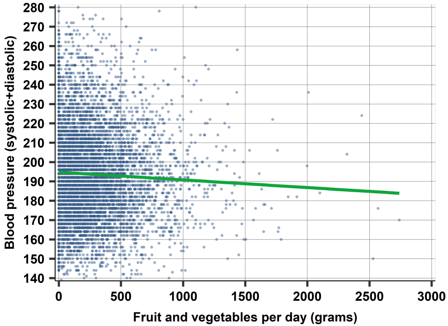
	Mean	Median	Std.Dev.	Min	Max	Obs
Blood pressure	194	192	24	129	300	7359
Fruit/vegetables	361	255	383	0	3153	7359

Source: food-health dataset, USA, ages 30 to 59, 2009–2013.

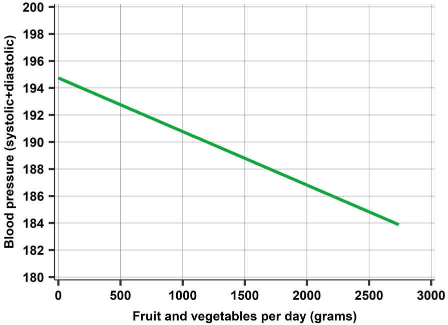
Causal map: relation between fruit-vegetable consumption and blood pressure



Correlations

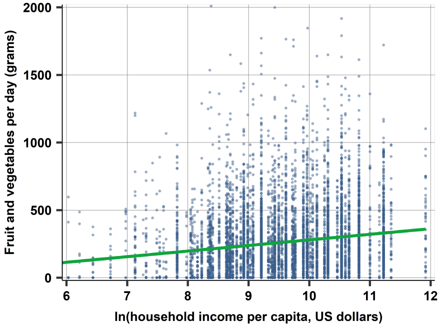


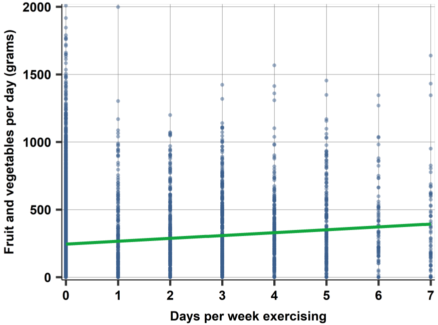
Scatterplot and regression line



Regression line ($\alpha = -0.004$)

Two variables, which affect fruit and vegetables consumption





Log household income

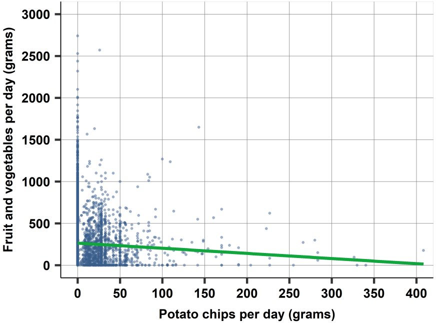
Days/week exercising

Unhealthy food consumption

- Should we control for the consumption of potato chips?

Unhealthy food consumption

- Should we control for the consumption of potato chips?
  - Yes. It is a good proxy for unhealthy living. It affects blood pressure even if we eat a lot of veggies
  - No. Veggie eating causes less chips eating that causes better health. Unwanted mechanism.



Summary

- Food and health correlated
- Many potential confounders
- Never be really causal
- But can offer insight and prompt experiments
- Can be informative - more likely causally true than not

Comparing pros and cons of approaches

- Causality can be established
  - Controlled experiment = great confidence
  - Natural experiment = good confidence, but needs proof
  - Conditioning on confounders = never certain

- This is about **internal validity**
  - How certain we can be that we indeed uncovered a causal relationship

External validity

- However, there is another aspect
- **External validity** is a measure of confidence about generalization
  - Will the causal relationship work in the future
  - Will the causal relationship work in other markets, countries, etc.
  - There is usually a trade-off between internal and external validity

Constructive skepticism

- No analysis is perfect
  - Weigh pros and cons of different approaches
- One can still learn from a well-designed analysis
  - Be that a controlled experiment or an observational study
- Solid knowledge from many studies
  - With different approaches
  - Pointing to similar conclusion if biases well understood
    - Some studies may be more biased than others
    - Need to take into account when summing up evidence from multiple studies