International comparison of labour market policies

Modern Causal Analysis. Rubin Causal Model und Directed Acyclic Graphs

Simon Ress

Ruhr-Universität Bochum

21.07.2021

Content

Chap. 1 Causal Hypotheses

"Simple" causal hypotheses

"Simple" causal hypotheses with two feature:

- Graphic representation of a "simple" hypothesis: causal effect of a feature D on a feature Y

D	Y
Treatment variable	Outcome variable
Explanatory variable	Explained variable
Independent variable	Dependent variable

Criteria for causal inference (D on Y)

1. **Distinct theoretical constructs**: No definitional dependency between D and Y
2. **Appropriate operationalizations and measurement** of both features: features → variables
3. **Theoretical plausibility**: Qualitative explanation (→ mechanism) of the causal effect necessary; reference to empirical studies not sufficient!
4. **Appropriate temporal order**: Theoretical justification needed (empirical order not sufficient, since anticipation effects can occur)
5. **Appropriate temporal distance**: Some effects take time to unfold and some effects weaken over time (theoretical justification needed)
6. **Identification of the causal effect**: E.g. by DAG
7. **Empirical association**: E.g. by regression analysis

Chap. 2 The counterfactual causal model

History and basic idea

History

- First approaches: John Stuart Mill 1843 & Gustav Theodor Fechner 1860)
- Formal concepts of experimental designs in statistics: Neyman 1923 & Fisher 1935
- Formalized causal analysis: Donald B. Rubin 1974, 1977 & 1978

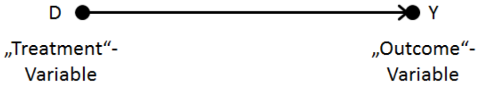
Basic idea

- Focus on the causality concept "Estimating the causal effect of D on Y"
- Based on experimental language: almost any situation can be described in non-experimental context at least as a thought experiment

Other names:

- Potential Outcome Model (POM)
- Rubin Causal Model (RCM)
- Modern Causal Analysis (MCA)

Definitions of treatment and outcomes

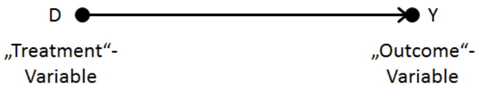


Simplified version of a binary “Treatment”:

- D=1: Treatment ("Experimental group")
- D=0: No Treatment ("Control group")

Note: The binary treatment assumption is a simplistic assumption; There are counterfactual causal models with polytomic treatments (nominal, ordinal, metric).

Central assumption for outcome Y



Each individual **i** can be observed in two potential states (depending on a potential treatment), which means that **two potential outcomes** are conceivable for each person **i**, regardless of the actual treatment status:

- Y_i^0 = potential outcome for person i in the case <u>without</u> the treatment
- Y_i^1 = potential outcome for person i in the case of the treatment

Note: The outcome is generally viewed as metric, but other scale levels can also be assumed.

Fundamental problem of causal inference

Fundamental problem of causal inference (Paul W. Holland, 1986):
$Y_i{}^0$ and $Y_i{}^1$ can never be observed simultaneously for a *single person i*

Table: Observability of various potential outcomes

	$\mathbf{Y_i{}^0}$	$\mathbf{Y_i{}^1}$
$\mathbf{D_i=0}$	Factual (=observable)	Counterfactual (=unobservable)
$\mathbf{D_i=1}$	Counterfactual (=unobservable)	Factual (=observable)

Observation rule: Given the actual state (D), only one potential outcome (Y) can be observed per person

Example of the problem

Table: Effect of (high) education on income

	$\mathbf{Y_i}^0$	$\mathbf{Y_i}^1$
$\mathbf{D_i=0}$	("What a low-educated person actually earns in a state of low-education")	("What a low-educated person would earn if higher education had been achieved")
$\mathbf{D_i=1}$	("What a high-educated person would earn if low education had been achieved")	("What a high-educated person actually earns in a state of high-education")

Interpretation (first line): If the person is actually low-educated ($\rightarrow$ no treatment), their income can be observed as if that person were low-educated. But it can never be observed what income this person would have with a higher education.

Impossibility of individual causal effects

The *individual causal effect* of the treatment on the outcome is defined for each individual as the difference between the two potential outcomes in the treatment and control status (given its actual D):

Individual causal effect of a person i $= Y_{i,D}{}^{1} - Y_{i,D}{}^{0}$

Implications:

- Theoretically the individual causal effect is defined
- In reality the individual causal effect can never be observed, because always one of the both components ($Y_i{}^1$ or $Y_i{}^0$) is contra factual ($\rightarrow$ red in the table) and therefore unobservable

Solution: Average causal effects

As social scientists, we are more interested in *average causal effects* and not in the individual causal effect:

Types of average effects:

- *ATT: Average Treatment Effect on the Treated*
- *ATU: Average Treatment Effect on the Untreated*
- *ATE: Average Treatment Effect*

Advanced notion of potential outcomes

Table: Observability of various potential outcomes

	$\mathbf{Y^0}$	$\mathbf{Y^1}$
$\mathbf{D_i=0}$	$E(Y^0 \mid D=0)$ Factual (=observable)	$E(Y^1 \mid D=0)$ Counterfactual (=unobservable)
$\mathbf{D_i=1}$	$E(Y^0 \mid D=1)$ Counterfactual (=unobservable)	$E(Y^1 \mid D=1)$ Factual (=observable)

E() denotes the expected value = a generalization of the weighted average

Description of the concepts: ATT

The *ATT* is the average causal effect of the treatment on the outcome for people who actually experienced the treatment.

	Y^0	**Y^1**	
D_i=0	$E(Y^0 \mid D=0)$ Factual (=observable)	$E(Y^1 \mid D=0)$ Counterfactual (=unobservable)	
D_i=1	$E(Y^0 \mid D=1)$ Counterfactual (=unobservable)	$E(Y^1 \mid D=1)$ Factual (=observable)	**ATT**

$$ATT = \underbrace{E(Y_i{}^1 \mid D_i = 1)}_{\text{observable}} - \underbrace{E(Y_i{}^0 \mid D_i = 1)}_{\text{unobservable}} \tag{1}$$

Research designs differ in how the unobservable/counterfactual situation is simulated / estimated.

Description of the concepts: ATU/ATNT

The *ATNT* is the average causal effect of the treatment on the outcome for people who didn't experienced the treatment.

	$\mathbf{Y^0}$	$\mathbf{Y^1}$	
$\mathbf{D_i=0}$	$E(Y^0 \mid D=0)$ Factual (=observable)	$E(Y^1 \mid D=0)$ Counterfactual (=unobservable)	**ATNT**
$\mathbf{D_i=1}$	$E(Y^0 \mid D=1)$ Counterfactual (=unobservable)	$E(Y^1 \mid D=1)$ Factual (=observable)	

$$ATNT = \underbrace{E(Y_i{}^1 \mid D_i = 0)}_{\text{unobservable}} - \underbrace{E(Y_i{}^0 \mid D_i = 0)}_{\text{observable}} \tag{2}$$

Research designs differ in how the unobservable/counterfactual situation is simulated / estimated.

Description of the concepts: ATE

The *ATE* is the average causal effect of the treatment on the outcome for all persons.

	$\mathbf{Y^0}$	$\mathbf{Y^1}$
$\mathbf{D_i=0}$	$E(Y^0 \mid D=0)$ Factual (=observable)	$E(Y^1 \mid D=0)$ Counterfactual (=unobservable)
$\mathbf{D_i=1}$	$E(Y^0 \mid D=1)$ Counterfactual (=unobservable)	$E(Y^1 \mid D=1)$ Factual (=observable)

ATNT x (1-π)

+

ATT x π

=

ATE

$$
\begin{aligned}
ATE &= \pi \times ATT + (1-\pi) \times ATNT \\
&= \pi \times (\underbrace{E(Y_i^1 \mid D_i = 1)}_{\text{observable}} - \underbrace{E(Y_i^0 \mid D_i = 1)}_{\text{unobservable}}) \\
&+ (1-\pi) \times \underbrace{E(Y_i^1 \mid D_i = 0)}_{\text{unobservable}} - \underbrace{E(Y_i^0 \mid D_i = 0)}_{\text{observable}}
\end{aligned}
$$

Example

The causal effect of higher education on income

- **ATT** = Average causal effect of higher education on income for more high-educated people
- **ATNT** = Average causal effect of higher education on income for low-educated people
- **ATE** = Average causal effect of higher education on income *(across all groups)*

Chap. 3 The naive estimator

Second Problem: Counterfactual states

Data is required for the empirical calculation of the effects. However, because some potential outcomes are counterfactual, it is not possible to calculate any of the average causal effects directly.

Nevertheless, we can only work with the existing data

	Y^0	Y^1	
$D_i=0$	$E(Y^0 \mid D=0)$ Factual (=observable)	$E(Y^1 \mid D=0)$ Counterfactual (=unobservable)	
$D_i=1$	$E(Y^0 \mid D=1)$ Counterfactual (=unobservable)	$E(Y^1 \mid D=1)$ Factual (=observable)	ATT

	Y^0	Y^1	
$D_i=0$	$E(Y^0 \mid D=0)$ Factual (=observable)	$E(Y^1 \mid D=0)$ Counterfactual (=unobservable)	ATNT
$D_i=1$	$E(Y^0 \mid D=1)$ Counterfactual (=unobservable)	$E(Y^1 \mid D=1)$ Factual (=observable)	

	Y^0	Y^1	
$D_i=0$	$E(Y^0 \mid D=0)$ Factual (=observable)	$E(Y^1 \mid D=0)$ Counterfactual (=unobservable)	ATNT $\times (1-\pi)$ +
$D_i=1$	$E(Y^0 \mid D=1)$ Counterfactual (=unobservable)	$E(Y^1 \mid D=1)$ Factual (=observable)	ATT $\times \pi$ = ATE

The naive estimator

Definition

The *naive estimator (NaivE)* is the difference between the factual (observable) outcomes

$$NaivE = \underbrace{E(Y_i^{1}|D_i = 1)}_{\text{observable}} - \underbrace{E(Y_i^{0}|D_i = 0)}_{\text{observable}}$$

	$\mathbf{Y^0}$	$\mathbf{Y^1}$
$\mathbf{D_i=0}$	$E(Y^0\|D=0)$ Factual (=observable)	$E(Y^1\|D=0)$ Counterfactual (=unobservable)
$\mathbf{D_i=1}$	$E(Y^0\|D=1)$ Counterfactual (=unobservable)	$E(Y^1\|D=1)$ Factual (=observable)

Naive estimator

The naive estimator: Example

Table: Effect of high education on income

	$\mathbf{Y_i}^{\mathbf{0}}$	$\mathbf{Y_i}^{\mathbf{1}}$
$\mathbf{D_i=0}$	10 €/h	20 €/h
$\mathbf{D_i=1}$	15 €/h	30 €/h

Calculation of the effect of high education on income:

- Naive estimator = 30 €/h - 10 €/h = **20 €/h**
- ATT = 30 €/h - 15 €/h = **15 €/h**
- ATNT = 20 €/h - 10 €/h = **10 €/h**
- ATE = 15 €/h$\times\pi$ + 10 €/h $\times(1-\pi)$ = **12.5 €/h**

Bias of the naive estimator

- In most of the real world scenarios, the naive estimator does not provide the true value of any causal effect. This means the result of the Naive Estimator is *biased*!
- This distortion results from the so-called **baseline difference**. This is the difference between the technically required but counterfactual potential outcome and the observable potential outcome used

Baseline difference regarding the **ATT**

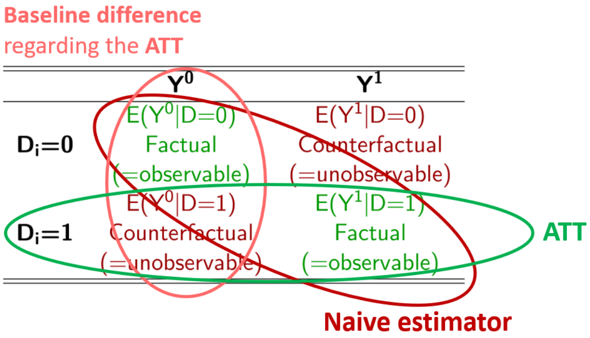
Baseline difference regarding the **ATNT**

	$\mathbf{Y^0}$	$\mathbf{Y^1}$
$\mathbf{D_i=0}$	$E(Y^0 \mid D=0)$ Factual (=observable)	$E(Y^1 \mid D=0)$ Counterfactual (=unobservable)
$\mathbf{D_i=1}$	$E(Y^0 \mid D=1)$ Counterfactual (=unobservable)	$E(Y^1 \mid D=1)$ Factual (=observable)

Bias decomposition: NaivE-ATT I

Bias decomposition of the naive estimator (regarding ATT):

$$\begin{aligned} NaivE &= E(Y_i{}^1|D_i=1) - E(Y_i{}^0|D_i=0) \\ &= ATT + \underbrace{(\underbrace{E(Y_i{}^0|D_i=1)}_{\text{unobservable}} - E(Y_i{}^0|D_i=0))}_{\text{baseline difference}} \end{aligned}$$



Bias decomposition: NaivE-ATT II

Calculation of the ATT from the naive estimator and the baseline difference.

Changing the formula of NaivE to ATT:

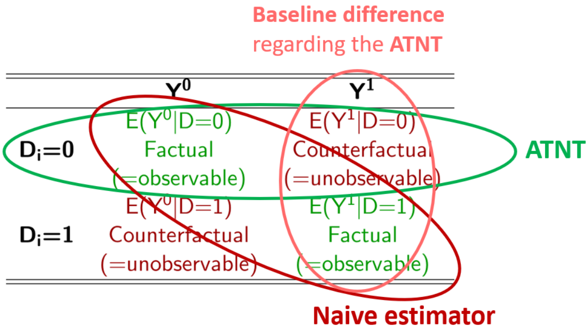
$$ATT = NaivE - \underbrace{(\underbrace{E(Y_i{}^0|D_i = 1)}_{\text{unobservable}} - E(Y_i{}^0|D_i = 0))}_{\text{baseline difference}}$$

Application:
In the example the baseline difference is (15 Euro/h - 10 Euro/h =) **5 Euro/h**. Naive estimator minus blaseline difference (20 Euro/h - 5 Euro/h) equals the true value of the ATT (15 Euro/h)

Bias decomposition: NaivE-ATNT I

Bias decomposition of the naive estimator (regarding ATNT):

$$\begin{aligned} NaivE &= E(Y_i{}^1|D_i = 1) - E(Y_i{}^0|D_i = 0) \\ &= ATNT + \underbrace{(E(Y_i{}^1|D_i = 1) - \underbrace{E(Y_i{}^1|D_i = 0)}_{\text{unobservable}})}_{\text{baseline difference}} \end{aligned}$$



Bias decomposition: NaivE-ATNT II

Calculation of the ATT from the naive estimator and the baseline difference.

Changing the formula of NaivE to ATNT:

$$ATNT = NaivE - \underbrace{(E(Y_i{}^1|D_i = 1) - \underbrace{E(Y_i{}^1|D_i = 0)}_{\text{unobservable}})}_{\text{baseline difference}}$$

Application:
In the example the baseline difference is (30 Euro/h - 20 Euro/h =) **10 Euro/h**. Naive estimator minus blaseline difference (20 Euro/h - 10 Euro/h) equals the true value of the ATNT (10 Euro/h)

Violation of the ignorability assumption

Reason for the bias of the naive estimator
→ Violation of the *Ignorability Assumption (IA)*

IA: $(Y^0, Y^1) \perp\!\!\!\perp D$
$\Rightarrow E(Y^0|D{=}0) = E(Y^0|D{=}1) = E(Y^0)$
$\Rightarrow E(Y^1|D{=}0) = E(Y^1|D{=}1) = E(Y^1)$

- Assignment of the treatment status D is independent of the potential outcome
- Caution: The ignorability assumption is often not fulfilled (due to a lack of randomization), since features in the background influence both individuals assignment of the treatment group and their potential outcomes → distortion of the naive estimator

Ignorability assumption

Implication of the ignorability assumption

- If it applies, the potential outcomes are independent of treatment status D.

Thus applies:
$E(Y^0|D{=}1) = E(Y^0|D{=}0)$ and
$E(Y^1|D{=}1) = E(Y^1|D{=}0)$

- From this, in turn, it follows that the baseline difference (the bias of NaiveS) is: $(E[Y_i{}^0|D_i{=}1] - E[Y_i{}^0|D_i{=}0] =)\ 0$
- *The NaiveS thus results in the true average causal effect of interest (ATT,ATNT or ATE)*

Ignorability assumption

Under which conditions is the ignorability assumption valid?

1. Treatment status D is in fact independent of the potential outcomes
2. Treatment status D is after conditioning a set of variables Z independent of the potential outcomes → *Conditional Ignorability Assumption (CIA)*

How can these conditions be verified?

1. With *Directed Acyclic Graphs (DAGs)* (1) and (2) can be checked
2. For (2), the required set of variables Z can be derived using DAGs

Chap. 4 Directed Acyclic Graphs (DAGs)

Directed Acyclic Graphs (DAGs)

Pearl 1995: Causal diagrams for empirical research

- Graphical tool to represent qualitative assumptions about causal relationships
- It is possible to derive causal conclusions from these assumptions and data

Notation

- Each point of the DAG represents a features/variable whose name is given
- Filled points represent *observed* features/variables
- Unfilled points represent *unobserved* features/variables
- One-sided arrows postulate a directed causal effect *(theoretically or based on empirical causal analysis)*

Directed Acyclic Graphs (DAGs)

Definitions

- **Path:** Every (uninterrupted) connection between two points, *regardless of the direction of the arrow*
- **Causal Path:** Every (uninterrupted) connection between two points, *with clear arrow direction*
- **Directed Graph:** Every connection between two variables is a one-sided arrow
- **Acyclic Graph:** There are no loops of causal paths
- **Back-door Path:** A path that connects D and Y and in which an arrowhead points to D.
- **Collider:** A variable pointed to by two arrows
- **(Un-)Blocked Path:** A path is *blocked* if there is at least one collider on it, otherwise this is *unblocked*.
- **Blocking Paths:** A path can be blocked by a variable on this path that is not a collider

Directed Acyclic Graphs (DAGs)

Back-door criterion

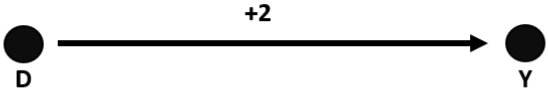
A set of variables Z fulfills the *back-door criterium* if:

- No variable in Z is causally influenced by D.
- By Z every *unblocked path* between D and Y becomes blocked

If the back-door criterion is met, the CIA is satisfied and a causal effect identified

DAG Examples I: Overview

No Confounding

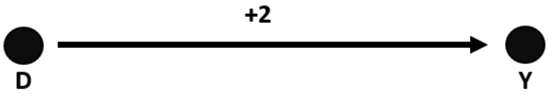


Question:

1. What is the causal path of interest?
2. What is/are the back-door path(s)? Are they open? How to block them?

→ Is the result of the naive estimator the true causal effect?

DAG Examples I: Paths



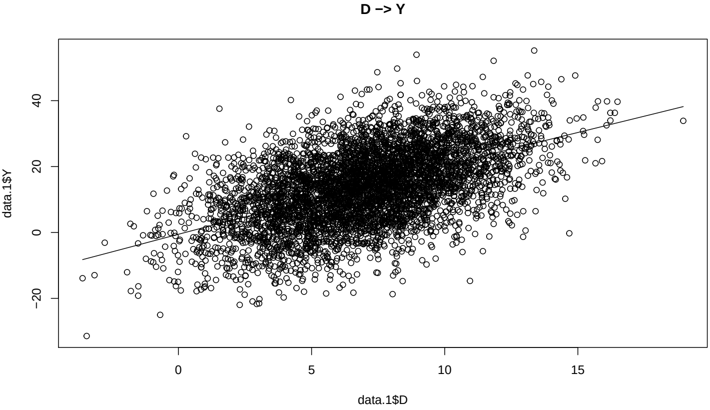
Paths:

- Causal path: D → Y

CIA satisfaction:

- Automatically satisfied because there are no back-door path

DAG Examples I: Scatterplot



DAG Examples I: Regression result

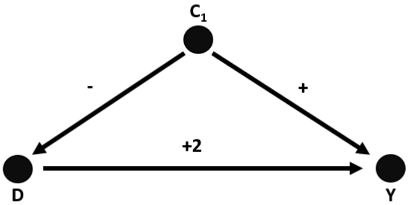
	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	-0.4878	0.3583	-1.36	0.1735
D	2.0578	0.0472	43.58	0.0000

Table: Regression: lm(Y ~ D)

	2.5 %	97.5 %
D	1.97	2.15

Table: 95-Percent confidence interval

DAG Examples II: (Partially) Covered

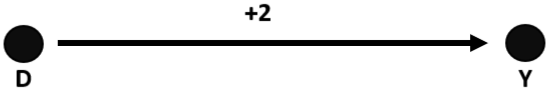


Question:

1. What is the causal path of interest?
2. What is/are the back-door path(s)? Are they open? How to block them?

→ Is the result of the naive estimator the true causal effect?

DAG Examples II: Paths



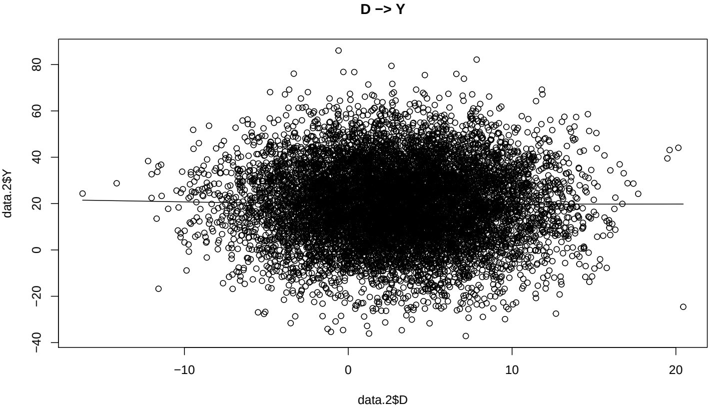
Paths:

- Causal path: D → Y
- Back-door path: D ← C → Y

CIA satisfaction:

- $Z1 = \{C\}$

DAG Examples II: Scatterplot



DAG Examples II: Regression result, NaivE

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	19.9093	0.1995	99.80	0.0000
D	-0.0443	0.0366	-1.21	0.2256

Table: Regression: lm(Y ˜ D)

	2.5 %	97.5 %
D	-0.12	0.03

Table: 95-Percent confidence interval: lm(Y ˜ D)

DAG Examples II: Regression result, CIA

Usage of Z1 to satisfy CIA

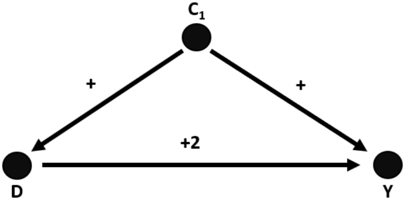
	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	-0.0086	0.4110	-0.02	0.9834
D	1.9542	0.0493	39.65	0.0000
C	3.4925	0.0651	53.62	0.0000

Table: Regression: lm(Y ~ D + C1)

	2.5 %	97.5 %
D	1.86	2.05

Table: 95-Percent confidence interval: lm(Y ~ D + C1)

DAG Examples III: Oversized

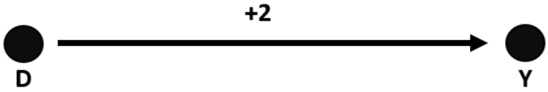


Question:

1. What is the causal path of interest?
2. What is/are the back-door path(s)? Are they open? How to block them?

→ Is the result of the naive estimator the true causal effect?

DAG Examples III: Paths



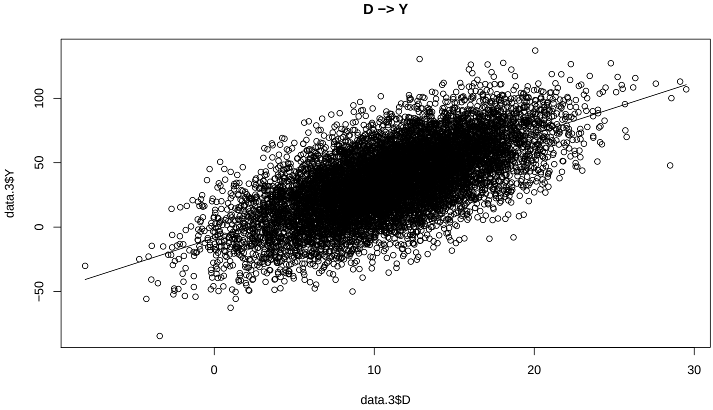
Paths:

- Causal path: $D \rightarrow Y$
- Back-door path: $D \leftarrow C \rightarrow Y$

CIA satisfaction:

- $Z1 = \{C\}$

DAG Examples III: Scatterplot



DAG Examples III: Regression result, NaivE

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	-8.4145	0.5506	-15.28	0.0000
D	4.0103	0.0461	87.01	0.0000

Table: Regression: lm(Y ~ D)

	2.5 %	97.5 %
D	3.92	4.10

Table: 95-Percent confidence interval: lm(Y ~ D)

DAG Examples III: Regression result, CIA

Usage of Z1 to satisfy CIA

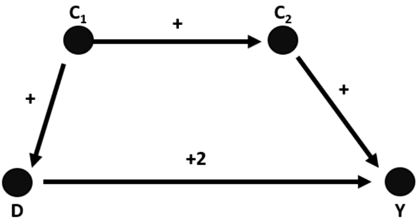
	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	-0.3580	0.5460	-0.66	0.5121
D	1.9886	0.0652	30.51	0.0000
C	3.5625	0.0868	41.02	0.0000

Table: Regression: lm(Y ˜ D + C1)

	2.5 %	97.5 %
D	1.86	2.12

Table: 95-Percent confidence interval: lm(Y ˜ D + C1)

DAG Examples IV

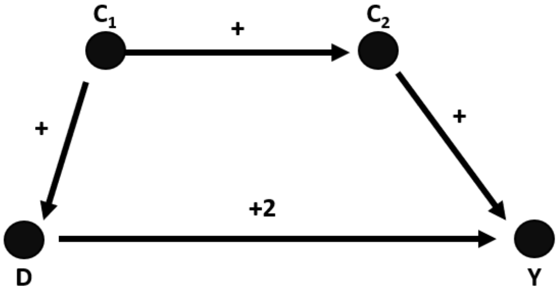


Question:

1. What is the causal path of interest?
2. What is/are the back-door path(s)? Are they open? How to block them?

→ Is the result of the naive estimator the true causal effect?

DAG Examples IV: Paths



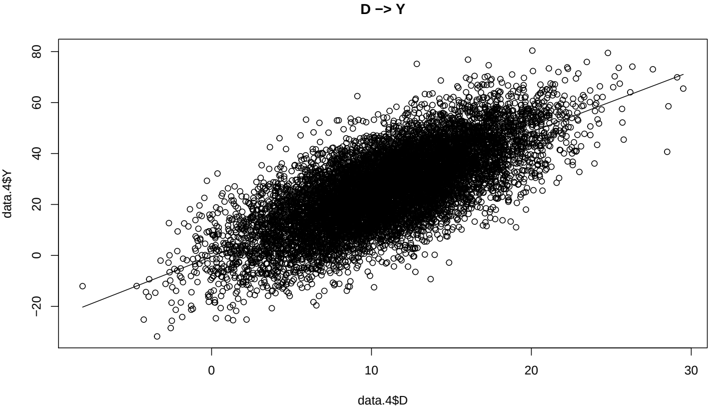
Paths:

- Causal path: $D \rightarrow Y$
- Back-door path: $D \leftarrow C_1 \rightarrow C_2 \rightarrow Y$

CIA satisfaction:

- $Z1 = \{C_1\}$ or
- $Z2 = \{C_2\}$

DAG Examples IV: Scatterplot



DAG Examples IV: Regression result, NaivE

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	-0.2863	0.2769	-1.03	0.3011
D	2.4241	0.0232	104.57	0.0000

Table: Regression: lm(Y ~ D)

	2.5 %	97.5 %
D	2.38	2.47

Table: 95-Percent confidence interval: lm(Y ~ D)

DAG Examples IV: Regression result, CIA I

Usage of Z1 to satisfy CIA

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	1.5029	0.2926	5.14	0.0000
D	1.9751	0.0349	56.54	0.0000
C1	0.7912	0.0465	17.00	0.0000

Table: Regression: lm(Y ˜ D + C1)

	2.5 %	97.5 %
D	1.91	2.04

Table: 95-Percent confidence interval: lm(Y ˜ D + C1)

DAG Examples IV: Regression result, CIA II

Usage of Z2 to satisfy CIA

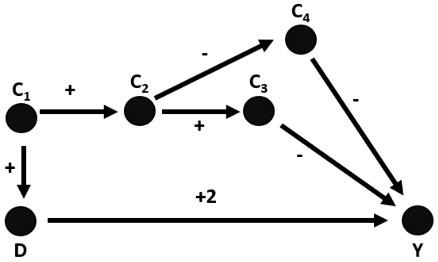
	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	-0.2489	0.2547	-0.98	0.3285
D	2.0035	0.0235	85.29	0.0000
C2	1.5313	0.0359	42.68	0.0000

Table: Regression: lm(Y ~ D + C2)

	2.5 %	97.5 %
D	1.96	2.05

Table: 95-Percent confidence interval: lm(Y ~ D + C2)

DAG Examples V: (Partially) Covered

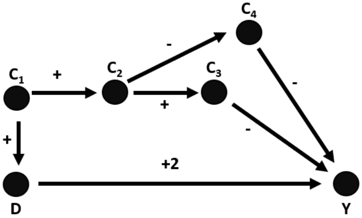


Question:

1. What is the causal path of interest?
2. What is/are the back-door path(s)? Are they open? How to block them?

→ Is the result of the naive estimator the true causal effect?

DAG Examples V: Paths



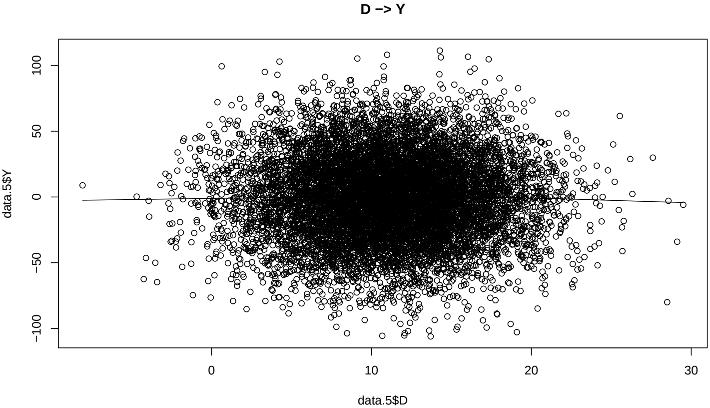
Paths:

- Causal path: $D \rightarrow Y$
- 1. Back-door path: $D \leftarrow C_1 \rightarrow C_2 \rightarrow C_3 \rightarrow Y$
- 2. Back-door path: $D \leftarrow C_1 \rightarrow C_2 \rightarrow C_4 \rightarrow Y$

CIA satisfaction:

- $Z1 = \{C_1\}$ or
- $Z2 = \{C_2\}$ or
- $Z3 = \{C_3, C_4\}$

DAG Examples V: Scatterplot



DAG Examples V: Regression result, NaivE

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	0.0195	0.8189	0.02	0.9810
D	-0.0488	0.0686	-0.71	0.4769

Table: Regression: lm(Y ~ D)

	2.5 %	97.5 %
D	-0.18	0.09

Table: 95-Percent confidence interval: lm(Y ~ D)

DAG Examples V: Regression result, CIA I

Usage of Z1 to satisfy CIA

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	-8.3448	0.8464	-9.86	0.0000
D	2.0501	0.1010	20.29	0.0000
C1	-3.6986	0.1346	-27.47	0.0000

Table: Regression: lm(Y ~ D + C1)

	2.5 %	97.5 %
D	1.85	2.25

Table: 95-Percent confidence interval: lm(Y ~ D + C1)

DAG Examples V: Regression result, CIA II

Usage of Z2 to satisfy CIA

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	-5.5728	0.7907	-7.05	0.0000
D	2.0116	0.0880	22.85	0.0000
C2	-2.4464	0.0708	-34.58	0.0000

Table: Regression: lm(Y ˜ D + C2)

	2.5 %	97.5 %
D	1.84	2.18

Table: 95-Percent confidence interval: lm(Y ˜ D + C2)

DAG Examples V: Regression result, CIA III

Usage of Z3 to satisfy CIA

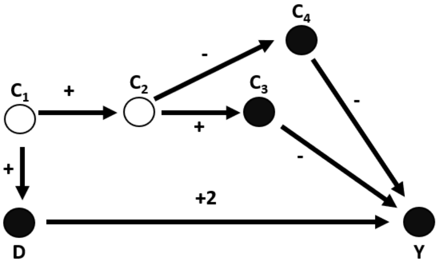
	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	-0.7023	0.8021	-0.88	0.3813
D	1.9977	0.0862	23.18	0.0000
C3	-1.4751	0.0385	-38.33	0.0000
C4	-0.9925	0.0732	-13.56	0.0000

Table: Regression: lm(Y ˜ D + C3 + C4)

	2.5 %	97.5 %
D	1.83	2.17

Table: 95-Percent confidence interval: lm(Y ˜ D + C3 + C4)

DAG Examples VI

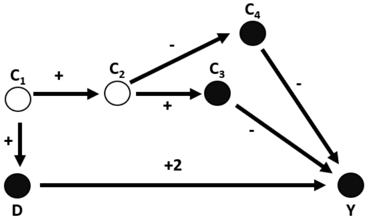


Question:

1. What is the causal path of interest?
2. What is/are the back-door path(s)? Are they open? How to block them?

→ Is the result of the naive estimator the true causal effect?

DAG Examples VI: Paths



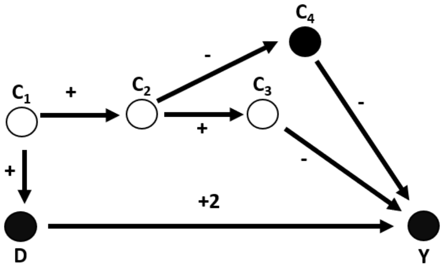
Paths:

- Causal path: D $\rightarrow$ Y
- 1. Back-door path: D $\leftarrow C_1 \rightarrow C_2 \rightarrow C_3 \rightarrow$ Y
- 2. Back-door path: D $\leftarrow C_1 \rightarrow C_2 \rightarrow C_4 \rightarrow$ Y

CIA satisfaction:

- Only Z3 $= \{C_3, C_4\}$ can be used because it is observed

DAG Examples VII

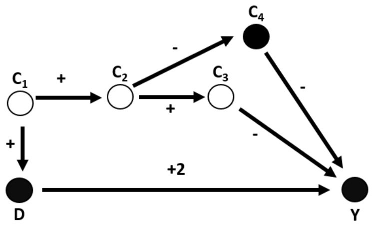


Question:

1. What is the causal path of interest?
2. What is/are the back-door path(s)? Are they open? How to block them?

→ Is the result of the naive estimator the true causal effect?

DAG Examples VII: Paths



Paths:

- Causal path: D → Y
- 1. Back-door path: D ← C_1 → C_2 → C_3 → Y
- 2. Back-door path: D ← C_1 → C_2 → C_4 → Y

CIA satisfaction:

- Not all back-door paths can be closed
- CIA is only partially met by $Z4 = \{C_4\}$
- The estimate will contain a bias despite the closing of one back door path
- The bias is induced by the back-door path that is still open

DAG Examples VII: Regression result, NaivE

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	0.0195	0.8189	0.02	0.9810
D	-0.0488	0.0686	-0.71	0.4769

Table: Regression: lm(Y ˜ D)

	2.5 %	97.5 %
D	-0.18	0.09

Table: 95-Percent confidence interval: lm(Y ˜ D)

DAG Examples VII: Regression result, CIA I

Usage of Z4 to partially satisfy CIA

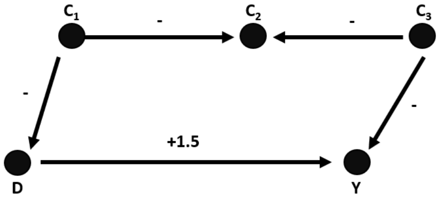
	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	-0.9820	0.8590	-1.14	0.2529
D	0.0676	0.0749	0.90	0.3666
C4	0.2687	0.0700	3.84	0.0001

Table: Regression: lm(Y ~ D + C3)

	2.5 %	97.5 %
D	-0.08	0.21

Table: 95-Percent confidence interval: lm(Y ~ D + C4)

DAG Examples VIII

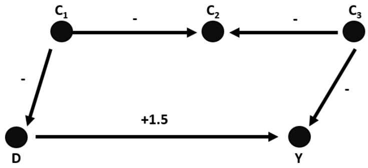


Question:

1. What is the causal path of interest?
2. What is/are the back-door path(s)? Are they open? How to block them?

→ Is the result of the naive estimator the true causal effect?

DAG Examples VIII: Paths



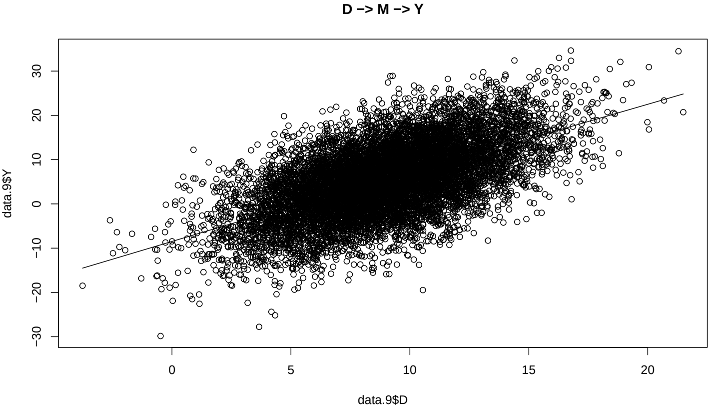
Paths:

- Causal path: D $\rightarrow$ Y
- 1. Back-door path: D $\leftarrow C_1 \rightarrow C_2 \leftarrow C_3 \rightarrow$ Y

CIA satisfaction:

- Automatically satisfied because of the collier on back-door path

DAG Examples VIII: Scatterplot



DAG Examples VIII: Regression result, NaivE

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	-8.2163	0.1962	-41.88	0.0000
D	1.5197	0.0207	73.57	0.0000

Table: Regression: lm(Y ˜ D)

	2.5 %	97.5 %
D	1.48	1.56

Table: 95-Percent confidence interval: lm(Y ˜ D)

DAG Examples VIII: Regression result, wrong control

Usage of Z1 to satisfy CIA

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	-6.2713	0.1713	-36.62	0.0000
D	1.8519	0.0185	99.84	0.0000
C2	0.8255	0.0137	60.09	0.0000

Table: Regression: lm(Y ˜ D + C1)

	2.5 %	97.5 %
D	1.82	1.89

Table: 95-Percent confidence interval: lm(Y ˜ D + C2)

Chap. 5 Experimental vs. non-experimental designs

Experimental designs

- The central feature of **experimental designs** is the random distribution of the observations (e.g. test subjects) between a treatment group and a control group (*randomization*)
- Due to the randomization, these are often seen as an ideal design for drawing causal inferences because the **IA is automatically satisfied**
- **Problem:** Often experimental designs cannot be implemented due to *practical and ethical problems*
- In addition, experiments can have problems *internal and external validity*

Non-experimental designs

- Frequent use in the social sciences of non-experimental methods of causal analysis based on observational data
- Data sources are e.g. general population surveys such as ALLBUS, SOEP or ESS
- **Problems:**
  - Non-random selection
  - CIA/IA needs to be satisfied

(Temporal) dataset structure I

Cross-sectional data

- No time dimension in the outcome variable (Y): measurement only once at time t1
- Treatment variable (D) is also recorded at the same point in time t1 ...
  - (A) ... but based on theoretical reasons it can be assumed to take place earlier, even if the measurement relates to time t1. (*Example: D:father's level of education, Y:respondent's income*)
  - (B) ... but can, as a retrospective question, relate to an earlier point in time t0 (or period)

(Temporal) dataset structure II

Longitudinal data

- At least two measurements of the Y variable in time to capture changes in Y.
- Additional condition for the implementation of a "fixed-effect" logic:
  At least one measurement of the Y variable must relate to a point in time before the treatment (= change over time in the D variable of interest) and at least one measurement of the Y variable must relate to a point in time after the treatment

Task for the next meeting

- Search for possible confounders of the chosen connections.
- Include this in the already created document.
- Inclusion of the confounder(s) in the regression analysis as control variable(s).

Do you still have any questions?