

VIDEN PÅ TVÆRS AF EFFEKTDESIGN

METTE DEDING, SFI CAMPBELL

PRIMÆR VS. SEKUNDÆR EFFEKTFORSKNING

- Primær effektforskning
 - Studium af grunddata. Undersøgellesdesign afhænger af problemstilling og datamuligheder.

- Sekundær effektforskning
 - Sammenfatning af resultater af primær studier inden for et specifikt område
 - Øget validitet, hvis mange studier peger i samme retning

PRIMÆR VS. SEKUNDÆR EFFEKTFORSKNING

Systematiske Campbell forskningsoversigter:

- En særlig type sekundær effektforskning
- Er én af brikkerne til forbedring af blandt andet socialt arbejde. Men det samlede puslespil bygger i sagens natur også på mange andre former for viden.

DISPOSITION

- Systematiske Campbell forskningsoversigter
- Hvordan måler man effekter
- Vurdering af forskellige studiedesigns
- Evidenstypologi vs. evidenshierarki

SYSTEMATISKE CAMPBELL FORSKNINGSOVERSIGTER

En systematisk Campbell forskningsoversigt er en videnskabeligt udført sammenfatning af *alle* tilgængelige effektstudier, der undersøger en bestemt indsats.

Formålet med forskningsoversigten er at give det bedst mulige svar på, om en bestemt intervention, rettet mod en bestemt gruppe og målt på eksplicitte kriterier, virker.

CAMPBELL METODEN (I)

1. Titelregistrering
2. Protokol (arbejdsplan)
3. Review
4. Update

5. Peer review proces

CAMPBELL METODEN (II)

1. Opstille kriterier for videnskabelighed i de studier, der medtages i analysen
 2. Søge systematisk efter studier vha. standardiseret søgestrategi
 3. Vurdere alle de fundne studier efter de opstillede inklusionskriterier
 4. Analyse af de studier, der lever op til inklusionskriterierne
- Et centralt element: hvordan kvalitetsvurderer vi de eksisterende effektstudier?

HVORDAN MÅLER MAN EFFEKTER: EFFEKT Forskningsens Grundspørgsmål (I)

1. Ændrer problemerne karakter i retning af større målopfyldelse, efter at en indsats (intervention, foranstaltning) er igangsat?
 - "bruttoeffekt"
2. Er indsatsen skyld i ændringen – eller skyldes det andre forhold?
 - "nettoeffekt" eller kausal effekt

EFFEKTFORSKNINGENS GRUNDSPØRGSMÅL (II)

Nettoeffekten kræver svar på effektmålingens præmiespørgsmål:

Hvordan går det personer, der modtager indsatsen, i forhold til hvordan det ville være gået, hvis de ikke havde modtaget indsatsen?

Hvilket samtidig er effektmålingens kerneproblem:

Vi ved ikke, hvordan det ville være gået personerne, hvis de ikke havde modtaget indsatsen (det kontrafaktiske spørgsmål)

(Fra brutto- til nettoeffekter: fjerne betydningen af forstyrrende forhold)

EFFEKTFORSKNINGENS GRUNDSPØRGSMÅL (III)

Præmiespørgsmålet for den systematiske Campbell forskningsoversigt bliver så:

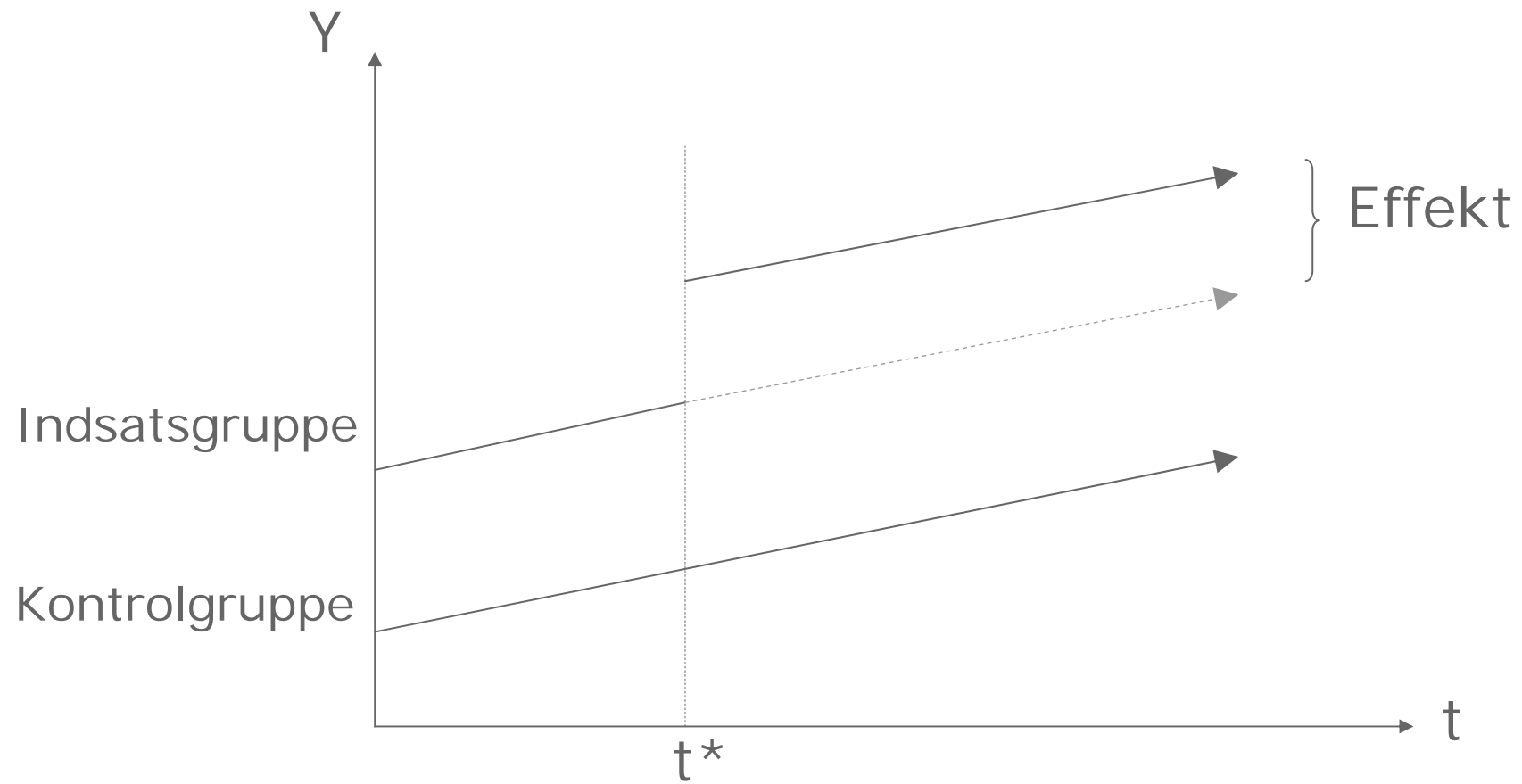
Hvor gode er de enkelte studier til at estimere nettoeffekter?
Hvordan håndterer de betydningen af "forstyrrende forhold"?

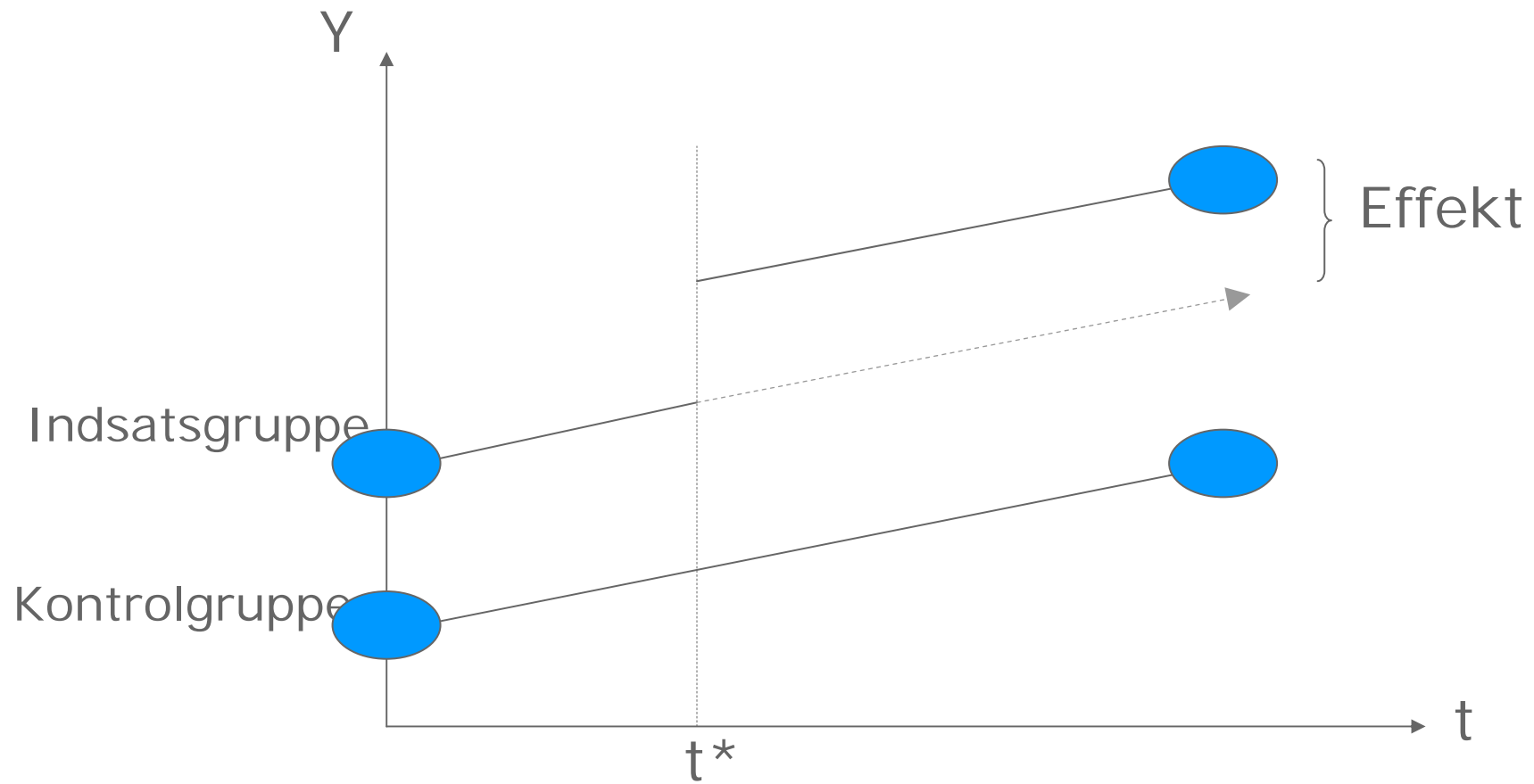
FORSTYRRENDE FORHOLD - UDEFRA

- Endogen forandring, "naturlige" processer
 - Fx går mange sygdomme over af sig selv
- Modningsprocesser
 - Fx bliver unge mindre kriminelle med alderen
- Generelle socio-økonomiske og demografiske udviklingstræk
 - Fx virker beskæftigelsestiltag bedre i højkonjunktur
- Forstyrrende begivenheder
 - Fx kunne 11. september eller Muhammed-krisen påvirke effekten af integrationstiltag
- Selvselektion
 - Fx melder de mest ressourcestærke sig ofte til lovende nye foranstaltninger

FORSTYRRENDE FORHOLD - INDEFRA

- Forhold knyttet til selve indsatsen
 - Hawthorne/placebo-effekt
 - Implementeringsproblemer
 - Den faktiske intervention svarer ikke til den planlagte
 - System-selektionsproblemer (creaming)
 - Særligt engageret/dygtigt personale ved forsøg
- Målefejl
 - Statistiske problemer, validitets- og reliabilitetsproblemer
- Generaliseringsproblemer





VURDERING AF FORSKELLIGE STUDIETDESIGNS

Forskellige effektdesigns lægger forskellig vægt på betydningen af forstyrrende faktorer og har forskellige metoder til at fjerne denne.

Kvaliteten af studier udført med forskellige metoder skal derfor vurderes forskelligt:

- Randomiserede kontrollerede forsøg (RCT)
- Quasi-randomiserede forsøg
- Quasi-eksperimenter
- Kvalitative studier

GENERELLE OVERVEJELSER

- For at sige noget om effekten må der nødvendigvis være et kontrolgruppedesign
- Mange primærstudier indeholder mangelfuld dokumentation i forhold til vurdering af den interne validitet
- Er resultaterne generaliserbare til andre settings?

Intern validitet: Måler undersøgelsen det, den har til hensigt at måle?

Giver undersøgelsen et ikke-biased estimat på den kausale effekt af intervention?

RANDOMISEREDE KONTROLLEREDE FORSØG (I)

Samme population udgør både indsatsgruppe og kontrolgruppe

Tilfældig lodtrækning afgør, om den enkelte person kommer i indsatsgruppe eller kontrolgruppe (fx kast med terning)

Statistisk er indsatsgruppen og kontrolgruppen derfor "ens" på alle områder, på nær indsats eller ej

Benævnt "guldstandard" for effektstudier, men ...

RANDOMISEREDE KONTROLLEREDE FORSØG (II)

Randomiserede forsøg er i princippet enkle, men i praksis vanskelige. Fx:

- Blev randomiseringen implementeret?
- Er der bortfald fra indsatsgruppen?
- Er der bortfald fra kontrolgruppen?
- Er der kontrol over, hvad kontrolgruppen foretager sig?

QUASI-RANDOMISEREDE FORSØG

Forsøg hvor indsats- og kontrolgruppe trækkes fra samme population, men hvor fordelingen til hhv. indsats- og kontrolgruppe ikke foregår ved lodtrækning

Fx fødselsdato, deltageres præferencer, professionel vurdering

Kræver nøje vurdering af, om fordelingen på grupper (selektionen) er korreleret med effekten af indsatsen

QUASI-EKSPERIMENTELT DESIGN (I)

Dækker over studier, der ikke bygger på forsøg, men på forskellige typer observationsstudier:

Survey data

Registerdata

Paneldata

Før- og eftermålinger

Økonometriske metoder anvendes til estimere kausale effekter

QUASI-EKSPERIMENTELT DESIGN (II)

NATURLIGE EKSPERIMENTER:

Hvor en indsats bliver indført på et givent tidspunkt for en delpopulation, men ikke for en anden delpopulation (fx i et geografisk område).

"Naturen" har dermed lavet en indsatsgruppe og en kontrolgruppe.

Difference-in-difference (DID) estimation valid, hvis uobserverbare faktorer er ens for de to grupper over tid.

QUASI-EKSPERIMENTELT DESIGN (III)

MATCHING:

Personer, der har modtaget en indsats, matches med personer, der ikke har modtaget indsatsen, for på den måde at skabe en kunstig kontrolgruppe.

Kræver for det første, at der kan matches på så mange variable, at der ikke er uobserverbar selektion mellem grupperne.

Kræver for det andet, at der er *common support*, dvs. at hver person i indsatsgruppen kan matches med mindst en person i kontrolgruppen.

KVALITATIVE STUDIER

Kvalitative studier kan i sagens natur ikke bruges til at vurdere effektstørrelser.

Men kvalitative studier kan supplere til forståelsen af kvantitative effekter, fx ved at analysere hvilke elementer af en given indsats, der virker hvordan.

Vi har ikke nogen standardredskaber til kvalitetsvurdering af kvalitative studier – endnu.

EVIDENSTYPOLOGI VS. EVIDENSHIERARKI

- Evidenstypologien søger at opstille mønster for, hvilke typer af studiedesigns der er gode til at svare på bestemte typer spørgsmål.
- Evidenshierarkiet for effektkvalitetsvurdering opstiller en rangorden for, hvor gode forskellige designs *generelt* er til at svare på én bestemt type spørgsmål.
 - Der behøver ikke at være modstrid mellem de to vinkler
 - Men evidenshierarkiet er for snævert og bør snarere være en vifte af muligheder

EVIDENSTYPOLOGI

Design: Research question:	Qualita- tive research	Survey	Case control studies	Cohort studies	RCTs	Quasi- experi- mental studies
Effectiveness Does this work? Does doing this work better than doing that?				+	++	+
Process of service delivery How does it work?	++	+				
Salience Does it matter?	++	++				
Safety Will it do more good than harm?	+		+	+	++	+
Acceptability Will children/parents be willing to or want to take up the service offered?	++	+			+	+
Cost-effectiveness Is it worth buying this service?					++	
Appropriateness Is this the right service for these children?	++	++				
Satisfaction with the service Are users, providers and other stakeholders satisfied with the service?	++	++	+	+		

“EVIDENSHIERARKI”

- Lodtrækningsforsøg
- Andre eksperimentelle studier
- Studier, der eksplicit kontrollerer for andre påvirkninger
- Før- og eftermålinger
- Eftermålinger
- ...

EVIDENS STUDIETDESIGN MULIGHEDER

- ❖ Lodtrækningsforsøg
- ❖ Quasi-randomiserede forsøg
- ❖ Matching modeller
- ❖ Naturlige eksperimenter
- ❖ (Kvalitative studier)
- ❖ ...

En liste over designs, der alle skal vurderes

KONKLUSION

Hvis vi gerne vil vide noget om effekten af sociale indsatser, kan vi ikke nøjes med at se på (randomiserede) kontrollerede forsøg

Inden for forskellige studiedesigns skal primær studier kvalitetsvurderes i forhold til den specifikke metode

Som det er nu, er det lettest at kvalitetsvurdere eksperimentelle designs, men vi skal hele tiden udvikle dette

TAK FOR OPMÆRKSOMHEDEN