

Effektmåling

Iben Bolvig og Rasmus Højbjerg Jacobsen

DES Årskonference, Comwell, Kolding

14.9.2017

Kontaktoplysninger

Iben Bolvig, seniorforsker, ibbo@kora.dk



Rasmus Højbjerg Jacobsen, seniorforsker, raja@kora.dk



Program for seminaret

DEL 1 Forståelsen af effekter – hvad er en effekt?

- Kausal ændring A \rightarrow B
- Logik bag lodtrækningsforsøg

PAUSE – ca. 10 min

DEL 2 Regression-analyse og IV-analyse

- Hvad er problemer med en standard regressionsanalyse?
- IV - principper, metoder og eksempler

PAUSE – ca. 10 min

DEL 3 Regression discontinuity og Regression Kink

- Hvilke antagelser er vigtige?
- Principper og metoder

|| Hvorfor vil vi gerne effektmåle?

Vi ønsker at vide, om en given indsats rent faktisk har haft den virkning, som var hensigten med indsatsen.

Og vi vil også gerne sikre os, at indsatsen ikke har en skadelig virkning.

|| Hvorfor vil vi gerne effektmåle?

Vi ønsker at vide, om en given indsats rent faktisk har haft den virkning, som var hensigten med indsatsen.

Og vi vil også gerne sikre os, at indsatsen ikke har en skadelig virkning.

EKSEMPEL:

Vi vil gerne vide, om en iværksat indsats over for unge kriminelle har reduceret deres kriminalitet.

|| Hvad er en effekt?

Når man taler om *effekt*, taler man også automatisk om *kausalitet*.

At en indsats har en (positiv) effekt, betyder:

- 1) at de personer, der har fået indsatsen har oplevet et positivt skift, som de ikke ville have oplevet, hvis de ikke havde fået indsatsen
- 2) at det er indsatsens indhold, der har forårsaget det positive skift

Hvad er en effekt?

Der er ofte mange tal i medier mm., som ikke er en effekt, fx



IERVISNING. Eleverne Tobias, Lasse og Leo fra 4.d på Brønshøj Skole i København er trætte i sidste skoletime or, de ligger ned. De har matematik og er i gang med at måle areal af en presenning.

8 ud af 10 elever synes, at deres skoledage er for lange

82 procent af eleverne i folkeskolen synes, at deres dage er al lange. Elever og forældre savner variation i undervisningen. M er klar med flere penge og ny lov.

Dette udsagn i sig selv siger ikke noget om effekt.

Hvad er en effekt?

Der er ofte mange tal i medier mm., som ikke er en effekt, fx



IERVISNING. Eleverne Tobias, Lasse og Leo fra 4.d på Brønshøj Skole i København er trætte i sidste skoletime or, de ligger ned. De har matematik og er i gang med at måle areal af en presentning.

8 ud af 10 elever synes, at deres skoledage er for lange

82 procent af eleverne i folkeskolen synes, at deres dage er al lange. Elever og forældre savner variation i undervisningen. M er klar med flere penge og ny lov.

Dette udsagn i sig selv siger ikke noget om effekt.

EKSEMPEL, fortsat:

2 pct. af alle 10-17-årige har været mistænkt eller sigtet for noget kriminelt i 2015 (Justitsministeriet, 2016).

|| Hvad er en effekt?

Det er altså kun en effekt, når vi kan tage højde for/vurdere, hvad der ville være sket, hvis der ikke havde været en indsats.

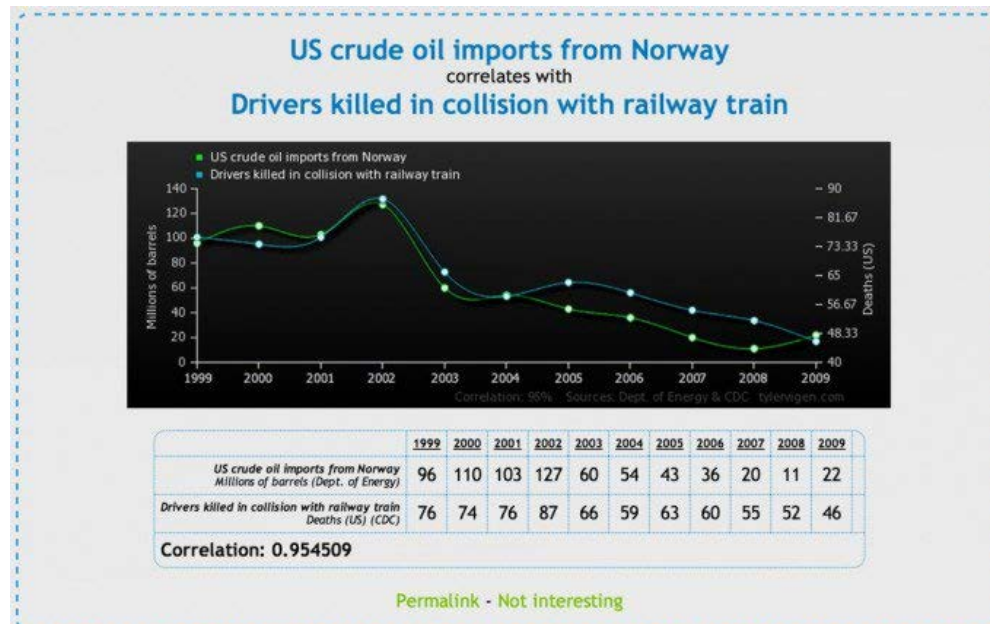
Det er relativt nemt at finde ud af, om der er en statistisk sammenhæng (korrelation) imellem indsats og udfaldsmål.

Hvad er en effekt?

Det er altså kun en effekt, når vi kan tage højde for/vurdere, hvad der ville være sket, hvis der ikke havde været en indsats.

Det er relativt nemt at finde ud af, om der er en statistisk sammenhæng (korrelation) imellem indsats og udfaldsmål.

MEN:



|| Hvad er en effekt?

Dvs. man skal have en sammenligningsgruppe (et kontrafaktisk udfald) for at kunne fastslå en effekt.

Sammenligningsgruppen kan dannes ud fra mange metoder, hvoraf vi skal se på nogen af dem i dag, fx

- Før-efter-måling (man er sin egen sammenligning)
- Matching
- Lodtrækning
- Etc.

Udfaldsmål

I mange tilfælde vil udfaldsmålet for en effektmåling relativt let at bestemme:

- Beskæftigelsesindsats
- Medicinsk behandling
- Stofmisbrugsbehandling

Men nogle udfald lader sig ikke så let måle (fx trivsel), og i andre tilfælde kan der være intermedieære udfald, som er vigtige at få med (måske især hvis der kan være tvivl om, hvorvidt endemålet vil blive nået).

EKSEMPEL:

Vi vil gerne vide, om en iværksat indsats over for unge kriminelle har reduceret deres kriminalitet.

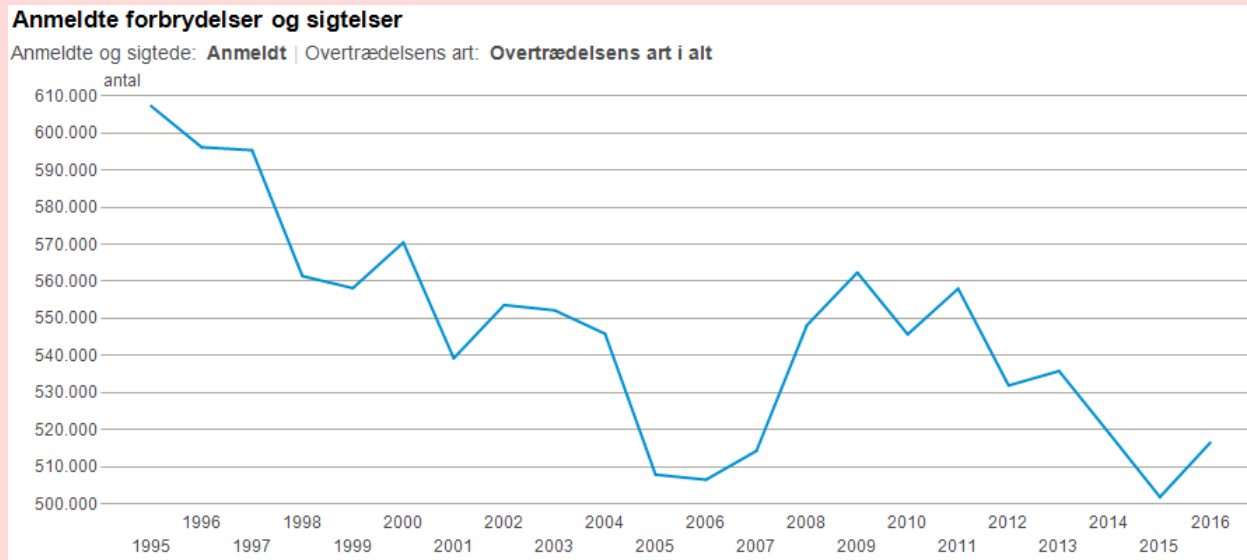
Lad os forestille os, at der i 2010 blev iværksat en indsats over for unge kriminelle.

Vi vil måle på, om deres kriminalitet er reduceret i 2015.

Som udfaldsmål vælges fx: andelen af de unge kriminelle, der har begået ny kriminalitet i perioden 2011-15.

EKSEMPEL:

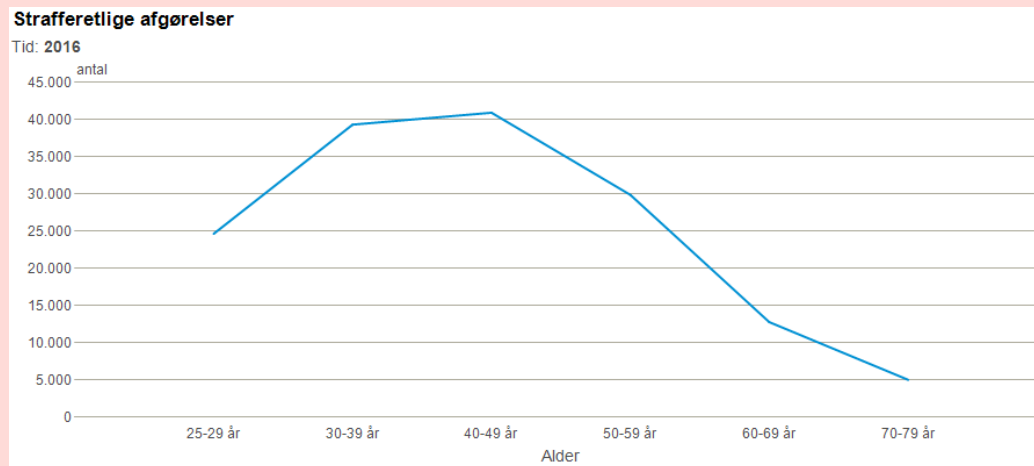
Det vil ikke være nok at følge deltagerne over tid. Der har fx været et stort fald i anmeldte forbrydelser fra 2010 til 2015:



Et fald i kriminalitet for deltagerne kunne altså lige så godt været forårsaget af det generelle fald i kriminaliteten.

EKSEMPEL:

Ud over problemet med et fald i kriminaliteten over tid, er der også en klar sammenhæng imellem kriminalitet og alder:



Måske er kriminaliteten blandt deltagerne ændret, blot fordi de er blevet ældre?

EKSEMPEL:

I dette eksempel kunne man altså relativt let stå med en kvantitativ måling, der viser, at kriminaliteten er faldet for deltagerne.

Så indsatsen må altså have været en succes? Nej, ikke nødvendigvis.

Dette eksempel er et godt eksempel på, hvorfor en før-efter-måling ikke nødvendigvis giver et retvisende billede.

Før-efter-måling kan ikke tage højde for faktorer, der ændrer sig over tid som fx alder og generelle tidstrends.

Derfor rangerer før-efter-måling ikke så højt på evidensstigen.

Lodtrækningsforsøg

Lodtrækningsforsøg (RCT – randomized control trial)

Idéen er simpel:

1. Identificer en gruppe af personer, der kan tilbydes en bestemt indsats/behandling.
2. For hver mulig deltager trækkes lod, om man skal have indsatsen eller ej.
3. Efter indsatsen måles, om der er forskel på udfaldet imellem indsatsgruppen og kontrolgruppen.

Lodtrækningsforsøg

I princippet er der ingen systematiske forskelle imellem indsats- og kontrolgruppe i en lodtrækning (i hvert fald ikke, hvis gruppen er stor nok).

Men det er ikke altid så nemt i praksis:

- Problemer med randomiseringen, hvis den der "trækker lod" kender udfaldet af lodtrækningen.
- Ethiske problemer (skal man ikke tilbyde en gavnlige indsats til alle?)
- Ex ante vs. ex post sammenligning (balancering)

EKSEMPEL:

Hvordan vil man udføre lodtrækning i eksemplet?

Flere mulige niveauer:

- Individ
- Afdelinger
- Fængsler

Afdeling/fængsel er måske nemmest i praksis, men det risikerer også let at give en ikke-balanceret sammenligning.

OPGAVER

Opgave 1:

Hvad vil I gerne måle effekt af?

Opgave 2:

Hvad er den kontrafaktiske situation?

Opgave 3:

Hvad er effektmålet? – og hvordan måles det?

Opgave 4:

Vil man kunne gennemføre lodtrækning? Hvordan?

Effektmåling – DEL 2

Iben Bolvig og Rasmus Højbjerg Jacobsen

DES Læringsseminar, Comwell, Kolding, 14.9.2017

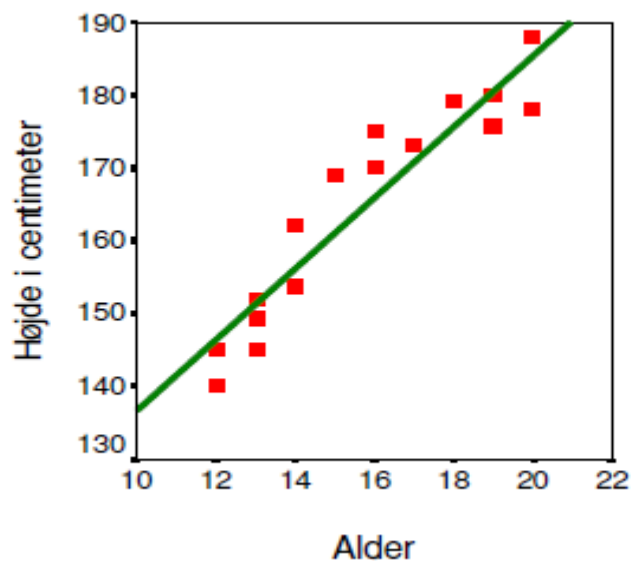
Regressionsanalyse

“I checked it very thoroughly, said the computer, and that quite definitely is the answer. I think the problem, to be quite honest with you, is that you’ve never actually known what the question is.”

Douglas Adams, *The Hitchhiker’s Guide to the Galaxy* (1979)
(Angrist & Pischke, 2008)

Regressionsanalyse

OLS regression



- $y_i = \alpha + \beta x_i + u_i$
- Beta viser hældningen på kurven
- Alfa viser niveauet på kurven.

Estimater

- β estimeres ved at minimere afstanden mellem datapunkterne og det estimerede output \hat{y} :

$$\text{MIN} \left((y - (\hat{\alpha} + \hat{\beta}x))^2 \right)$$

$\rightarrow \hat{\beta}$ og $\hat{\alpha}$

- Dvs. regression giver den bedste lineære prædiktion af den afhængige variabel

Anvendelse af lineær regression i effektmåling

- Antag at deltagelse i indsatsen er tilfældig - betinget på et sæt af observerbare karakteristika.
- $Y_i = \alpha + \gamma X_i + \rho D_i + \mu_i$
- $\hat{\rho}$ kan siges at estimere den kausale effekt, *hvis* deltagelsesbeslutningen er ukorreleret med μ_i
- Med andre ord, der må ikke findes uobserverede karakteristika, der kan have betydning for deltagelse eller ikke-deltagelse og som samtidig påvirker output.

Long-run outcome of a Danish welfare-to-work experiment (Sofie Theilade Nyland Brodersen (2014))

Table 3.4. OLS estimates on main outcomes I (with time fixed effects)

	Employment	Self-sufficiency	Social assistance	Disability pension
Treated	-0.003 (0.006)	-0.002 (0.004)	-0.024** (0.011)	0.028*** (0.010)
Male	0.001 (0.006)	0.011*** (0.004)	-0.025** (0.012)	0.009 (0.011)
Married	-0.002 (0.007)	0.007 (0.005)	-0.038*** (0.013)	0.014 (0.012)
Danish native	-0.005 (0.017)	-0.018 (0.015)	-0.009 (0.035)	0.049* (0.029)
Non-western immigrant	-0.016 (0.018)	-0.022 (0.016)	0.054 (0.036)	0.040 (0.030)
Age 29-39	0.068*** (0.007)	0.005 (0.012)	0.200*** (0.028)	-0.028 (0.024)
Age 40-49	0.057*** (0.006)	-0.011 (0.012)	0.157*** (0.028)	0.055** (0.024)
Age 50-59	0.027*** (0.006)	-0.012 (0.012)	0.096*** (0.029)	0.135*** (0.026)

Problemer med udeladte variable

- Sand model:

$$test_i = \alpha + \beta \cdot klassestørrelse_i + \gamma \cdot evne_i + u_i$$

- Men vi kender ikke $evne_i$, så den estimerede model:

$$y_i = \alpha + \beta \cdot klassestørrelse_i + \varepsilon_i$$

- Dvs. $\hat{\varepsilon}_i = \gamma \cdot evne_i + \hat{u}_i$
- Er det et problem at $E(\hat{\varepsilon}_i) \neq 0$?

Problem med udeladte variable

- Ja, hvis *evne* og *klassestørrelse* er korrelerede
- $\text{Min}(\hat{\epsilon}) = \text{Min}((\gamma \cdot \text{evne}_i + \hat{u}_i)^2)$
- $\rightarrow \hat{\beta} = \beta + \frac{\text{COV}(\text{evne}, \text{klassestørrelse})}{\text{var}(\text{klassestørrelse})}$
- Dvs. medmindre *evne* og *klassestørrelse* er uafhængige vil beta estimatet blive skævvredet.

TABLE 1 Unstandardized Coefficients From Linear Regression Analyses of Self-Reported Life Satisfaction in the 2000 American National Election Study.

Predictor	Step 1	Step 2	Step 3
Political conservatism	0.25 (0.06)***	0.19 (0.06)**	0.15 (0.06)*
Income		0.13 (0.11)	0.15 (0.11)
Education		0.19 (0.08)**	0.16 (0.08)*
Sex		-0.00 (0.04)	-0.02 (0.04)
Age		0.08 (0.06)	0.04 (0.06)
Age squared		0.21 (0.05)***	0.20 (0.05)***
Marital status		0.14 (0.04)***	0.12 (0.04)**
Employment status		-0.28 (0.12)*	-0.27 (0.12)*
Church attendance			0.20 (0.05)***
Need for cognition			
Rationalization of inequality			

Note. Robust standard errors are given in parentheses. This analysis is based on a sample size
⁺ $p < .10$. * $p < .05$. ** $p < .01$. *** $p < .001$.

Napier J L , and Jost J T Psychological Science
 2008;19:565-572

Opgave

- Sæt jer sammen to og to og se på omdelte regressionsmodel
- Hvad viser estimerne?
- Hvilken sammenhæng er der mellem elevernes familiebaggrund og deres karakterer?
- Hvad kan man sige om valg af forklarende variable?
- Kan regressionen sige hvad effekten af et skoleskift er på karakterniveauet?
 - Hvad kan problemet være?

Instrument variabel metode

Anything that happens, happens.

Anything that, in happening, causes something else to happen,
causes something else to happen.

Anything that, in happening,
causes itself to happen again, happens again.

It doesn't necessarily do it in chronological order, though.

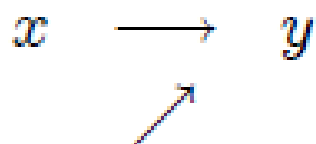
Douglas Adams, *Mostly Harmless* (1992)

(Angrist & Pischke, 2008)

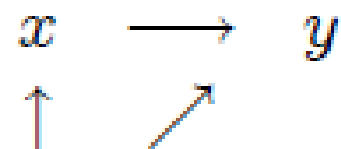
Ideen bag

- Håndtering af problem med udeladt variable
- IV-metoden anvender instrumenter, der er stærkt korrelerede med variabelen af interesse (den endogene variabel) og som samtidigt ikke er korreleret med den udeladte variabel.
- Dermed kan β estimeres uden at blive skævvredet.

Ideen bag

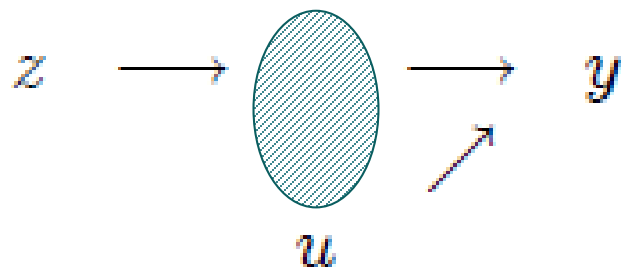


u



u

$$\frac{dy}{dx} = \beta + \frac{du}{dx}$$



- Dvs. z påvirker kun y gennem x .

Modellen med instrument

- $test_i = \alpha + \beta \cdot klassestørrelse_i + \gamma \cdot evne_i + u_i$
- Problem: vi har ikke data på evne
- Løsning: find et instrument, der er korreleret med klassestørrelse, men som ikke er korreleret med evne.
- Den første betingelse beskriver instrumentets styrke
 - Ved lav korrelation/styrke er instrumentet evt. ikke stærkt nok til at identificere effekten
- Den anden betingelse er validitetsbetingelsen. Den SKAL være opfyldt for at metoden løser problemet.

Modellen med instrument

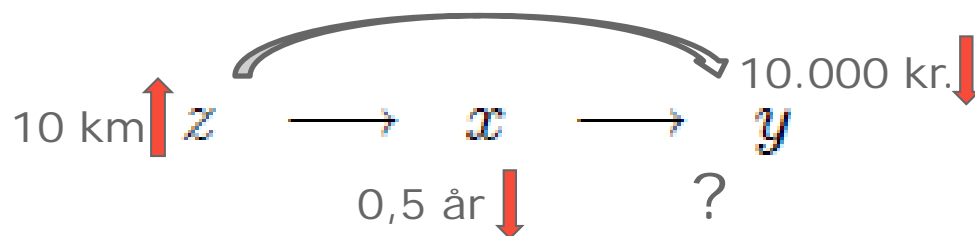
- Tænk over et muligt instrument, der er korreleret med klassestørrelse, men som ikke er korreleret med evne.
- Fx en reform, der ændrer klassekoefficienten
- Skolesammenlægning, der ændrer elevgrundlag

Et intuitivt eksempel – beregning af effekten

- Betydning af uddannelsesniveau på lønnen.
- Problem: *kompetencer* er udeladt variabel, som både har betydning for uddannelsesniveau og løn.
- Vi bruger flytning af uddannelsesinstitutioner, som tilfældig variation, så:
 - afstand til uddannelsesinstitution agerer instrument for uddannelsesniveau
 - Dvs. ændring i afstand til uddannelsesinstitution påvirker uddannelsesniveaue (styrken af instrument)
 - Men afstand til uddannelsesinstitution er ikke korreleret med kompetenceniveau (fordi afstand skyldes flytning af uddannelsessteder)

Et intuitivt eksempel – beregning af effekten

- Beregning af sammenhængen:
 - hvis 10 km længere afstand (z) vil sænke uddannelseslængden med 0,5 år
 - Og 10 km længere afstand (z) vil sænke lønnen med 10.000 kr. om året (y)



- \rightarrow et ekstra år i skole vil hæve lønnen med 20.000 kr. om året.
- Svarende til $-10.000 / -0.5 = 20.000$

Eksempel

Angrist & Evans (1998). Children and their parents' labor supply.

- Hvad er effekten af et ekstra barn på forældres udbud af arbejdskraft?
- Hvad kan det potentielle problem ved blot at estimere dette være?
- Udnytter variation i antal af børn som følge af mix af køn for de første to børn.

Angrist & Evans (1998). Children and their parents' labor supply.

Sex of first two children in families with two or more children	All women			
	1980 PUMS (394,835 observations)		1990 PUMS (380,007 observations)	
	Fraction of sample	Fraction that had another child	Fraction of sample	Fraction that had another child
one boy, one girl	0.494	0.372 (0.001)	0.495	0.344 (0.001)
two girls	0.242	0.441 (0.002)	0.241	0.412 (0.002)
two boys	0.264	0.423 (0.002)	0.264	0.401 (0.002)
(1) one boy, one girl	0.494	0.372 (0.001)	0.495	0.344 (0.001)
(2) both same sex	0.506	0.432 (0.001)	0.505	0.407 (0.001)
difference (2) – (1)	—	0.060 (0.002)	—	0.063 (0.002)

Angrist & Evans (1998). Children and their parents' labor supply.

Gns. Forskel i ssh. mellem "første to af samme køn" og "første to af forskellig køn".

Variable	Mean difference by <i>Same sex</i>	Wald estimate using as covariate:	
		<i>More than 2 children</i>	<i>Number of children</i>
<i>More than 2 children</i>	0.0600 (0.0016)	—	—
<i>Number of children</i>	0.0765 (0.0026)	—	—
<i>Worked for pay</i>	-0.0080 (0.0016)	-0.133 (0.026)	-0.104 (0.021)
<i>Weeks worked</i>	-0.3826 (0.0709)	-6.38 (1.17)	-5.00 (0.92)

$$z \rightarrow Y / z \rightarrow x$$

Angrist & Evans (1998). Children and their parents' labor supply.

- Hvad siger estimaterne?
- Hvad skal man være opmærksom på ift. at generalisere resultatet?

	All women		
	(1)	(2)	(3)
Estimation method	OLS	2SLS	2SLS
Instrument for <i>More than 2 children</i>	—	<i>Same sex</i>	<i>Two boys, Two girls</i>
Dependent variable:			
<i>Worked for pay</i>	-0.176 (0.002)	-0.120 (0.025)	-0.113 (0.025) [0.013]
<i>Weeks worked</i>	-8.97 (0.07)	-5.66 (1.11)	-5.37 (1.10) [0.017]

Estimation i to niveauer (2SLS)

- I første stadie estimeres sammenhængen mellem instrumentet (fx en reform) og variabelen af interesse (fx klassestørrelsen).
- Denne sammenhæng bruges til at beregne de forventede værdier af klassestørrelse
- Forventede værdier fra stage 1:
- $\hat{S}_i = X_i' \hat{\pi}_{10} + \hat{\pi}_{11} Z_i$
- $Y_i = X_i' \alpha + \rho \hat{S}_i + \eta_i + \rho (S_i - \hat{S}_i)$
- ρ er konsistent så længe X_i og Z_i er ukorrelerede med både η_i og $S_i - \hat{S}_i$

Two Stage Least Square

- Fordelen ved denne estimator er, at den kan håndtere flere instrumenter for hver variabel.
- $$s_i = X_i' \pi_{10} + \pi_{11} Z_{1i} + \pi_{12} Z_{2i} + \pi_{13} Z_{3i} + \xi_{1i}$$
- Fx dummies for fødselsdag i første anden og tredje måned, jf. efterlønsordningen.

Valg af instrumenter

- Der gælder som nævnt følgende krav for valide instrumenter (z):
 1. Z er ukorreleret med de udeladte variable (og dermed η)
 2. Z er korreleret med den endogene variabel (*klassestørrelse*)
 3. Z er hellere stærkt korreleret med *klassestørrelse* end svagt korreleret.

Valg af instrumenter

- Se på de teoretiske sammenhænge – giver det mening?
- Anvend, hvis der er en reform eller andet chok, der kan udnyttes.
- Ellers, anvend fx 2. og 3. ordens sammenhænge, dvs. x^2 eller x^3
- Anvend interaktioner af modelens øvrige variable fx $køn \cdot afstand$ (mænds afstand til uddannelse har større betydning end kvinders, mens der ikke er forskel i afstandens betydning for lønnen mellem de to køn).

|| Test af instrument

- Relevanskriteriet (er instrumentet korreleret med variabelen af betydning) kan testes i 1. stage modellen.
- Mens validitetskriteriet (er instrumentet ukorreleret med alle andre betydende variable) kun kan testes, hvis antal instrumenter $>$ antal endogene variable (overidentificeret).

Test af instrumenter

- Test af validitet kun muligt med flere instrumenter
 - Sargan's test for overidentificerede restriktioner
- Overordnet tommelfingerregel for relevans af instrumenter er en F-statistik i 1. stage >10 .
- MEN selv med dette opfyldt kan der vise sig at være svage instrumenter i modellen.
- Desuden: Anderson's rank-test (rangen af Z matricen) eller Stock & Yogo's test af relative bias
- Alle test kan findes i Stata.
- **Svage og mange instrumenter kan føre til bias lige stor som OLS bias!**

5 gode råd

1.

Angiv 1. stage og tænk over om det giver mening! Er retning og størrelse af estimatet som forventet?

- hvis ikke er din hypotese måske ikke sand og du har måske bare været heldig!

5 gode råd

2.

Rapporter F-statistikken for instrumenterne. Jo større denne er, jo bedre. Forslag $F > 10$, men ikke en "sikker" regel.

5 gode råd

3.

Vælg dit bedste instrument og rapporter estimater baseret kun på dette instrument. Præcis-identificeret IV er langt bedre end at bruge en stribe svage instrumenter.

5 gode råd

4.

Undersøg koefficienter, t-statistik og F-statistik i regression af de afhængige variable på instrumenterne (dvs Y på Z).

Estimaterne i denne reducerede form er proportionale med den kausale sammenhæng af interesse. Hvis du ikke kan finde en reduceret sammenhæng, så findes den kausale sammenhæng nok heller ikke! (fx ingen sammenhæng ml. løn og afstand til uddannelsesinstitution).

|| 5 gode råd

5.

Brug altid statistik software!

Estimation i Stata

- Brug standard software programmer – der er mange fordele, fx
 - korrigerede standard afvigelser
 - undgår du fejl i de fittede værdier
 - robusthedstjek med forskellige specifikationer (IV, 2SLS, GMM, LIML)
 - test af relevans og overidentifikation.
- **ivreg2** depvar varlist1 (varlist2=varlist_iv)
- Hvor depvar=Y, varlist1=exogene forklarende variable, varlist2=endogene variable, varlist_iv=instrument variable.

Effektmåling – DEL 3

Iben Bolvig og Rasmus Højbjerg Jacobsen

DES Læringsseminar, Comwell, Kolding, 14.9.2017

Regression discontinuity (RD)

Grundlæggende er RD en kvasiekperimentel metode, dvs.

- vi vil gerne opnå nogle af de samme statistiske egenskaber som en lodtrækning
- der bliver ikke foretaget lodtrækning

Idéen er, at der findes en underlæggende allokeringsvariabel, som påvirker individets deltagelse i en indsats, men hvor der ved en given tærskelværdi er forskel på, om man får indsatsen eller ej. Denne variabel kaldes også "forcing"-variablen.

Regression discontinuity

Det er vigtigt, at allokeringsvariablen er uafhængig af udfaldsvariablen.

Allokeringsvariablen kan fx være alder, testscore, geografiske grænser mv.

Typisk er det sådan, at hvis D angiver deltagelse, så

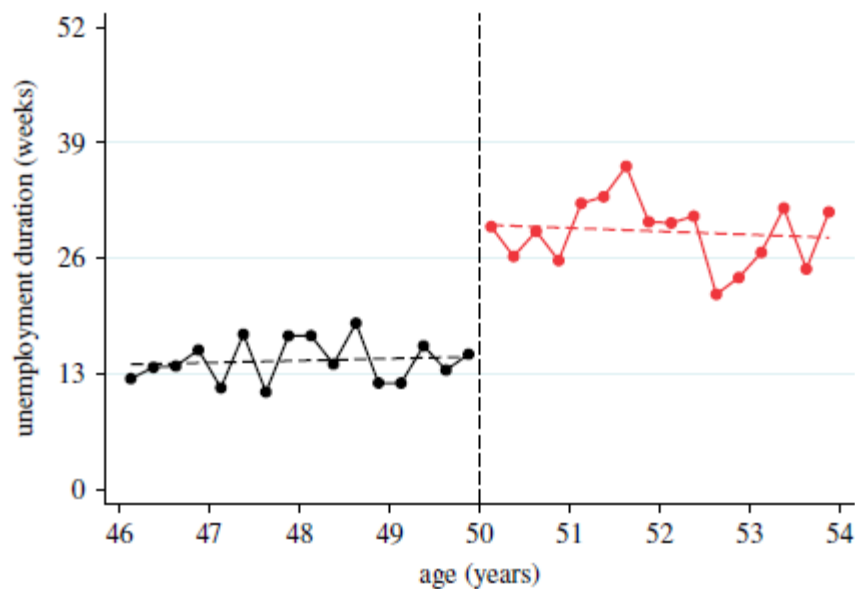
$$D = \begin{cases} 0 & \text{hvis } x \leq c \\ 1 & \text{hvis } x > c \end{cases}$$

hvor c er tærskelværdien for allokeringsvariablen.

Regression discontinuity

Eksempel (Lalive, 2008):

Østrig indførte et program, der tildelte personer over 50 år ret til længere arbejdsløshedsunderstøttelse:



Discontinuity at threshold = 14.798; with std. err. = 1.928.

Regression discontinuity

Generelt findes sådanne diskontinuiteter inden for

- Skattesystemet
- Offentlige overførsler
- Skolesystemet

Det er relativt let at se, om der er en ikke-kontinuitet ved at lave en figur, hvor allokeringsvariablen er på X-aksen og udfaldsvariablen er på Y-aksen.

Regression discontinuity

Det er vigtigt, at

- personerne ikke har fuld kontrol over, om de befinder sig på den ene eller den anden side af grænsen. Oftest vil fx skattesystemet kunne give incitamenten til at vælge, hvis man kan.

Hvis denne antagelse er opfyldt, er der imidlertid store fordele ved RD:

- Allokeringen til indsats/kontrol er helt tilfældig lige omkring tærskelpunktet. Det svarer til lodtrækning.
- Man kan teste, om antagelsen om tilfældighed er opfyldt. Hvis den er det, må der ikke være systematisk variation imellem indsats- og kontrolgruppe.
- Metoden kan (som et startpunkt) let lade sig fremstille i en grafisk præsentation. MEN: det siger ikke noget om effekt.

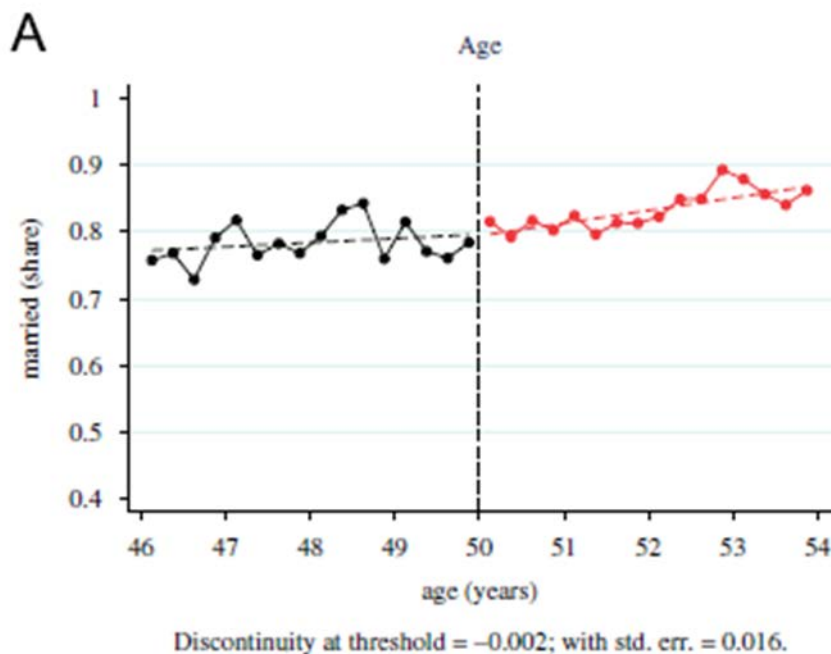
Regression discontinuity

En måde at undersøge, om individerne på hver side af tærskelværdien er tilfældigt fordel, er ved at lave graf med allokeringsvariablen på X-aksen og forklarende baggrundsvariable (fx køn, alder, uddannelse osv.) på Y-aksen.

Disse må ikke vise et spring ved tærskelværdien.

Figur fra Lalive (2008), andel gifte som funktion af alder.

Man skal huske at teste statistisk!



Estimation af RD

Ved en simpel lineær ikke-kontinuitet estimeres RD ved en ligning som denne ved en ikke-parametrisk estimation

$$Y = \alpha + \tau D + \beta_1(X - c) + \beta_2 D(X - c) + u$$

Her er τ den variabel, vi er interesseret i. Ved estimation vælges en båndbredde på hver side af tærskelværdien, c .

Fortolkning af RD-estimer

Der er to mulige fortolkninger af RD-estimatoren:

1. Hvis man finder en effekt ved tærskelværdien, så gælder den effekt, der observeres kun for individer med værdier meget tæt på tærskelværdien, fx for elever med en testscore tæt på værdien.
2. Estimatet kan fortolkes som en mere bred effekt af deltagelse ved at se estimatet som en vægtet værdi, med højere vægt, jo tættere man er på tærskelværdien.

Regression discontinuity

Uanset, om effekten kan fortolkes lidt bredere, så vil RD estimatoren kun kunne sige meget lidt om effekten væk fra tærskelværdien.

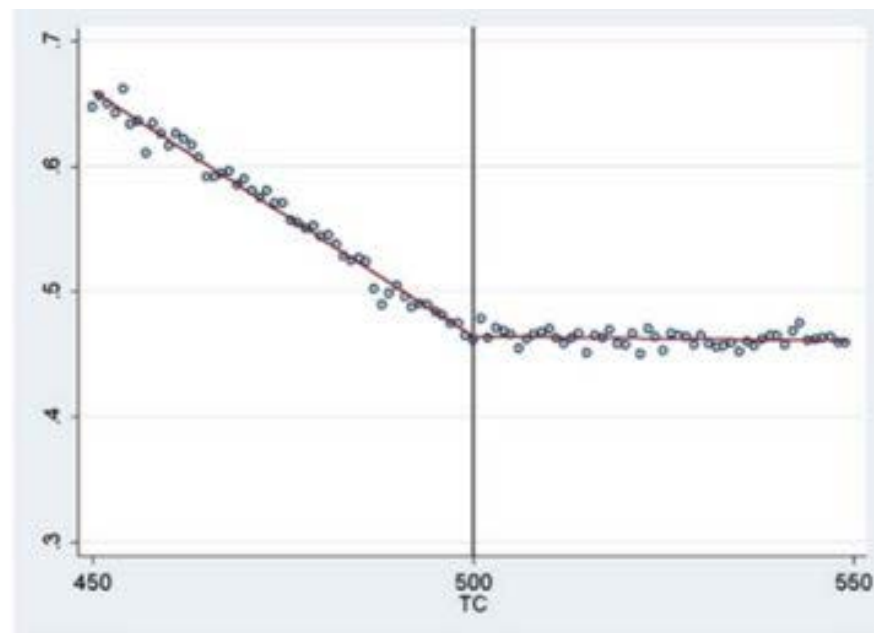
Dette er fælles med lodtrækningsforsøg og andre kvasiekperimentelle metoder som fx matching.

Regression kink design

Regression kink design er et specialtilfælde af regression discontinuity, hvor der i stedet for en ikke-kontinuitet er et "knæk" for den variabel, der undersøges. Dvs. der er en ikke-kontinuitet i de første afledede mht. forcing variabelen:

Fx viser Simonsen et al. (2015):

Totale omkostninger ved
receptpligtig medicin.



Afsluttende pointer om RD og RK

1. Lav altid en grafisk præsentation og inspektion. Dette giver en god idé om, hvorvidt der er en ikke-kontinuitet, og om hvilken form denne har. Er den fx lineær eller ikke-lineær?
2. Er der visuelt et tegn på, at der optræder en effekt ved tærskelværdien? Dvs. er der forskel på, hvordan individer på hver side af grænsen optræder? Hvis ikke, så er der næppe nogen effekt.
3. Husk altid at teste for, om grupperne på hver side af tærsklen er ens. Man kan evt. starte med en "grafisk test".

Eksempler på RD i Danmark

Skolestart: Efter reglerne skal man starte i skole i det år, man fylder 6 år. Derfor vil der være nogle, der er født lige omkring d. 1. januar, som enten starter i skole eller ikke gør det (Simonsen et al., 2015).

Efterløn: Under indfasning af efterlønsreformen var der aldersgrænser for, om man var berettiget til efterløn eller ej.

Jobpræmie: Ledige, der havde været ledige mere end 47 uger kunne få en præmie på op til 4% af jobbet's månedsløn (Kolodziejczyk og Arendt, 2017).

Geografi: Visse udkantskommuner har forhøjet befordringsfradrag for pendlere.

Opgaver

Opgave 1:

Kan I tænke på nogen mulige diskontinuiteter inden for de områder, I normalt beskæftiger jer med?

Opgave 2:

Hvordan vil I undersøge, om antagelserne til RD er opfyldt?

Opgave 3:

Tror I individerne vil kunne påvirke deres allokering på den ene eller den anden side af tærskelværdien?