



## Dagens program

14:00 – 15:20 Godt at vide + RCT

15:20 – 15:30 Pause

15:30 – 16:00 Klassiske kvasiekperimentelle metoder

16:00 – 17:00 Cases

- Formålet er at opnå øvelse i at forstå forskellige designs fordele og ulemper i forskellig kontekst.

## Dagens program

- Stil gerne spørgsmål
- Litteratur
- Præsentation udleveres

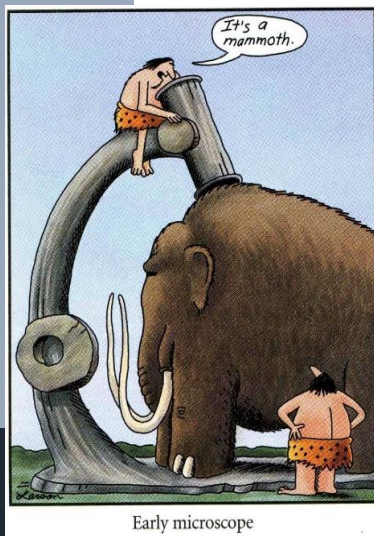
# 1. Godt at vide forud for at beskæftige sig med effekter



# 1. Godt at vide forud for at beskæftige sig med effekter

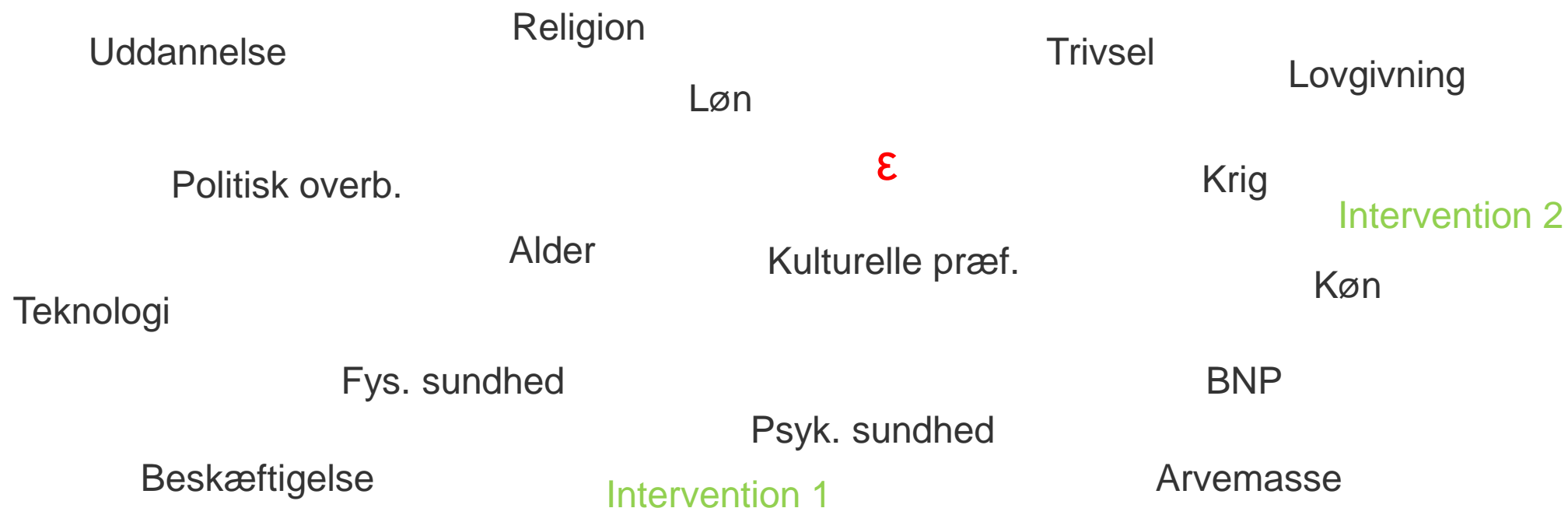
- Videnskabsteoretiske standpunkter ifht. kausalitet:
  - Instrumentalisme/positivisme: kausale effekter vs pseudovidenskab
  - Kritisk rationalisme: falsifikation
  - Socialkonstruktivisme: (social) verden = konstruktion, kausalitet ikke observerbar
  - Kritisk realisme: mekanismer, transitive/intransitive

Diskussion: Hvilken teori er den stærkeste?



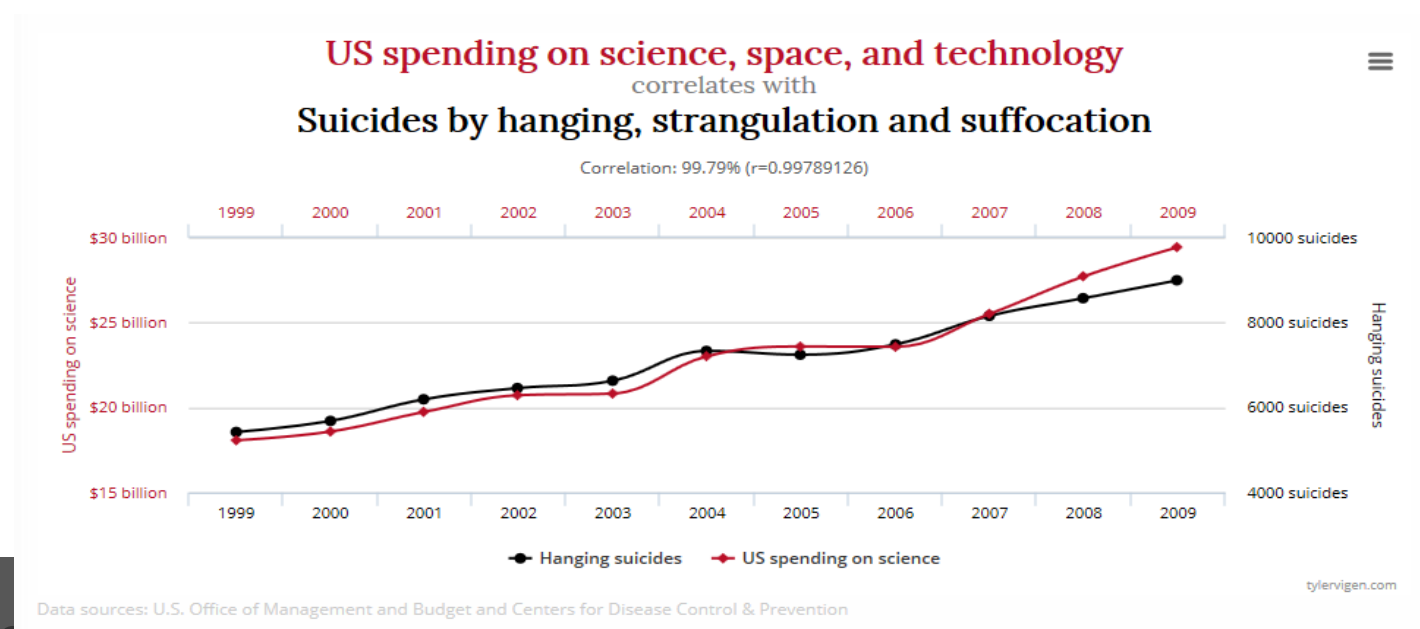
# 1. Godt at vide forud for at beskæftige sig med effekter

- Samfundets myriade af kausaliteter: Øvelse: tegn pile. Hvilke variable kan der være afhængigheder mellem?



# 1. Godt at vide forud for at beskæftige sig med effekter

- Samvariation, spuriøse sammenhænge
- Den naive estimator
- Storke og fødsler (teoretisk sammenhæng, teoriens styrke)





# 1. Godt at vide forud for at beskæftige sig med effekter





# 1. Godt at vide forud for at beskæftige sig med effekter

- Den kontrafaktiske tilstand: behov for flere tidsmaskiner!
- Effektmåling går ud på at simulere den kontrafaktiske tilstand
- Test- og kontrolgruppe
- Effekten er altid et udfald relativt til noget andet.

# 1. Godt at vide forud for at beskæftige sig med effekter

- Effektmåling handler om middelværdier for grupper, ikke individ-effekter
- Subpopulationer kan gemme på stor variation

# 1. Godt at vide forud for at beskæftige sig med effekter

En ufuldstændig liste over kilder til bias:

- Placebo
- Hawthorne
- Naturlig progression/maturation (tid/alder), spontan remission (ex. forkølelse)
- Regression mod middelværdi
- Kontamination (Stable Unit Treatment Variation Assumption)
- Frafald



# 1. Godt at vide forud for at beskæftige sig med effekter

En ufuldstændig liste over kilder til bias, II:

- Lav validitet ved outcome-målet
- Lav reliabilitet ved outcome-målet kombineret med lille n
- Lav sensibilitet ved outcome-målet
- Målefejl
- Analysefejl
- Tidsforskydelse ml test- og kontrol
- Type 3-fejl

# 1. Godt at vide forud for at beskæftige sig med effekter

En ufuldstændig liste over kilder til bias, III:

- Selektion, selektion og atter selektion!
  - Systematisk støj: **problem**
  - Usystematisk støj: **ok ved højt N**
  - Confounders (Hubert M. Blalock)

# 1. Godt at vide forud for at beskæftige sig med effekter

Hvad er pointen med at nævne alle disse former for bias?



# 1. Godt at vide forud for at beskæftige sig med effekter

- Simple vs komplekse indsatser
  - Simple: Oversvømmelse, justering af skat, fartgrænse
  - Komplekse: Kognitiv terapi, psykoterapi, uddannelse, ændret sociallovgivning
- Måling med simple vs komplekse outcomes
  - Simple: En persons vægt, dødsfald, domfældelse
  - Komplekse: Latente faktorer (adfærd, kognition, psyke, læring, trivsel), samfundsforandringer
- Kan indsatser og outcomes blive så komplekse, at effekten er umulig at måle?



## 2. Det randomiserede kontrollerede forsøg

## 2. Det randomiserede kontrollerede forsøg

- Treatment/allokering er (sand) tilfældig, aka sand eksogen



## 2. Det randomiserede kontrollerede forsøg

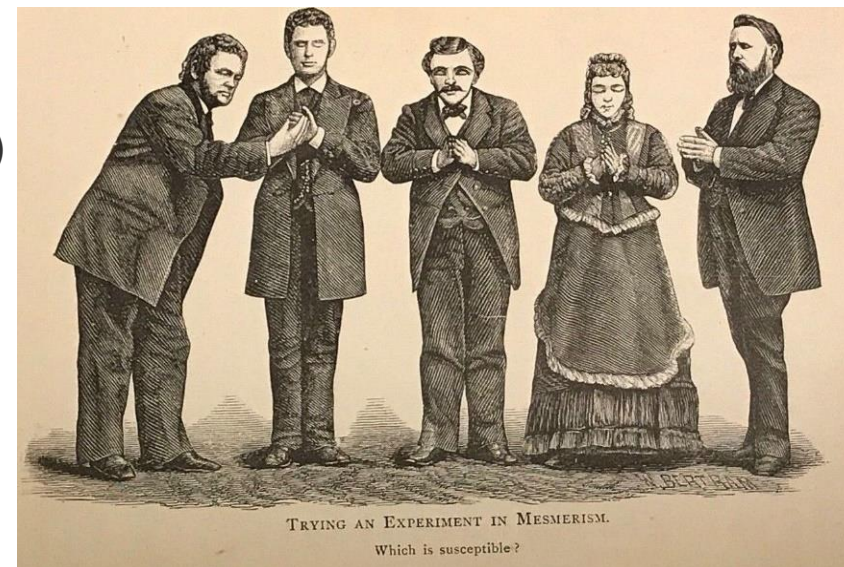
### Etik

- Er det etisk forsvarligt at eksperimentere med mennesker vs etisk forsvarligt at lade være?
- Stanford prison experiment
- Milgram experiments
- Diskussion: Hvor går grænsen?



## 2. Det randomiserede kontrollerede forsøg

- RCT (Ronald A. Fisher):
  - Average Treatment Effect (ATE)
    - Den sande effekt (intet bortfald)
  - Intention to Treat (ITT)
    - Effekten af at være tildelt allokering (i tilfælde af bortfald)



## 2. Det randomiserede kontrollerede forsøg

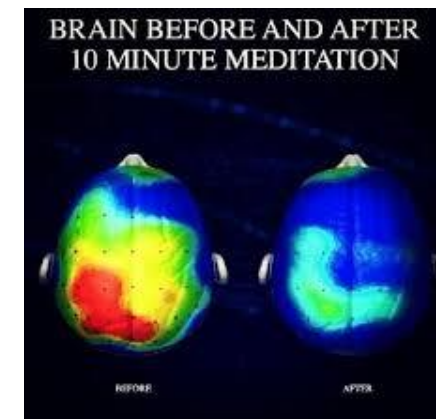
- |  |  |
|--|--|
| <ul style="list-style-type: none"><li>• Designs:<ul style="list-style-type: none"><li>• Simpel lodtrækning</li><li>• Venteliste</li><li>• Løbende rekruttering</li><li>• Cross-over</li><li>• Cluster/multisite</li><li>• Factorial design</li><li>• Effekt vs effectiveness</li></ul></li></ul> | <ul style="list-style-type: none"><li>• Blok/cluster-randomisering</li><li>• Stratificering</li><li>• Efficacy (TAU)</li><li>• Stepped wedge</li><li>• “Biased coin” diskutér om dette design giver bias</li><li>• Blindet/åbent</li></ul> |
|--|--|

## 2. Det randomiserede kontrollerede forsøg

### Afsluttende diskussioner

- Hvordan analyserer man data, hvis nogen falder fra?
- Hvad gør man, hvis man ikke kan identificere en valideret skala, som kan måle det relevante outcome?
- Hvad gør man, hvis man ikke kan opnå deltagere nok?
- Hvad gør man, hvis praktikere ikke bakker op?
- Hvad gør man, hvis det bliver for dyrt at gennemføre?

**Pause til kl 15:30**





## 3. Kvasi-eksperimentelle studier

### 3. Kvasi-eksperimentelle studier

- Tilfældig allokering er ikke foretaget men antaget
- → Efter publicering kan det senere vise sig, at antagelsen om tilfældig allokering var problematisk
- Diskussion: Hvordan ruster man sig mod denne problematik?

### 3. Kvasi-eksperimentelle studier

- Selektion, confounders
- Proxy

### 3. Kvasi-eksperimentelle studier

- Eksempler:
  - Naturlige eksperimenter (Trygve Haavelmo)
  - Instrumentmetode (LATE) (Philip G. Wright)
  - Difference-in-differences (MTE) (John Snow)
  - Multilevel (Larzarsfeld & Menzel)
  - Tvillingestudier (Francis Galton)
  - Regression discontinuity (Thistlewaite & Campbell)

### 3. Kvasi-eksperimentelle studier

- Det generelle udgangspunkt for kvasi-eksp studier:
  - Samfundet begrænser hvad vi kan eksperimentere med
  - Derfor må vi finde alternativer til setup med rent eksogen treatment

### 3. Kvasi-eksp

Aspekt	RCT	Kvasi
<u>Design</u>		
Hvad afgrænser emner for måling?	Etik	Tilfældigheder og tilgængelige data
Population/stikprøve	Lille	Stor
Estimator	ATE/ITT	TOT, LATE, MTE m.m.
Random allokering	Ja	Antages
Data	Typisk primært survey	Ofte register, evt. paneldata
Bias fra placebo-effekt	Muligt uden blinding	Muligt
Hawthorne-effekt	Muligt uden blinding	Nej
<u>Konsekvenser</u>		
Intern validitet	Høj	Middel/høj
Ekstern validitet	Lav	Høj
Afspejler virkeligheden?	Bedste bud v factorial designs og reviews	Høj ssh ved stærkt design
Compliance	Afhænger af fokus	-
Dataindsamling	prospektiv	retrospektiv
Kan estimere sand kausal fixed effekt	Ja	Ja
Kan estimere sand kausal random effekt	Ved review / factorial design	Nej
Analytisk kompleksitet	Afhængigt af design (psykometri, styrketest)	Statistisk kompleksitet
Gennemførelsestid	Ofte høj	Lav
Pris	Høj	Lav
Egner sig til komplekse interventioner	Ja	Nej
Egner sig til estimering af meget små effekter	Nej	Ja

## 3.1 Naturlige eksperimenter

- Samme princip som RCT, men hvor "naturen" (dvs. ikke researcheren) har tildelt allokering, som kan forventes at være eksogen ifht  $x \rightarrow y$ .
- Derfor betragtes naturlige eksp. som en kategori mellem rct og kvasi.
- Eksempler: vejrfænomener (vind, regn, strømforhold), geografi, topografi, diskontinuiteter pba lovgivning ifht fx kommuner/politikredse/sogn.
- Eksempler
  - Snow: Kolera i Soho 1854



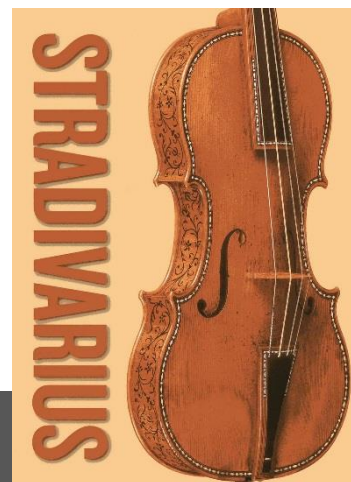


## 3.1 Naturlige eksperimenter

- SAS data
- <https://support.sas.com/documentation/tools/sashelpug.pdf>
- Mediation Effect of Low Birth Weight on Infant Mortality
- Sashelp.heart
- Framingham Heart Study
- [http://support.sas.com/documentation/cdl/en/etsug/63939/HTML/default/viewer.htm#etsug\\_syslin\\_sect007.htm](http://support.sas.com/documentation/cdl/en/etsug/63939/HTML/default/viewer.htm#etsug_syslin_sect007.htm)

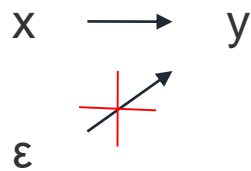
## 3.2 Instrumentmetode

- Metoden står og falder på, om man kan identificere et stærkt instrument (eksogent ifht problemstillingen)
- IKKE eksogenitet på treatmentvariabel – i stedet på instrumentvariabel
- Lineære effektfunktioner
- Local Average Treatment Effect (LATE) Philip G. Wrigt
- IV-reg med 2SLS

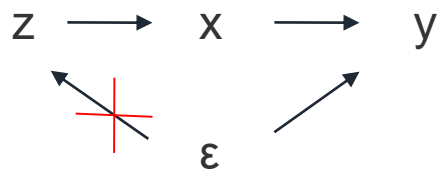


## 3.2 Instrumentmetode, IV-reg

- RCT: Selektionsbias udelukkes via randomisering

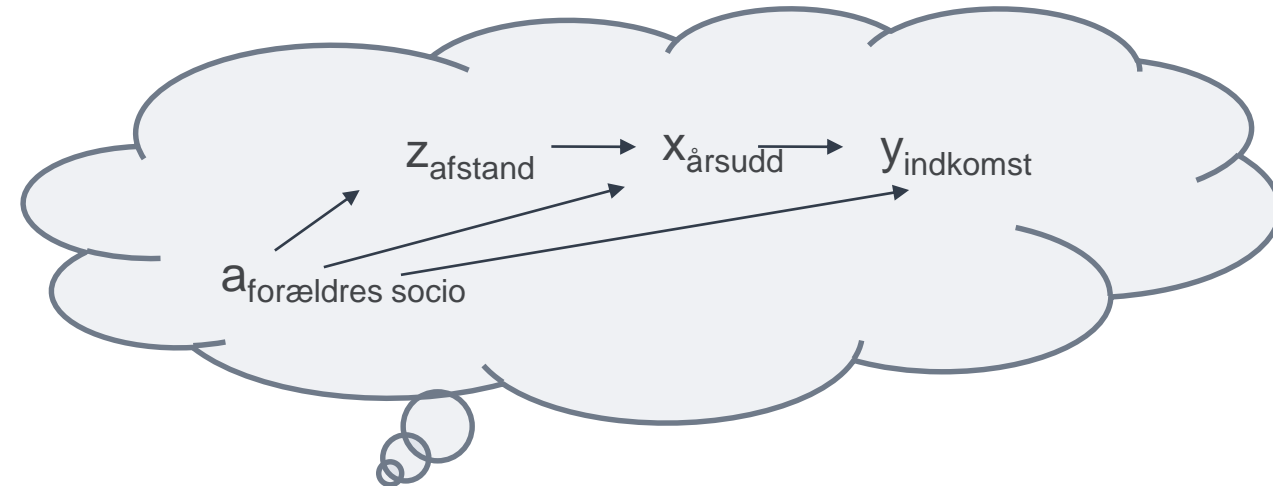
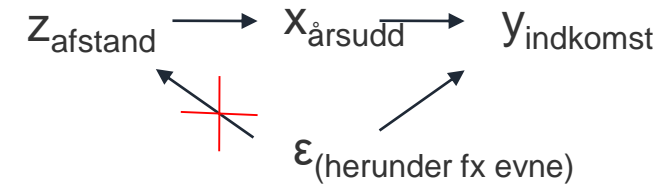


- IV-reg: Selektionsbias udelukkes ved at udelukke støj fra instrument



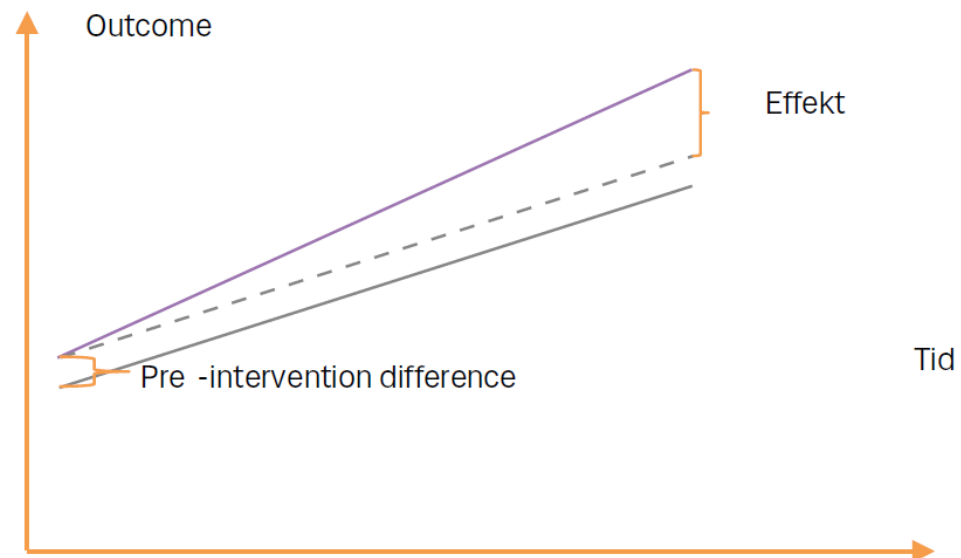
## 3.2 Instrumentmetode

- Ani Katchova
- <https://www.youtube.com/watch?v=bSLEGjbuit8>
- Års uddannelse  $\rightarrow$  Indkomst
- Stærkt instrument?



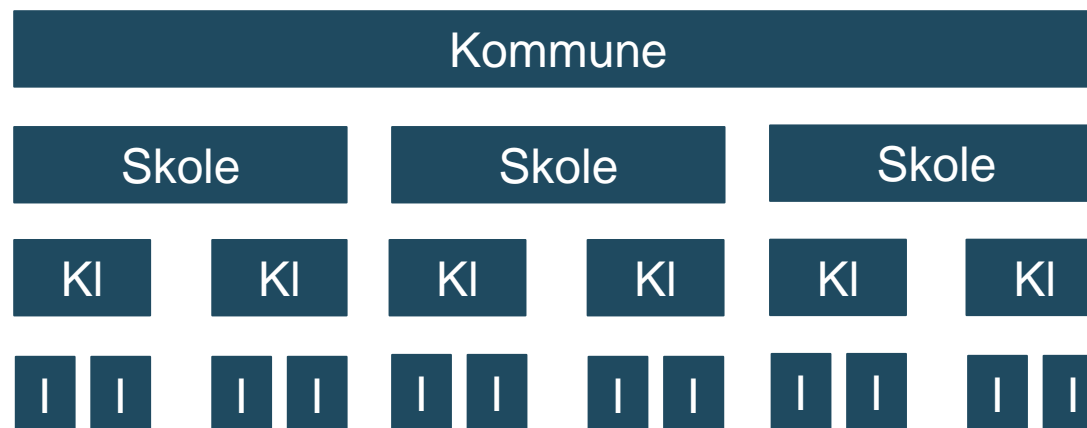
## 3.3 Difference-in-difference

- Difference-in-difference (MTE)
- Ligner dette design RCT?



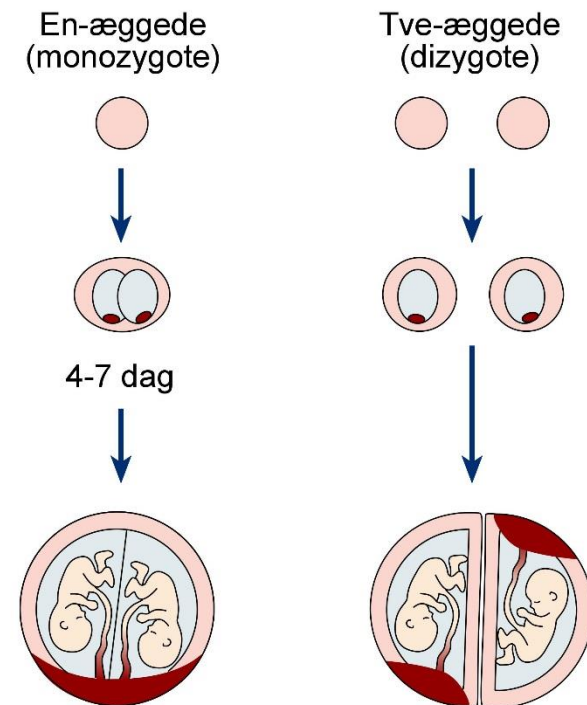
## 3.4 Multilevel

- Nastede data, strukturelle effekter (bias vs undersøgelsesfokus)
- Flere eksempler på, hvor man har strukturelt + individ-niveau?



## 3.5 Tvillingestudier

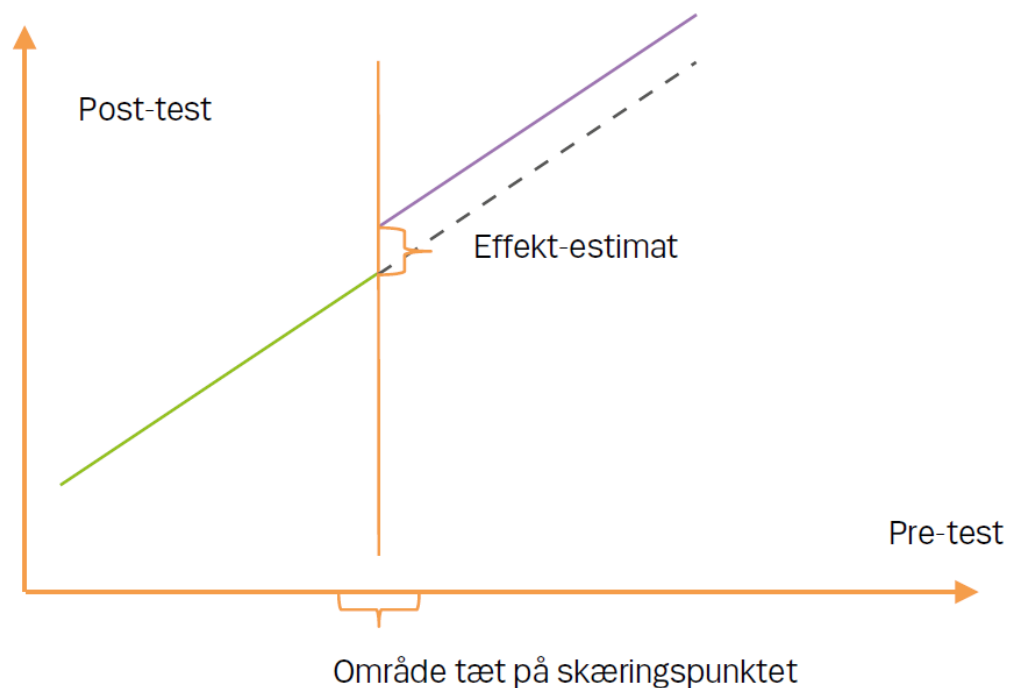
- Monozygote vs dizygote
- Hvad kan man bruge den forskel til?



© Birgitte Lerche-Barlach 2010

## 3.6 Regression discontinuity

- Ligner dette design difference-in-difference? Bemærk at det ikke er tid men effektmål på x-aksen!





## Grænselandet omkring klassisk effektmåling

- Progressionsmåling: Ikke effekt
- Pre-post studier: Ingen sammenligningsgruppe
- Simple komparative studier: Selektionsbias
- Inferens, logiske modeller: Induktive ræsonnementer + empiri
- Propensity score matching: Matche individer
- Sti-modeller: Teoretisk funderet regressionsmodel
- Dynamiske modeller: Variabel effekt
- Strukturelle ligningsmodeller: Sti-modeller med latente faktorer
- Systematiske reviews: Fixed og random effects

## Opsamling

- Hvad er den stærkeste metode?
  - RCT?
  - Faktorielle studier?
  - Kombination af metoder?
  - Review?
  - ?
- Formålet her har været at vise en række metoder med forskellige styrker og svagheder, som hver især har deres bidrag til at opbygge viden om effekter.

## Øvelse: Diskutér muligheder for effektmåling ved cases

- Cases udleveres
- Hver gruppe diskuterer en case og kommer med bud på effektestimat

## Cases

- **Print cases – Case for hver metode, forskellige sværhedsgrader**
- Design et effektstudie / diskutér design af et eksisterende effektstudie. (evt. flere runder)
- Cases: Fly-design → Jetlag, Ulighed → Trickle down, Hund → Eksamensangst, Skoler → Faglig kvalitet, Hulebyg → Resiliens, Overklasse → Overdreven selvtillid, Træning i kildekritik → Kildekritik, Smileys → Oplevet intelligens, Skole → Karakter, Være i rummet → Påvirkning af kroppen

## Cases

### Hjælpe spørgsmål til drøftelse af cases (jeg lader spørgsmålene stå):

- Er emnet relevant? (søg evt. på internet)
- Er mekanismen plausibel?
- Er der etiske udfordringer?
- Hvilken metode egner sig til at afdække emnet? (se s. 20 + 25)
- Er population/stikprøve afgrænset meningsfuldt?
- Hvilke outcomes egner sig? (søg evt. på internet)
- Hvilke typer bias kan der være i designet? (se s. 11-13)
- Er datagrundlaget stort nok?
- Er indsatsen ens for alle forsøgsdeltagere?
- Kan indsatsen forventes at være implementeret korrekt?

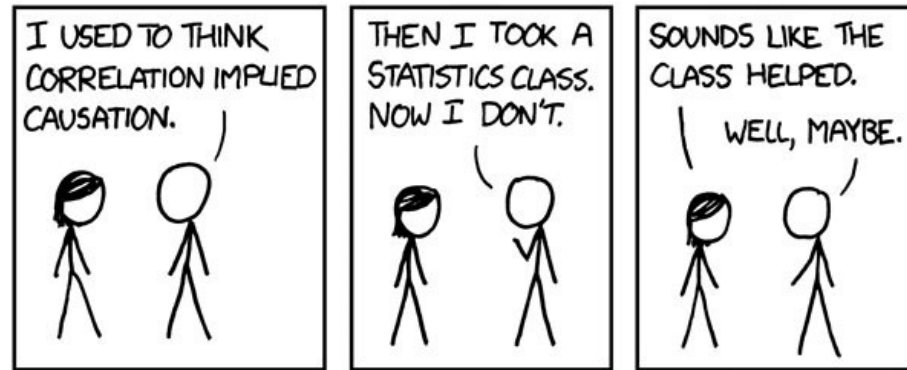
## Cases

- Effekten af nox-gasser på menneskers dødsfald: <https://www.dr.dk/nyheder/viden/klima/miljoetjek-er-diesel-vaerre-end-benzin>
- Effekten af partipolitisk tilknytning for huslejenævns afgørelser <https://www.dr.dk/nyheder/penge/formand-huslejenaevn-er-ogsaa-kasserer-i-konservative-ingen-politisk-indblanding>
- Effekten af kemo som behandling af cancer
- Flytype → Jetlag <https://www.dailymail.co.uk/sciencetech/article-4138362/Boeing-Airbus-design-planes-end-jet-lag.html>
- Ulighed → Spill-over (Piketty)
- Hund → Eksamensangst: <https://www.tv2nord.dk/artikel/hunden-sif-hjaelper-elever-med-eksamensangst>
- IV: Skoler → Faglig kvalitet: <https://www.sciencedirect.com/science/article/pii/S0272775717307860?via%3Dihub>
- Multilevel: European Social Survey, lande-effekter, skole-effekter, arbejdsplads-effekter, European Value Survey <https://www.vive.dk/da/udgivelser/faglig-skolekvalitet-ifoelge-krevi-og-undervisningsministeriet-8608/>
- Difference-in-differences, Resultatløn → Behandlingskvalitet <https://www.bmj.com/content/339/bmj.b3047.short>

## Cases

- Hulebyg → Resiliens: <https://theconversation.com/forest-schools-how-climbing-trees-and-making-dens-can-help-children-develop-resilience-117920>
- Overklasse → Overdreven selvtillid: <https://videnskab.dk/kultur-samfund/forskere-opvaekst-i-overklassen-giver-overdreven-selvtillid>
- Træning i kildekritik → Kildekritik: <https://videnskab.dk/krop-sundhed/effektmaaling-rigsarkivet-traener-elever-i-kildekritik-med-historiske-kilder-og>
- Smileys → Oplevet intelligens: <https://research.cbs.dk/en/publications/smiley-and-your-students-will-smile-with-you-the-effects-of-emoti>
- Skole → Karakter <https://www.vive.dk/da/udgivelser/folkeskolens-faglige-kvalitet-9748/>
- Skolestart → Mentalt helbred <https://www.vive.dk/da/udgivelser/the-gift-of-time-school-starting-age-and-mental-health-6487/>
- Være i ydre rum → Påvirkning af kroppen: <https://videnskab.dk/krop-sundhed/unikt-tvillingestudie-saadan-blev-astronauts-krop-paavirket-af-at-vaere-i-rummet>

# Spørgsmål?



XXKD



## Litteratur

Angrist, J. D. & Pischke, J. (2009) *Mostly harmless econometrics – an empiricist's companion*. New Jersey: Princeton.

Blalock, H. M., Jr. (1971) *Causal Models in the Social Sciences*. Chicago: Aldine.

Jadad, A. J. & Enkin, M. W. (2008) *Randomized Controlled Trials – questions answers and musings*. Chichester: John Wiley & Sons.

Pocock, S. J. (1983) *Clinical Trials – A practical approach*. Chichester: John Wiley & Sons.

Zachariae, B. (1998) *Det vellykkede eksperiment*. København: Munksgaard.