

07:08

Connie Nielsen
Peter Thisted Dinesen
Lars Benjaminsen
Jens Bonke

EFFEKT MÅLING



07:08

EFFEKT MÅLING

Connie Nielsen
Peter Thisted Dinesen
Lars Benjaminsen
Jens Bonke

KØBENHAVN 2007
SOCIALFORSKNINGSINSTITUTTET

EFFEKT MÅLING

Afdelingsleder: Ole Gregersen
Afdelingen for Socialpolitik og velfærdsydelser

ISSN: 1396-1810
ISBN: 978-87-7487-850-6

Layout: Hedda Bank

© 2007 Socialforskningsinstituttet

Socialforskningsinstituttet
Herluf Trolles Gade 11
1052 København K
Tlf. 33 48 08 00
sf@sf.dk
www.sf.dk

Socialforskningsinstituttets publikationer kan frit citeres med tydelig angivelse af kilden. Skrifter, der omtaler, anmelder, henviser til eller gengiver Socialforskningsinstituttets publikationer, bedes sendt til instituttet.

INDHOLD

	FORORD	5
	RESUMÉ	7
	Tankegangen bag effektmåling	7
	Forskellige effektmålingsmetoder	9
	Hvordan kan effekter måles i praksis?	9
1	FORMÅL MED EFFEKTMÅLING	11
	Evidensbaserede indsatser	13
	Et eksempel på usikkerhed om evidens	14
	Hvornår skal vi gennemføre effektmålingsstudier?	15
2	HVAD ER EFFEKTMÅLING?	19
	Procesevaluering og effektmåling	19
	Hvordan måler man effekter?	21
	Er effekter individuelle?	25
	Hvem skal tilbydes deltagelse?	27
3	SELEKTIONSSKÆVHED	29
4	EKSPERIMENTELLE METODER	35

	Intern og ekstern validitet af randomiserede forsøg	35
5	IKKE-EKSPERIMENTELLE METODER	39
	Statistisk kontrol – regression	40
	Matching	44
	Instrument Variabel (IV) estimation	46
	Quasi-eksperimentelle metoder	50
	Valg af effektmålingsmetode	52
	Begrænsninger ved Effektmålingsmetoder	54
6	KLASSIFIKATION AF SATSPULJEPROJEKTER	57
7	TO EKSEMPLER PÅ SATSPULJEPROJEKTER	61
	Indsatsen over for kriminelle unge, som opholder sig på sikrede institutioner	61
	Handlingsplan for forebyggelse på ældreområdet – forebyggende hjemmebesøg	66
8	EFFEKT MÅLING AF 06-SATSPULJEPROJEKTER	71
9	IMPLEMENTERING AF EFFEKT MÅLING PÅ DET SOCIALE OMRÅDE	85
	Barrierer for implementering	85
	Anbefalinger	90
	LITTERATUR	93
	SFI-RAPPORTER SIDEN 2006	97

FORORD

Der er en stigende opmærksomhed omkring, hvordan forskellige tiltag på det sociale område virker. Derfor er der et politisk ønske om at forbedre vores muligheder for at vurdere, om de enkelte tiltag har en positiv virkning på brugerne. Det betyder, at det ikke er tilstrækkeligt at vide, om tiltagene er sat i værk, og om de involverede parter udtrykker tilfredshed med dem. Det er også ønskeligt at vide, om brugerne som følge af det konkrete tiltag er blevet bedre stillet.

På det grundlag blev Socialforskningsinstituttet af Socialministeriet i 2005 bl.a. bedt om at udarbejde en effektmålingsmanual. Formålet var at udvikle et redskab, der gør det muligt for Socialministeriets medarbejdere konkret at vurdere, om forskellige tiltag kan og bør effektmåles, idet udgangspunktet var en stillingtagen til Satspuljeprojekterne for 2006 og 2007.

I forbindelse med udarbejdelsen af manualen er forslag og udkast hertil blevet fremlagt på møder i Socialministeriets Effekt Taskforce vedrørende evalueringer på det sociale område. Der er tillige afholdt en række møder med Effekt- og Dokumentationsprojektets sekretariat, bestående af Eva Husum Schmidt, Signe Tholstrup og Carsten Skovgaard Nielsen, og Effekt- og Dokumentationsprojektets ansvarlige chef Aksel Meyer, som ud over konstruktiv kritik også har bidraget med forslag til projekter til afprøvning af manualen.

Manualen er udarbejdet af forsker Connie Nielsen, som har stået for kapitlerne om effektevalueringsmetoderne, forsker Lars Benjaminsen, som har bidraget med viden om evalueringer af satspuljeprojekter, forskningsassistent Peter Dinesen, som har udarbejdet kapitlet om implementering af effektmålingsresultater, og programleder Jens Bonke, som har ledet undersøgelsen. Samtlige deltagere har kommenteret hinandens bidrag og har dermed haft indflydelse på udformningen af manualen. Under udarbejdelsen af manualen er der blevet afholdt seminarer med professor Thomas Crossley, University of Cambridge, og senior fellow Charles Michalopoulos, Manpower Demonstration Research Corporation, som bl.a. har fortalt om canadiske og amerikanske erfaringer med effektevalueringer. Adjunkt Marianne Simonsen, Århus Universitet, har været referee på manualen og har givet værdifulde kommentarer og kritik. Manualen er finansieret af Socialministeriet.

København, marts 2007

Jørgen Søndergaard

RESUMÉ

Denne manual udgør et forsøg på at opbygge et fælles grundlag for teoretikere og praktikere til at tænke effektmåling ind i tiltag på det sociale område. Manualen giver et overblik over de metoder, det er muligt at anvende, når man ønsker at udforme et effektmålingsdesign. Ved at kende disse metoder vil det i højere grad være muligt at afklare, hvilken indsats der virker bedst. For hver enkelt indsats er det ideelt, hvis man kan finde frem til, dels om effektmåling vil være nyttig, dels hvilken effektmålingsmetode der så vil egne sig godt. Ved at få gode idéer til et godt gennemtænkt effektmålingsdesign vil det være nemmere at isolere effekten af en indsats. Det kan hjælpe til at opbygge en evidensbaseret viden om, hvilken indsats der virker bedst. Det er således håbet, at denne manual kan bidrage til, at vi begynder at opbygge en tradition for effektmålinger på det sociale område i Danmark.

TANKEGANGEN BAG EFFEKTMÅLING

Der findes flere måder at måle effekter på, og manualen giver en oversigt over de forskellige metoder. Effektmålingsmetoderne bygger grundlæggende på tankegangen om den kontrafaktiske situation. Når en person har deltaget i en indsats, er det fx vigtigt at vide, hvad den kontrafaktiske

situation ville have været, dvs. hvad der ville være sket, hvis samme person ikke havde deltaget i den pågældende indsats.

Når der er fundet en tilfredsstillende måde at måle den kontrafaktiske situation på, kan det bruges til at sammenligne med det faktisk observerede resultat. Det svarer til at sammenligne en deltager- og en ikke-deltagergruppe, hvor det kontrafaktiske resultat fx kan bruges som et udtryk for resultatet ved ikke at deltage, hvis man kender det faktisk observerede resultat for en deltager.

Når en deltagergruppe sammenlignes med en ikke-deltagergruppe, er det endvidere vigtigt at vurdere, om personerne er forskellige. Hvis alle personer, der deltager, er ens, vil der være tale om homogene deltageseffekter, og da er det forholdsvist enkelt at beregne en gennemsnitlig effekt af at deltage. Er personerne derimod forskellige, er der tale om heterogene deltageseffekter, og det er meget sandsynligt, at den gennemsnitlige effekt af deltagelse blandt deltagerne er forskellig fra den gennemsnitlige effekt af deltagelse for gruppen af ikke-deltagere.

Hvis personerne er forskellige, kan det også føre til overvejelser om, hvem der skal tilbydes at deltage i en given indsats. Det kan fx ske ud fra en politisk prioritering eller ud fra en betragtning om, hvad der økonomisk set er mest efficient, hvilket svarer til at fokusere på de personer, der opnår det bedste resultat. Det kan tænkes, at der også bevidst spekuleres i, at indsatsen gives til de personer, der opnår et relativt godt resultat, for at opnå gode effekter, hvilket i litteraturen benævnes *cream skimming*.

Det forhold, at personer kan være forskellige, fører også til begrebet om selektionsskævhed. Det bør således overvejes nøje, om der er systematiske forskelle mellem de personer, der deltager i en indsats, og de personer, der ikke deltager. I manualen vises det, hvordan det er muligt at få et udtryk for selektionsskævheden. Det fremkommer som forskellen i resultatet af ikke at deltage set ud fra en deltager, dvs. den kontrafaktiske situation, og i det faktiske resultat fra en tilsvarende ikke-deltager. Senere diskuteres det, hvordan de enkelte effektmålingsmetoder kan korrigere for denne selektionsskævhed.

FORSKELLIGE EFFEKTMÅLINGSMETODER

Manualen giver et overblik over de metoder, der kan anvendes til at måle effekten af en indsats. Disse metoder kan overordnet set inddeles i de eksperimentelle og de ikke-eksperimentelle metoder.

De eksperimentelle metoder omfatter randomiserede forsøg, hvor der trækkes tilfældigt lod blandt en gruppe af personer til bestemmelse af, om de skal tilhøre en deltager- eller ikke-deltagergruppe. Det gør, at der i eksperimentets design i sig selv ikke skulle være selektionsproblemer, hvorfor der ofte refereres til randomiserede forsøg som ”guldstandarden”. Der kan dog også opstå problemer, når denne type af metoder anvendes, hvilket diskuteres i manualen.

De ikke-eksperimentelle metoder udnytter de tekniske færdigheder, der er opbygget inden for kvantitative metoder i løbet af de sidste årtier. Det kan fx ske på baggrund af registerdata, surveys eller andre eksisterende data. Det er også muligt at kombinere de ikke-eksperimentelle metoder med et eksperimentelt design. De ikke-eksperimentelle metoder omfatter regression, matching, instrument variabel estimation (IV) og quasi-eksperimentelle metoder. Når der vælges mellem de enkelte metoder, er det fx afgørende, hvordan de kontrafaktiske resultater konstrueres, og hvordan de estimerede effekter af deltagelse aggregeres og sammenlignes.

HVORDAN KAN EFFEKTER MÅLES I PRAKSIS?

For at systematisere de grundlæggende tankegange, der er beskrevet for de enkelte effektmålingsmetoder, er der i manualen udfærdiget et klassifikationsskema. Det er tænkt som et hjælperedskab til at finde et godt konkret effektmålingsdesign for en given indsats. Det er således tanken, at man med et bestemt satspuljeprojekt i tankerne kan arbejde sig igennem skemaet og se, hvad der er vigtigt at huske, når man ønsker at bestemme, hvilken effektmålingsmetode der vil egne sig godt.

Det er også muligt at udnytte den foreliggende viden om effektmålinger til bedre at kunne præcisere, hvordan et satspuljeprojekt ønskes beskrevet. Med udgangspunkt i to idéer til satspuljeprojekter for satspuljeforhandlingerne i efteråret 2006 gennemgås, hvordan projekterne er defineret, og der kommer med forslag til, hvad der kan gøre dem

mere målrettede til at måle effekter. Det første projekt er indsatsen over for kriminelle unge, som opholder sig på sikrede institutioner, og det andet projekt er en handlingsplan for forebyggelse på ældreområdet – forebyggende hjemmebesøg.

Herudover gennemgås samtlige satspuljeprojekter med henblik på en overordnet klassificering af, om projektet er egnet til effektmåling eller ej. Hvis projektet anses for egnet til effektmåling, markeres det, om det umiddelbart er mest hensigtsmæssigt med et eksperimentelt eller et ikke-eksperimentelt design.

Den evidensbaserede politik på det sociale område og effektmålingsmetoderne ligger fjernt fra mange af de traditionelle metoder, der er anvendt inden for det sociale område. Men for at udnytte diverse effektmålingsmetoder i praksis anbefaler manualen fx at indgå i en dialog med uddannelsesinstitutioner for socialarbejdere, at øge dialogen mellem forskere og praktikere på det sociale område og i det hele taget at få øget fokus på betydningen af evidensbaserede tiltag på det sociale område.

FORMÅL MED EFFEKTMALING

I de senere år er der sket en betydelig udvikling af forskellige typer af sociale tilbud til socialt svage og udsatte grupper. Særligt er der gennem satspuljemidlerne igangsat et stort antal sociale projekter ofte på forsøgsbasis og for en tidsbegrænset periode. Gennem evalueringer ved vi en del om indholdet og forløbet af disse projekter, hvorimod vi kun ved lidt om de opnåede effekter for brugerne af disse tilbud. Det hænger bl.a. sammen med, at der i Danmark generelt ikke er tradition for at gennemføre egentlige effektmålingsstudier af sociale indsatser.

Effektmålingsstudierne har til formål at give en viden om, hvorvidt der rent faktisk kommer de tilsigtede effekter ud af den indsats, der gennemføres. Overordnet sigter effektmålinger således mod at besvare spørgsmål som:

- Hvilken effekt har indsatsen for brugerne sammenholdt med, hvis brugerne ikke havde modtaget indsatsen?
- Hvad betyder selektion af brugere af indsatsen for de effekter, vi kan iagttage?

Effektmålinger gør det muligt at sammenholde effekterne af forskellige indsatser, så der efterhånden fremkommer et mønster af, hvorvidt bestemte indsatstyper er mere hensigtsmæssige at anvende end andre.

Ved effektmålingsstudiet forsøger vi at opnå en sikkerhed for, at det rent faktisk er indsatsen og ikke andre forhold, der medfører forandringen. Med effektmålingsstudiet udvider vi således vores fokus fra det givne og observerbare til den kontrafaktiske situation og dermed spørgsmålet om, hvad der ville være sket, hvis vi ikke havde gennemført indsatsen. Effekstudiet sigter således mod at specificere den kausale effekt af indsatsen, så det kan sandsynliggøres, at det er den pågældende indsats, der isoleret set har ført til den opnåede effekt.

Da der i Danmark først i de senere år er opstået politisk interesse for mere præcist at kunne måle effekter, er det både nødvendigt at indføre arbejdsgange, så de nødvendige informationer tilvejebringes, og at få erfaring med at analysere effekter. Ved at opbygge en tradition for at udføre effektmålingsstudier i Danmark vil det være muligt at forbedre beslutningsgrundlaget for tildeling af midler til sociale projekter.

I den angelsaksiske litteratur er der i løbet af de sidste 40 år udviklet metoder til at måle effekter, der formentlig mest hyppigt går under betegnelsen impact evaluation. Denne udvikling har fundet sted, da det har været anset for nødvendigt at forbedre det evidensbaserede grundlag for at beslutte, hvilke politiske tiltag der virker bedst.

Selvom der i udlandet allerede er en længere tradition for at gennemføre effektmålinger af politiske tiltag og initiativer, er der stadig udviklingsmuligheder inden for dette felt. Det fremgår for eksempel af Cobb-Clark & Crossley (2003):

While academic researchers are increasingly focused on assessing the strengths and weaknesses of the evaluation methodology itself, policy makers are turning to the results to provide the foundations for evidence-based policy.

Den nævnte artikel giver et godt overblik over, hvad der i dag er standard inden for effektmålingsstudier, hvis man ønsker at gå i gang med at gennemføre effektmålingsstudier. Vi har bl.a. ladet os inspirere af artiklen i den følgende fremstilling. Vi ønsker, at denne fremstilling kan forstås af læsere, der ikke har specielle forkundskaber i effektmåling.

I denne manual om effektmåling vil vi formidle tankegangen i de forskningsbaserede metoder. Vi gennemgår først de enkelte effektmålingsmetoder for ud fra dem at kunne udvikle mere praksisorienterede redskaber. Det sker dels i form af at opstille en oversigt over de væsentli-

ge forskelle mellem de enkelte effektmålingsmetoder, dels i form af at udvikle et klassifikationsskema. Dette klassifikationsskema er tænkt som et hjælperedskab til at finde et godt effektmålingsdesign, når man sidder med et konkret satspuljeprojekt. Ved at arbejde sig systematisk igennem skemaet og se, om forskellige kriterier er opfyldt, vil det til sidst føre til et konkret forslag til et effektmålingsdesign eller evt. en konklusion om undervejs i klassifikationsskemaet, at effektmåling ikke er egnet.

I forbindelse med projektet, der ligger til grund for udfærdigelsen af denne manual, har udgangspunktet været en stillingtagen til Satspuljeprojekterne for 2006 og 2007. Disse projekter har selvfølgelig udgjort en stor inspirationskilde til, hvordan det er muligt at udnytte den teoretiske viden i praksis for et konkret projekt, hvor man ønsker at gennemføre et effektmålingsstudie, hvorfor de også nævnes i manualen. Vi har imidlertid bestræbt os på at gennemgå de enkelte effektmålingsmetoder, så de kan forstås ud fra deres grundlæggende præmisser. Det skyldes, at effektmålingsmetoderne har en generel interesse for projekter, der iværksættes på det sociale område.

EVIDENSBASEREDE INDSATSER

At have en evidensbaseret tilgang er det samme som at samle hidtil dokumenteret viden, hvilket også kaldes evidens, for en politisk beslutning. Der kan være mange typer af evidens, der tilsammen udgør et evidensbaseret grundlag for at føre politik (Davies et al., 2006). Effektmålingsstudier kan ses som en form for evidens, hvor der som tidligere nævnt er rige muligheder for at udvide vores viden i Danmark.

Effektmålingsstudier kan bidrage til, at vi får opbygget en mere klar struktur over de forventede sammenhænge mellem indsatser og effekter, så vi opnår en systematisk viden om, hvordan forskellige indsatsformer virker. Denne systematiske viden refererer vi til som det evidensbaserede grundlag. Da det langt fra altid er eksakt, hvad der virker, og hvad der ikke virker inden for det sociale område, kan de forventede sammenhænge mellem indsatser og effekter, der opbygges i det evidensbaserede grundlag, ses som hypoteser.

Fremgangsmåden svarer til den traditionelle inden for naturvidenskab, hvor der opstilles hypoteser om, hvordan de naturvidenskabelige spørgsmål skal forstås. Sættet af hypoteser udgør en naturvidenskabe-

lig forklaring, så længe de ikke er modbevist ved fx nye eksperimentelle forsøg.

Inden for det sociale område kan vi også tænke os, at en hypotese fungerer, så længe den ikke bliver udfordret. For eksempel kan det tænkes, at vi for en periode er tilfredse med det evidensbaserede grundlag, der ligger til grund for vores indsatser. Dette kan også formuleres sådan, at det ikke har været muligt at afvise eksisterende hypoteser, der ligger til grund for det eksisterende evidensbaserede grundlag. Først når der opstår usikkerhed om effekten af en indsats, vil det være særligt interessant at designe effektmålingsstudier. Når der mangler viden om, hvad der virker bedst, eller der opstår ny viden om indsatser, der kan tænkes at virke bedre, vil vi gerne kunne dokumentere, hvilken indsats der virker bedst. På den måde kan vi løbende udfordre den viden og de hypoteser, der udgør det evidensbaserede grundlag, så vi får en stadigt klarere dokumentation af effekterne af diverse indsatser på det sociale område.

ET EKSEMPEL PÅ USIKKERHED OM EVIDENS

For at illustrere hvordan vi kan kvalificere en effektorienteret problemstilling ved at kombinere viden fra praksis med teoretisk viden, vil vi nu inddrage et eksempel fra indsatsen for de socialt udsatte i forbindelse med hjemløshed. På dette område er det nemlig stadig usikkert, om den bedste indsatsstrategi er at sørge for, at borgere først og fremmest får stillet en bolig til rådighed, inden der sker en social indsats, eller om det skal indgå som en del af et program, hvor det at få en bolig er et mål, man skal kvalificere sig til.

Inden for den teoretiske litteratur sondres der mellem såkaldte housing first- og housing ready-strategier. Housing first-tilgangen lægger vægt på, at der skal ske en tidlig etablering af en permanent boligløsning, før der kan arbejdes med fx misbrugsproblemstillinger o.l., mens housing ready-strategien består i en trappetrinsmodel, der rummer et element af kvalifikation eller fremskridt i behandlingen, før det vurderes, om det er realistisk, at der kan etableres en permanent boligløsning. En hypotese i den internationale forskning på området er, at housing first-strategien med sit helhedsfokus på spillet mellem boligsituationen og den sociale situation bidrager til højere succesrater, når det fx gælder om at fast-

holde hjemløse i bolig og at afhjælpe misbrugsproblematikker (jf. Tsemberis et al., 2004).

I praksis er der i Danmark i de senere år sket en ekspansion i antallet af botilbud. Det er sket, selvom vi ikke har en systematisk viden om, hvilke indsatser der virker bedst over for brugerne fx set i forhold til en målsætning om på længere sigt at kunne etablere og fastholde den hjemløse i eget hjem eller at opnå en øget livskvalitet. Det er således også tidligere påvist, at der i Danmark er en regional variation mellem kommunerne, når det gælder forskellige typer af sådanne tilbud (Fabricius et al., 2005). Hvor man i nogle kommuner primært satser på indsatsformer, der inkluderer en permanent løsning af den hjemløses boligproblem relativt tidligt i et indsatsforløb, har man i andre kommuner lagt vægten på indsatser efter trappetrinsmodellen. Der er endvidere kommuner, hvor der kan identificeres strategier og indsatser, der i højere grad ligger imellem de to idealtyper af indsatser, som teorien definerer. Vi ved imidlertid ikke nok om, hvilken af de to strategier – housing first eller housing ready – der virker bedst for brugeren på sigt. Her kommer spørgsmålet om en systematisk viden om effekterne af indsatsen altså ind.

HVORNÅR SKAL VI GENNEMFØRE EFFEKT MÅLINGSSSTUDIER?

Det er langt fra altid hensigtsmæssigt eller muligt at gennemføre effektmålingsstudier i forbindelse med sociale indsatser. Der er således brug for at skærpe opmærksomheden på, i hvilke tilfælde det er relevant og værd at gennemføre.

Det er vigtigt, at effektmålingsstudiet ikke blot anlægger en rent empirisk tilgangsvinkel, men også inddrager teoretiske overvejelser, da teorien bidrager til at identificere og kvalificere relevante effektmålingsstudier. Effektmålingsstudier handler således ikke blot om at kunne identificere, hvor stor effekten af en given indsats er, men også om at kunne identificere, hvorfor bestemte effekter indtræffer, og hvilke omstændigheder der kan bidrage til at styrke eller svække effekterne af en given indsats. Her bidrager den eksisterende litteratur til, at vi kan opstille hypoteser om effekter af indsatsen på det sociale område. Når vi skal identificere et muligt effektmålingsstudie, kan konkrete praktiske problemstillinger kvalificeres gennem inddragelsen af social teori. Gennem inddra-

gelsen af teori kan det fx være muligt at sondre mellem forskellige typer af indsatser, som i det tidligere eksempel med de forskellige boligstrategier, *housing first* og *housing ready*. Teorien kan bidrage til at afgrænse bestemte typer af indsatser, som vi ønsker at sammenligne med andre indsatser med henblik på at afdække, hvilken indsats der virker bedst, og til at specificere, hvilke sammenhænge vi antager, at der er mellem indsatser og effekter.

Hvis et effektmålingsstudie først synes velbegrunderet i forhold til at bidrage med en relevant viden, der kan styrke det evidensbaserede grundlag for indsatsen, er det også nødvendigt at tage hensyn til, at det at gennemføre effektmålingsstudier kan være omkostningskrævende. Hvis man har fundet frem til, at det er nyttigt at gennemføre et effektmålingsstudie, er det derfor også nødvendigt at undersøge, om det vil være penge værd at gennemføre det. Et relevant spørgsmål er således, hvad det vil koste at gennemføre et konkret effektmålingsdesign. Som regel vil det være nødvendigt at afveje, hvor omstændeligt og dyrt det vil være at gennemføre i forhold til, hvad der kan komme ud af det.

Det kan også være, at der er usikkerhed om, hvilket resultat der kan opnås, hvis et givet effektmålingsdesign gennemføres. Denne situation er særlig udpræget her i begyndelsen, hvor vi ikke har de store erfaringer med at gennemføre effektmålingsstudier i Danmark. Det kan derfor i nogle tilfælde være en god idé at forsøge sig frem i mindre skala i form af pilotprojekter. I stedet for at begynde med at indhente oplysninger for alle personer ved fx at gennemføre interview, kan man overveje at begynde med at interviewe et færre antal personer. Så vil man kunne vurdere, om det på det sociale område i det hele taget kan vurderes, om det er muligt at indhente den information, der er nødvendig, og hvordan man kan tænkes at gennemføre en analyse på baggrund af den pågældende information.

Det er også værd at overveje, om der i praksis kan identificeres variationer i indsatsen. Det kan fx tænkes, at to ellers sammenlignelige kommuner har hver deres praksis, så det er muligt at gennemføre et effektmålingsstudie, der tilnærmelsesvist kan give en idé om, hvorvidt den ene indsats virker bedre end den anden. I nogle tilfælde vil der være et tilstrækkeligt grundlag for at gennemføre et egentligt effektmålingsstudie på dette grundlag, og i andre tilfælde kan det udgøre et forstudie, der gør, at det bliver muligt at gennemtænke et bedre effektmålingsdesign.

Det kan være fordelagtigt, at effektmålingsdesignet er godt gennemtænkt på forhånd, hvis det fx er dyrt at tilvejebringe nye data.

Endelig vil det være fornuftigt at overveje, om resultaterne fra et effektmålingsstudie kan tænkes at have en mere generel interesse for andre typer af projekter, da det så kan være værd at satse flere ressourcer på det. I forlængelse heraf er det også værd at overveje, om det er relevant at udbrede den viden om effekter, der erhverves, til en bredere kreds af personer, fordi det kan tænkes at have almen interesse i befolkningen.

HVAD ER EFFEKTMÅLING?

PROCESEVALUERING OG EFFEKTMÅLING

I denne manual vil vi sondre mellem procesevaluering og effektmåling, der er to metoder med hver deres berettigelse, når det gælder om at tilvejebringe vigtig information til beslutningstagere.¹ Hvert af begreberne kan forstås bredt, idet procesevaluering kan dække eftermålinger, før- og eftermålinger, *performance management* og resultatbaseret styring, og effektmåling kan indbefatte flere slags metoder, som vi vil beskrive i denne manual.² I dette afsnit vil vi kort sammenholde effektmåling og procesevaluering, hvorefter vi udelukkende vil fokusere på effektmåling, så læseren får en idé om og et overblik over, hvad der på nuværende tidspunkt er muligt på dette område med de teknikker, som vi kender.

Procesevaluering omfatter metoder til at afgøre, om et politisk tiltag er implementeret, som det var hensigten. I forbindelse med denne type evaluering bruges ofte udtrykket *accountability*, der dækker over det, at det arbejde, der udføres, er i overensstemmelse med de aftalte regler og standarder for projektet. Det kan registreres ved, at det udførte

¹ Vi har valgt at bruge de danske betegnelser procesevaluering og effektmåling om det, der i den engelsksprogede litteratur ofte kaldes hhv. *process evaluation* og *impact evaluation* (Angrist, 2004).

² Performance management er kursiveret, da det er den betegnelse, vi synes, at vi hører mest i Socialministeriet.

arbejde afrapporteres så præcist og retfærdigt som muligt i forhold til planer og opdrag for projektet (Kusek & Rist, 2004).

Det, som vi i denne manual kalder effektmåling, er baseret på metoder til at måle, hvor godt målsætningerne med det politiske initiativ er opfyldt. Inden for effektmåling vil man typisk forsøge at finde den kausale sammenhæng mellem en indsats og effekten af denne. Det er vigtigt at isolere effekten af den konkrete indsats, da det fx kan tænkes, at et givet resultat er opnået, hvor også andre faktorer end den konkrete indsats har ført til dette resultat. Der er således studier, der påviser, at der ikke altid er den sammenhæng mellem resultat og indsats, som man skulle tro. Inden for procesevaluering vil man ofte være nødt til at antage, at det er indsatsen, der fører til resultatet.

Forskellen mellem procesevaluering og effektmåling kan illustreres med et eksempel, der tager udgangspunkt i en syg patient. Antag, at vi ønsker at evaluere, om et givet behandlingsforløb vil gavne en syg patient. I dette tilfælde vil procesevaluering gå ud på at styre behandlingsforløbet, så patienten får de behandlinger, der er tiltænkt til de rigtige tidspunkter, mens effektmåling mere vil fokusere på, om denne behandlingstype gør, at patienten fx vil kunne leve længere.

Der kan også tænkes at være utilsigtede effekter af en indsats. Da kan det diskuteres, hvordan metoderne hver for sig bedst er i stand til at indfange disse. Procesevaluering følger som regel et på forhånd udvalgt antal indikatorer for at undersøge, om de opfylder det ønskede resultat. Disse indikatorer vil ikke altid fange utilsigtede effekter. Til gengæld vil man under en procesevaluering i højere grad end i et effektstudie have kontakt til projektmagerne, så der er mere tilknytning til, hvad der sker i praksis.

Forhåbentlig har beskrivelsen i dette afsnit allerede på et overordnet plan klargjort berettigelsen af både procesevaluering og effektmåling, hvor effektmålingen kan ses som en udvidelse af de teknikker, der i dag er mulige i forhold til de resultater, der kan opnås ved hjælp af procesevaluering alene. Det er klart, at de teknikker, der anvendes inden for effektmåling, kan være forholdsvis abstrakte til at begynde med. Forhåbentlig vil manualen her formå at afhjælpe en del af disse problemer. At effektmåling udvider metodeapparatet i forhold til procesevaluering, har vi set anskueliggjort flere gange ved hjælp af den såkaldte evidensstige (Konnerup, 2006). Evidensstigen angiver, hvordan forskellige evaluer-

ringsmetoder kan indplaceres i forhold til hinanden ud fra et ønske om at opnå den bedst mulige evidens for virkningen af en indsats.

HVORDAN MÅLER MAN EFFEKTER?

Når det drejer sig om effektmåling, kan der være mange måder at gribe forskellige problemer an på. Der er imidlertid nogle bestemte tankegange, der er nyttige at kende for at kunne afgøre, hvilket effektmålingsdesign der vil være egnet i et konkret tilfælde. Derfor vil vi i denne manual fokusere på disse tankegange. Det vigtigste overordnede aspekt i forsøget på at måle effekten er at sammenligne resultatet af et givet program eller en given politik med en situation, hvor man forestiller sig, at der ikke er igangsat noget, hvilket i litteraturen kaldes den kontrafaktiske situation. I Angrist et al. (2004) beskrives dette således:

The causal effect of a treatment on a single individual or unit of observation is the comparison (e.g. difference) between the value of the outcome if the unit is treated and the value of the outcome if the unit is not treated.

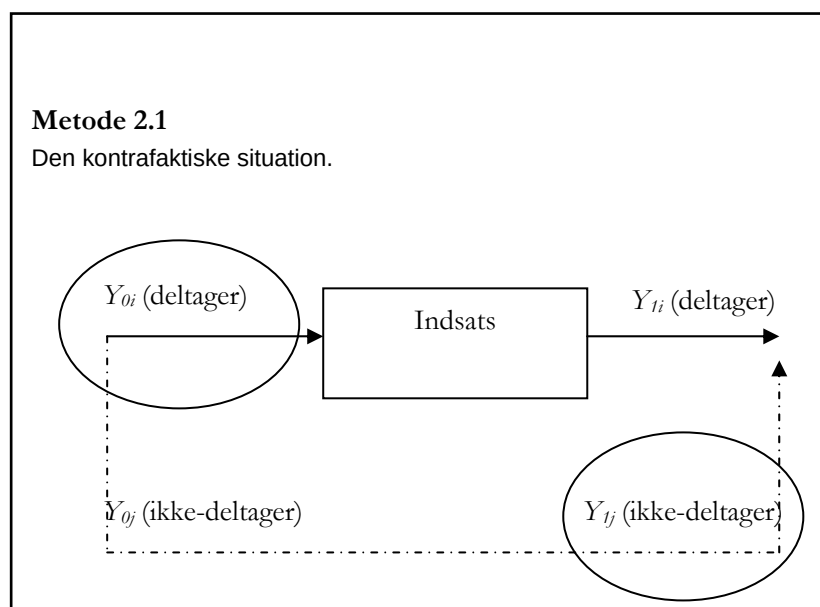
Hvis *treatment* oversættes direkte til ”behandling” på dansk, får det ofte en lidt uheldig klang, hvorfor vi i stedet for har valgt at kalde det ”deltagelse i en indsats”.

Outcome har vi valgt at oversætte til ”resultat”, da det er et almindeligt ord på dansk. Det ligger dog nok mere tydeligt i ordet på engelsk, at man også skal tænke det som et udfald.

The causal effect of a treatment vil vi i teksten beskrive som forskellen mellem resultatet for en deltager- og en ikke-deltagergruppe.

Siden 1938 har man i lægeverdenen systematisk forsøgt at teste virkningen af forskellige medicinalprodukter. Det er typisk forsøg, hvor en gruppe af patienter får det nye medicinalprodukt, mens en tilsvarende gruppe af patienter får en placebotablet, der ikke har nogen virkning. Ved disse forsøg bruges ofte betegnelsen kontrolgruppe for det, vi har valgt at kalde ikke-deltagergruppe. Vi har valgt at holde os til ikke-deltagergruppe, idet overførslen af tankegangen til det sociale område gør, at det bliver et mere bredt anvendelsesområde.

Tankegangen om den kontrafaktiske situation er illustreret i figuren i metode 2.1:



I metode 2.1 skitseres to situationer, hvor der enten tages udgangspunkt i en deltager eller i en ikke-deltager, der hhv. svarer til den øverste fuldt optrukne linje og den nederste stiplede linje. Det skal bemærkes, at der altid sammenlignes langs en af disse linjer, så vi enten ser det ud fra en person, der deltager, som vi indekserer med i , eller ud fra en person, der ikke deltager, som vi indekserer med j . I begge tilfælde er *baseline*-scenariet angivet ved et nul, og det at deltage i en indsats er angivet ved et et-tal. Vi har bestræbt os på at tænke på nul-indekset som *baseline*-scenariet i stedet for at tale om udgangspunktet før deltagelse. Det skyldes, at forskellen mellem resultatet Y_{0i} og Y_{1i} i princippet ikke har noget med tid at gøre, således at læseren må være varsom med at tænke i før- og eftermålinger. For eksempel indebærer tankegangen med den kontrafaktiske situation i sig selv, at der ikke er tale om en tidsforskel. I metode 2.1 er de kontrafaktiske situationer markeret med en cirkel, dvs. hhv. resultatet af ikke at deltage for en deltager og resultatet af at deltage for

en ikke-deltager. De kontrafaktiske resultater er ikke til at observere direkte.

Først ser vi på den situation, hvor en person deltager i en socialpolitisk indsats. Hvis det er tilfældet, kan vi følge den fuldt optrukne linje i metode 2.1, hvor vi kan observere resultatet af deltagelse, der er Y_{1i} (deltager). Da vi ser det ud fra en deltager, ved vi imidlertid ikke, hvilket resultat den samme person ville have opnået, hvis personen ikke havde deltaget. Dette kontrafaktiske resultat, der benævnes Y_{0i} (deltager), er i metode 2.1 markeret med en cirkel. I praksis er det ofte aktuelt at se det ud fra en deltager, idet det ofte er en opgave at evaluere effekten af en socialpolitisk indsats, der er iværksat, hvor der foreligger faktisk observerede resultater for deltagerne.

Når vi følger den stiplede linje i metode 2.1 er udgangspunktet en person, der ikke deltager i en indsats. I det tilfælde kan Y_{0j} (ikke-deltager) observeres, og Y_{1j} (ikke-deltager) er det kontrafaktiske resultat, der ikke kan observeres. Vi ved nemlig ikke, hvad der ville være sket, hvis personen havde deltaget.

Indtil nu har vi understreget, at det er vigtigt at sammenligne resultaterne langs de vandrette linjer i metode 2.1 i stedet for at sammenligne den fuldt optrukne linje (deltagerne) med den stiplede linje (ikke-deltagerne). I praksis vil der som regel eksistere faktisk observerede resultater for deltagerne, der deltager, og for ikke-deltagerne, der ikke deltager, dvs. resultaterne uden cirkler i metode 2.1. Det gør, at der er mange studier, der sammenligner disse resultater. Det svarer til, at de, for at beregne effekten af at deltage, sammenligner Y_{1i} (deltager) med Y_{0j} (ikke-deltager). Det kan imidlertid forårsage problemer, at det er forskellige personer, der sammenlignes, hvilket er vist i form af indekseringen i og j i metode 2.1. De to typer af personer kan have væsentligt forskellige karakteristika, der gør, at de ikke har mulighed for at opnå samme resultat ved at deltage. I det tilfælde taler vi om, at der er et selektionsproblem. Det opstår, hvis det ikke er de samme typer af personer, der deltager og ikke deltager. Problemet er, at effekten af at deltage også vil være påvirket af, at personerne er forskellige, så vi ikke får den rene effekt af at deltage, når vi sammenligner deres resultater.

Det kan vises, at det er nødvendigt, at udgangspunktet for en person, der står for enten at deltage eller ikke at deltage, ikke må afhænge af, om han/hun rent faktisk vælger at deltage eller ej. Med notationen fra metode 2.1 svarer det til, at Y_{0i} (ikke-deltager) skal stemme overens med

Y_{0i} (deltager), hvor det bemærkes, at det ses ud fra den samme person, så begge størrelser her er indekseret med et i . På baggrund af tankegangen med den kontrafaktiske situation er det muligt at opstille et udtryk for effekten af indsatsen, som det ses i metode 2.2.

Metode 2.2

Et udtryk for effekten af indsatsen.

Antag, at vi ønsker at måle effekten af en indsats på en gruppe personers indkomst. En faktisk observeret indkomst kan enten være en indkomst for en person, der deltager, eller en person, der ikke deltager. Vi kan skrive den faktisk observerede indkomst Y_i , så den består af både den observerede indkomst og den kontrafaktiske indkomst. Alt efter om personen deltager eller ej, kan vi udelade den kontrafaktiske indkomst. Det gøres i det følgende udtryk ved at sætte $D_i = 1$, hvis personen deltager, og $D_i = 0$, hvis personen ikke deltager:

$$Y_i = D_i Y_{1i} + (1 - D_i) Y_{0i}$$

I modsætning til procesevaluering, der primært fokuserer på *monitoring* af resultatet fra Y_{0i} til Y_{1i} vil effektmåling gå ud på at vurdere virkningen af programmet, som vi her kan kalde effekten $Y_{1i} - Y_{0i}$. Ved at rykke lidt rundt på udtrykkene i ligningen ovenfor uden at ændre den, kan vi få effekten til at indgå mere direkte i ligningen:

$$Y_i = Y_{0i} + D_i (Y_{1i} - Y_{0i})$$

Det er imidlertid vigtigt at huske, at det aldrig er muligt at observere effekten direkte, da vi ikke kender den kontrafaktiske situation, dvs. man kan ikke både deltage og ikke deltage på samme tid. Der findes derfor en række måleredskaber for, hvorvidt en effekt er opnået eller ej.

I Danmark har vi et eksempel på et effektmålingsstudie, der blev iværksat som en opgave fra den nuværende regering, der gerne ville vide mere om, hvad folk får ud af at deltage i de tilbud om videregående efteruddannelseskurser, som iværksættes (Clausen et al., 2006). Finansministeriet og Amternes og Kommunernes Forskningsinstitut (AKF) fik til opgave af regeringen i dette effektmålingsstudie at vurdere, hvad videregående efteruddannelseskurser betyder for de enkelte personers indtægtsmuligheder på længere sigt. Indsatsen er tilbuddet om de videregående efteruddannelseskurser. Målgruppen er arbejdstagerne, og den forventede effekt af indsatsen er de øgede indtægtsmuligheder, personerne vil opnå, fordi de har deltaget i efteruddannelseskurser. Mere præcist er effekten af indsatsen forskellen mellem den indtægt, personen har, når denne har gennemført et kursus, hvilket er resultatet Y_{1i} (deltager), i forhold til den indtægt, den samme person ville have haft, hvis personen ikke havde gennemført kurset, Y_{0i} (deltager), dvs. den kontrafaktiske situation.

I forhold til de redskaber, som vi kan anvende, er det vigtigt at bemærke, at effekten af at deltage kan være individuel, hvilket vi ser i notationen ovenfor, da deltagelseffekten $Y_{1i} - Y_{0i}$, der er specificeret i metode 2.2, har en indikator for den enkelte person, nemlig i . Som vi tidligere har været inde på, er det nødvendigt at være opmærksom på, at personerne kan have forskellige karakteristika. De individuelle karakteristika kan både være observerbare, som fx alder, køn, uddannelse, eller uobserverbare, som fx evner. Denne heterogenitet blandt personer kræver, at vi anvender metoder, der kan tage højde herfor.

Det er også nødvendigt at præcisere, at individuelle forskelle kan stamme fra både heterogenitet i Y_{0i} og i Y_{1i} . Når vi ønsker at gøre en indsats for en gruppe, kan det meget vel tænkes, at de enkelte personers *baseline* er forskellige, så fx $Y_{0i} \neq Y_{0j}$. Det er dog ikke et problem i forhold til selektion, at personerne er forskellige, hvis der kan korrigeres for systematiske forskelle mellem personer, der deltager, og personer, der ikke deltager, jf. kapitlet om selektionsskævhed.

ER EFFEKTER INDIVIDUELLE?

Da effekten af deltagelse kan være forskellig for de enkelte personer, er det nødvendigt at være præcis med, hvordan man ønsker at måle effek-

ten for en gruppe personer. Når vi bruger notationen fra ovenfor, kan vi udtrykke heterogenitet ved, at effekten af at deltage for person i kan være forskellig fra effekten af at deltage for person j , dvs.,

$$Y_{1i} - Y_{0i} \neq Y_{1j} - Y_{0j}$$

Hvis det at deltage virker ens på alle personer, er der homogene effekter af deltagelse, og der vil være lighedstegn i ligningen ovenfor. Hvis det er tilfældet, vil de tre målemetoder, som gennemgås nedenfor bevirke, at estimaterne af effekterne er de samme.

Da antagelsen om homogene deltagelseseffekter ikke er realistisk, er det meget vigtigt at være præcis med, hvad det er, man ønsker at undersøge og for hvilken population. I det følgende vil vi antage, at vi har fundet en god måde at måle effekten af en indsats på, hvilket indebærer, at vi har et godt mål for den kontrafaktiske situation. Herudover antager vi, at vi både kan måle effekten for den gruppe i populationen, der deltager, og den gruppe, der ikke deltager. De følgende tre afsnit gennemgår hver deres måde at beregne effekten for en samlet gruppe på:

Den gennemsnitlige effekt af deltagelse for den gruppe, der deltager

Vi ønsker at beregne den gennemsnitlige effekt af at deltage på baggrund af den gruppe af personer, der faktisk deltager. Det kaldes *Average Treatment Effect on the Treated* i den engelsksprogede litteratur, som derfor forkortes til ATET. Det svarer til, at vi beregner effekten, idet vi følger den fuldt optrukne linje i metode 2.1, hvor det ses på baggrund af en deltager. Her kan vi måle resultatet for en deltager og sammenholde det med det kontrafaktiske resultat for, hvad personen ville have opnået, hvis personen ikke havde deltaget. Det kontrafaktiske resultat kan vi selvfølgelig ikke observere, men vi har metoder, der forsøger at opstille mål for det, og som beskrives senere i denne manual. Vi kan forestille os, at vi gør dette for en gruppe af personer, der deltager. Den estimerede effekt af at deltage, givet at man er en af deltagerne, fremkommer herefter ved deltagernes gennemsnitlige resultat fratrasket deltagernes gennemsnitlige kontrafaktiske resultat.

Et vejet gennemsnit af effekten af deltagelse for deltagere og ikke-deltagere

Vi ønsker at beregne den gennemsnitlige effekt af deltagelse på baggrund af både deltagelse og ikke-deltagelse, svarende til at der tages højde for begge situationer i metode 2.1. Den gennemsnitlige effekt for deltagerne blev beregnet som ATET ovenfor. Dernæst skal vi bruge den beregnede effekt for deltagelse, givet at en person ikke deltager, hvilket svarer til at følge den nederste stiplede linje i metode 2.1. Sammenvejer vi alle individuelt beregnede effekter for hver af de to grupper vægtet med sandsynligheden for at deltage eller ikke-deltage, får vi et vejet gennemsnit af effekten af deltagelse for både dem, der deltager, og dem, der ikke deltager. Denne samlede sammenvejede gennemsnitlige effekt for både deltagere og ikke-deltagere kaldes ATE for *Average Treatment Effect*.³

Den gennemsnitlige effekt for marginale deltagere

Det kan også tænkes, at man ønsker at udvide en indsats, så der er flere, der bliver tilbudt at deltage. I det tilfælde kan det være relevant at se på den gennemsnitlige effekt af deltagelse for de nye deltagere, dvs. effekten for de personer, der tages marginalt ind fra gruppen af ikke-deltagere.

HVEM SKAL TILBYDES DELTAGELSE?

Når man på en fornuftig måde skal vurdere, hvem der bør tilbydes at deltage i en indsats, er det nødvendigt at tage hensyn til, hvad det politiske ambitionsniveau er. Overordnet set er der to hovedprincipper til bedømmelse af, hvem der bør tilbydes en indsats. For det første kan det være et spørgsmål om politisk prioritering, og for det andet kan det være en forudsætning, at det er økonomisk efficient. De to principper kan både tænkes at være i overensstemmelse med hinanden og i modstrid med hinanden.

På den ene side kan det være en politisk prioritering, at der i samfundet skal være en sikring af, at alle individer kan modtage hjælp,

³ Matematisk skrives ATET og ATE:

$$ATET = E[Y_{1i} - Y_{0i} \mid D_i = 1]$$

$$ATE = E[Y_{1i} - Y_{0i}]$$

hvis de er uheldige at opnå en uacceptabelt lav social position. Det svarer til, at det primært er personer, der som baseline har et dårligt resultat, dvs. en relativt lav værdi af Y_{0i} , der bør tilbydes en indsats. Det kan fx være en politisk beslutning, at en udsat gruppe bør tilbydes et værested, fordi det skønnes, at det har en positiv effekt for den udsatte gruppe, selvom det kan være svært at måle en effekt for samfundet i øvrigt. Den overordnede tankegang er i så tilfælde, at det er politisk uacceptabelt, at der er personer i samfundet, der har det dårligere end et vist defineret niveau.

På den anden side kan det være et udgangspunkt, at der skal komme så meget ud af en given sum penge som overhovedet muligt. Her vil man i højere grad anlægge det synspunkt, at det skal være økonomisk efficient. Hvor gør pengene med andre ord mest nytte? Det samfundsmæssigt set mest økonomisk efficiente resultat får vi ved at maksimere summen af den forventede effekt for alle individer, der deltager, og fratrække de omkostninger, der har været forbundet med deltagelse. Det giver det nettoresultat, som deltagelse giver. Hvis vi forestiller os, at vi kan sammenligne mange forskellige indsatstyper, der hver kræver de samme omkostninger, vil det være økonomisk efficient at vælge den løsning, der maksimerer dette nettoresultat. I hvert fald vil vi betænke os ved at tilbyde en indsats efter dette princip, hvis den individuelle effekt er mindre end omkostningerne forbundet hermed, hvorfor vi ser bort fra denne mulighed.

De to principper om politisk prioritering og økonomisk efficiens er i overensstemmelse med hinanden, hvis effekten af deltagelse er størst for de individer, der har den mindste værdi af Y_{0i} . Omvendt må det prioriteres, hvilket princip der bør vælges efter, hvis det er dem, der har højere værdier af Y_{0i} , der får mest ud af at deltage. Det bliver således et politisk spørgsmål, hvad der skal prioriteres. I denne situation kan det fx være betænkeligt alene at bedømme projektmagere på, om de individer, der er involverede i deres projekt, opnår et godt resultat eller ej. Hvis de har mulighed for det, vil det så nemlig være en fordel for projektmagerne at udvælge relativt velstillede individer til deres projekt for at opnå størst mulig succes. Dette kaldes i litteraturen ofte *cream skimming*.

SELEKTIONSSKÆVHED

Formålet med mange af de metoder, der anvendes inden for effektmåling, er at finde gode måder at konstruere den kontrafaktiske situation på, da den ikke kan observeres i virkeligheden. I litteraturen anvendes der både eksperimentelle og ikke-eksperimentelle metoder til dette. Der er stor interesse for at udvikle disse metoder, da det er den eneste måde, hvorpå man kan isolere effekten af et givet politisk tiltag for en gruppe personer, når der iværksættes et projekt til forbedring af deres forhold. Problemet er, at de personer, der udtages til deltagelse, sjældent kan sammenlignes direkte med de personer, der ikke deltager. Fagfolk taler i dette tilfælde om, at der er en selektionsskævhed, hvilket der kan gives et eksempel på. Hvis et socialt initiativ går ud på at stille et frivilligt uddannelsesstilbud til rådighed for en udsat gruppe af mennesker, kan det nemt tænkes, at det vil være de relativt bedst stillede inden for gruppen af de udsatte, der har overskud til at deltage. Herved opstår der en selektionsskævhed, idet gruppen af deltagere er forskellig fra gruppen af ikke-deltagere, så det ikke er muligt at sammenligne de to grupper for at finde effekten af at deltage, idet den også vil være påvirket af, at personerne er forskellige.

På det sociale område kan der tænkes at være visitering til deltagelse, idet en indsats ofte vil have til formål at komme en særligt udsat gruppe til gode. Hvis der er knappe ressourcer, så det ikke er muligt at tilbyde alle visiterede et tilbud, kan der trækkes lod om deltagelse, hvis

det vurderes, at alle i princippet burde have lige ret til tilbuddet. De, der ikke trækkes ud, vil i så fald udgøre en oplagt ikke-deltagergruppe. Hvis alle visiterede er akkurat lige berettigede, vil man således undgå et selektionsproblem ved at trække lod blandt de visiterede.

Vi vil gerne måle effekten som det resultat, personen opnår ved at deltage, fratrullet det kontrafaktiske resultat, som personen ville have opnået ved ikke at deltage. Det svarer til at følge den fuldt optrukne streg i figuren i metode 2.1. Vi kan skrive det således:

$$\text{Effekt}_i = Y_{1i}(\text{deltager}) - Y_{0i}(\text{deltager})$$

Y_{1i} (deltager) er det resultat, person i opnår ved at deltage, og Y_{0i} (deltager) er det kontrafaktiske resultat, samme person kunne have opnået, hvis han/hun ikke havde deltaget.

Gennemsnittet af alle de individuelle effekter, ATET, fremkommer herved som:

$$\text{ATET} = \text{gns}(Y_{1i}(\text{deltager})) - \text{gns}(Y_{0i}(\text{deltager}))$$

Herved får vi således et udtryk for den gennemsnitlige effekt af deltagelse for den gruppe, der deltager (ATET), ved at tage gennemsnittet af resultaterne for deltagerne og fratrække gennemsnittet af deltagernes kontrafaktiske resultater af ikke at deltage.⁴ Da det ikke er muligt at observere det kontrafaktiske resultat Y_{0i} (deltager), er vi nødt til at overveje, hvad der så er mest hensigtsmæssigt at gøre.

For at gøre det enkelt kan vi indledningsvist antage, at vi har to ens individer i og j , der som baseline har samme karakteristika ($Y_{0i} = Y_{0j}$). Hvis de udsættes for en indsats, vil de også reagere ens, så $Y_{1i} = Y_{1j}$. Hvis det ene individ deltager, og det andet ikke gør, vil vi i dette specielle tilfælde få en præcis vurdering af, hvad effekten af deltagelse er for person i ved at sammenligne resultaterne for de to personer:

⁴ Den tilsvarende matematiske notation er:

$$\text{ATET} = E[Y_{1i} | D_i = 1] - E[Y_{0i} | D_i = 1]$$

Vi har valgt at bruge gennemsnittet i stedet for forventningsoperatoren i teksten for at gøre det mere forståeligt for den ikke-matematiske læser. Til den matematiske læser skal det imidlertid medgives, at gennemsnittet blot er et skøn på den forventede værdi.

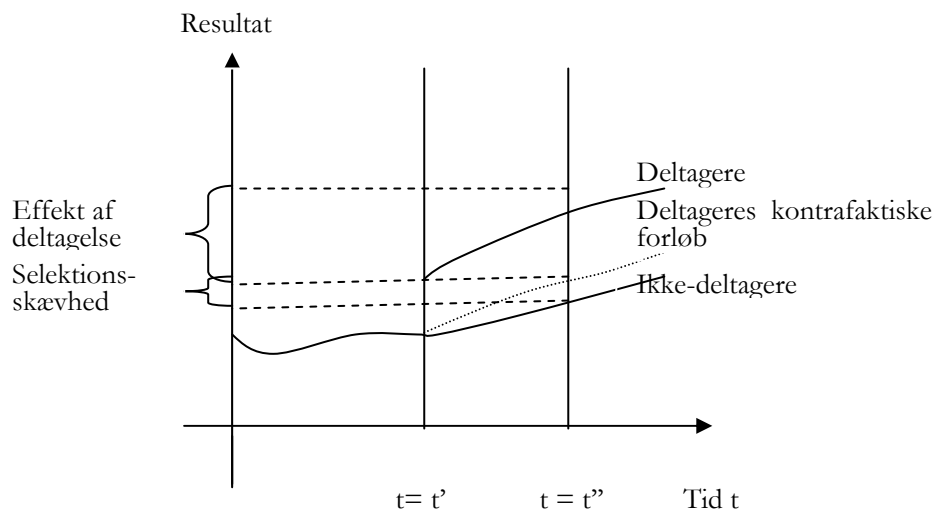
$$\text{Effekt}_i = Y_{1i}(\text{deltager}) - Y_{0j}(\text{ikke - deltager})$$

Da de to personer reagerer ens, kan vi beregne effekten for person i ved at bruge ikke-deltagerresultatet for person j som mål for, hvad ikke-deltagerresultatet ville have været for person i .

Da det sjældent er muligt i praksis at finde identiske personer, så der kan sammenlignes, er det nærmeste, vi kan komme derpå, at opstille en gruppe, der tilbydes en indsats og en gruppe, der så vidt muligt svarer til deltagergruppen, der ikke tilbydes en indsats. Herved er vi i den situation, hvor vi sammenligner de faktisk observerede resultater for en deltager- og en ikke-deltagergruppe, dvs. de resultater uden cirkler omkring i metode 2.1, hvilket blev omtalt i afsnittet om at måle effekter ovenfor. I det omfang de to sammenligningsgrupper ikke svarer til hinanden, bortset fra om de deltager eller ej, vil der opstå en selektionsskævhed, hvilket er illustreret i figur 3.1.

Figur 3.1

Selektionsskævhed mellem deltagere og ikke-deltagere.



I figur 3.1 følger vi en gruppe af deltagere og en gruppe af ikke-deltagere over tid. Den øverste fuldt optrukne linje viser et gennemsnit af de fakti-

ske resultater over tid for de enkelte personer i deltagergruppen (jf. Y_{1i} (deltager) i metode 2.1). Den nederste fuldt optrukne linje viser det gennemsnit af de faktiske resultater for de enkelte personer i ikke-deltagergruppen (jf. Y_{0j} (ikke-deltager) i metode 2.1). Vi forestiller os, at der iværksættes en indsats til tid $t = t'$, hvilket den første vertikale linje markerer. Når der er gået et stykke tid, dvs. fx til tid $t = t''$, kan vi måle, hvad resultaterne er for hver af grupperne.

Hvis vi herefter forestiller os, at deltagerne ikke havde deltaget, er det muligt, at de kunne have opnået et bedre resultat end ikke-deltagerne. Der kan fx være en selektion i, at det er de bedst kvalificerede, der vælger at deltage i et uddannelsesstilbud. Hvis resultatet måles på den gennemsnitlige indkomst, ville disse deltagere ifølge figuren således også have opnået højere indkomster end ikke-deltagerne, hvis de ikke havde deltaget i kurset. Det er det gennemsnitlige kontrafaktiske resultat for deltagerne, der er markeret ved den mellemste tæt stiplede linje i figur 3.1. Effekten af at deltage i en indsats for deltagergruppen (ATET) bør derfor måles som forskellen mellem den øverste kurve for deltagerne og deltagerne kontrafaktiske forløb.

Figur 3.1 illustrerer, at den deltagereffekt, vi ønsker at beregne, også kan være påvirket af en selektionsskævhed, hvis vi bruger ikke-deltagerne resultat som skøn for deltagerne kontrafaktiske resultat Y_{0i} (deltager), hvilket svarer til at sammenligne de fuldt optrukne linjer for hhv. deltagere og ikke-deltagere i figur 3.1. Vi kan præcisere, at det er et skøn for den beregnede effekt af at deltage ved at tilføje en "hat" over ATET:⁵

$$\widehat{ATET} = gns(Y_{1i}(\text{deltager})) - gns(Y_{0j}(\text{ikke-deltager}))$$

Skønnet på ATET indebærer, at det er et forsøg på at give et bud på den sande værdi af ATET. Fejlen, der kan opstå ved dette skøn, er selektionsskævheden, der er markeret i figur 3.1. Den opstår således, fordi vi

⁵ Matematisk kan det skrives:

$$\begin{aligned}\widehat{ATET} &= E[Y_{1i} | D_i = 1] - E[Y_{0i} | D_i = 0] \\ &= E[Y_{1i} | D_i = 1] - E[Y_{0i} | D_i = 1] + \{E[Y_{0i} | D_i = 1] - E[Y_{0i} | D_i = 0]\}\end{aligned}$$

hvilket svarer til udtrykket, der fremkommer i metode 3.1.

bruger $\text{gns}(Y_{0j} \text{ (ikke-deltager)})$ som erstatning for det ikke-observerbare kontrafaktiske gennemsnitlige resultat for deltagerne $\text{gns}(Y_{0i} \text{ (deltager)})$.

I metode 3.1 ses, hvordan der kan fås et mere præcist udtryk for selektionsskævheden:

Metode 3.1

Et udtryk for selektionsskævheden.

Vi kan præcisere den selektionsskævhed, der opstår ved at bruge det følgende skøn på ATET:

$$\hat{ATET} = \text{gns}(Y_{1i} \text{ (deltager)}) - \text{gns}(Y_{0j} \text{ (ikke - deltager)})$$

Fra dette udtryk kan vi omskrive uden at ændre indholdet ved at lægge det gennemsnitlige kontrafaktiske resultat for deltagerne fra og til:

$$\begin{aligned} \hat{ATET} = & \left(\text{gns}(Y_{1i} \text{ (deltager)}) - \text{gns}(Y_{0i} \text{ (deltager)}) \right) \\ & + \left(\text{gns}(Y_{0i} \text{ (deltager)}) - \text{gns}(Y_{0j} \text{ (ikke - deltager)}) \right) \end{aligned}$$

Det første led er altså ATET, og det andet led er netop et udtryk for den selektionsskævhed, der kan være imellem de personer, der deltager, og de personer, der ikke deltager, hvis de to grupper ikke har samme resultat uden deltagelse, jf. også figur 3.1. Herved ses det også, at en randomiseret udvælgelse ville føre til, at selektionsskævheden bliver nul, så:

$$\text{gns}(Y_{0i} \text{ (deltager)}) = \text{gns}(Y_{0j} \text{ (ikke - deltager)})$$

Fra metode 3.1 vil vi referere til kravet om, at der ikke er selektionsskævhed, hvilket kræver, at *baseline* for både deltager- og ikke-deltagergruppe svarer til hinanden, hvis de ikke deltog:

$$\text{gns}\left(Y_{0i} \text{ (deltager)}\right) = \text{gns}\left(Y_{0j} \text{ (ikke - deltager)}\right) \quad (1)$$

Problemet er bl.a., at selektionsskævheden i sig selv kan bevirke, at estimatet for effekten er positivt, selvom der reelt ikke er en effekt af at deltage (dvs. ATET er nul, så $Y_{1i} = Y_{0i}$).

Det, at der kan være problemer med at skønne, hvad selektionsskævheden er, kan også udtrykkes ved, at der er en trussel mod det, der kaldes intern validitet, dvs. evnen til at vurdere den reelle deltagereffekt for dem, der deltager. Hvis det er muligt at lave et perfekt eksperimentelt design for det sociale eksperiment, vil der imidlertid ikke være problemer, da ligning (1) ovenfor så vil være opfyldt. I det næste kapitel gennemgås imidlertid nogle af de problemer, der kan opstå, hvis man ønsker at tilrettelægge et sådant randomiseret forsøg.

EKSPERIMENTELLE METODER

Der findes både eksperimentelle og ikke-eksperimentelle metoder til at tage højde for den kontrafaktiske situation. En eksperimentel metode kan fx være et randomiseret forsøg, hvor der trækkes lod mellem, hvem der deltager, og hvem der ikke deltager. Hvis der er trukket tilfældigt lod, kan vi antage, at sammensætningen af personer i den ene gruppe svarer til sammensætningen i den anden gruppe, hvorved effekten kan beregnes som forskellen mellem resultaterne for hver af de to grupper.

INTERN OG EKSTERN VALIDITET AF RANDOMISEREDE FORSØG

Reaktionen på problemer med selektionsskævhed kan nemt blive, at man laver randomiserede forsøg, hvor den tilfældige udvælgelse burde føre til, at man undgår problemet. Det er baggrunden for, at der ofte refereres til de randomiserede forsøg som "guldstandard". Det har imidlertid vist sig, at der også kan være problemer forbundet med at gennemføre randomiserede forsøg. I litteraturen nævnes i denne forbindelse, at der kan være trusler mod både den interne og den eksterne validitet af et randomiseret forsøg.

Der opstår problemer med den interne validitet, hvis vi ikke kan isolere effekten af indsatsen, typisk fordi det alligevel ikke er tilfældigt,

hvilke personer blandt de tilfældigt udtrukne der ender med at deltage. Herved opstår der en selektionsskævhed, der ligger internt i det randomiserede forsøg. Den eksterne validitet vedrører, hvorvidt det er muligt at generalisere resultaterne. Her vil en eventuel selektionsskævhed således ligge eksternt i det randomiserede forsøg i den forstand, at de effekter, der måles for de personer, der deltager i forsøget, ikke nødvendigvis kan overføres til den situation, hvor man gerne vil bruge de målte effekter. Det er således spørgsmålet om, hvad der sker, hvis forsøget gentages, udvides eller udsættes for en anden gruppe.

Trusler mod intern validitet

- Det er vigtigt, at alle deltagere i et eksperiment har forstået og er indforståede med betingelserne. Ellers vil man ikke nødvendigvis opnå det tilsigtede resultat. Den oprindelige allokering til gruppen af deltagere og gruppen af ikke-deltagere kan tænkes at blive ændret, hvis det fx lykkes personer fra ikke-deltagergruppen at komme til at deltage alligevel eller evt. i noget, der ligner uden for forsøget. Det kan også tænkes, at der er personer i deltagergruppen, der ikke gør det, de bliver bedt om, eller ikke gennemfører forsøget.
- Ud over problemer med deltagerne i forsøget kan der være problemer, hvis de, der styrer forsøget, påvirker resultatet tilsigtet eller utilsigtet. Det kan fx være, at de ikke er godt nok instrueret og uddannet, eller de har etiske problemer med ikke at tilbyde personer i ikke-deltagergruppen deltagelse.

Trusler mod ekstern validitet

- For at teste indsatsens resultater kan det være hensigtsmæssigt at udføre flere forsøg. Der kan imidlertid i praksis være grænser for, hvor mange det er muligt at gennemføre. Det kan også tænkes, at det er nødvendigt at begrænse gruppen af forsøgspersoner i forhold til den gruppe af personer, man senere ønsker at tilbyde deltagelse. I det tilfælde er det ikke nødvendigvis muligt at generalisere resultatet, idet den udvidede gruppe med personer kan tænkes at adskille sig fra de personer, der er i den oprindelige gruppe.
- De personer, der er villige til at deltage i et forsøg, kan adskille sig fra de personer, der senere gerne vil deltage.

Det skal således konkluderes, at der kan være flere grunde til, at det er svært at gennemføre randomiserede forsøg. Medmindre et forsøg er designet helt ideelt, kan det tænkes, at der tilvejebringes et bedre informationsgrundlag ved brug af de ikke-eksperimentelle metoder med hensyn til information om den kontrafaktiske situation. Ud over de allerede nævnte forhold i dette kapitel er der også ved den eksperimentelle metode problemer med muligheden for at bedømme tilbagespilseffekter. Det er et problem, både når det gælder de eksperimentelle og de ikke-eksperimentelle metoder, jf. afsnittet om begrænsningerne ved metoderne nedenfor.

Brugen af eksperimentelle metoder er imidlertid øget inden for de sidste fem år i fx USA og England. I Angrist (2004) angives to hovedargumenter for at udføre randomiserede forsøg. For det første nævner han, at det kan diskuteres, om det er uetisk ikke at tilbyde alle en indsats, da udgangspunktet som regel er, at det ikke er muligt at stille det konkrete tilbud til rådighed. For det andet mener han ikke, at det er problematisk, hvis der alligevel er begrænsninger i, hvor mange der kan tilbydes deltagelse. Som det blev omtalt i kapitlet om selektionsskævhed, er det ikke ualmindeligt, at der er flere visiterede til et tilbud end dem, der med de givne ressourcer kan få et tilbud om deltagelse.

Et væsentligt argument til fordel for eksperimentelle metoder er vel også, at det kan være uetisk at blive ved med at afholde udgifter til indsatser, hvis der ikke er tilstrækkelig god dokumentation for, at de virker efter hensigten.

IKKE-EKSPERIMENTELLE METODER

I dette kapitel beskrives de ikke-eksperimentelle metoder, der anvendes i forbindelse med effektmåling i dag. Det kan være ønskværdigt at supplere med en eksperimentel metode, men det kommer vi ikke ind på her. Derimod forsøger vi at begrænse os til at forklare metoderne hver for sig, så terminologien bliver klar. De ikke-eksperimentelle metoder er:

- regression
- matching
- Instrument Variabel (herefter benævnt IV, da det er en almindelig forkortelse)
- quasi-eksperimentelle metoder.

Der er et par forhold, der er vigtige, når fordele og ulemper for de enkelte metoder gøres op. For det første er det for hver af de enkelte metoder vigtigt at se på, hvordan personerne er forskellige uden iværksættelsen af en indsats, idet det har betydning for, hvordan de kontrafaktiske resultater for deltagerne bør konstrueres. For det andet er det væsentligt at se på, hvordan man bør aggregere og sammenligne de estimerede effekter af deltagelse.

STATISTISK KONTROL – REGRESSION

Ved en regression prøver vi at få et samlet overblik over, hvad effekten af deltagelse er, når der er mange individer. I regressionen finder vi således en sammenhæng mellem det at deltage og betydningen for en resultatvariabel som fx den faktisk observerede indkomst. Regressionen svarer til, at vi finder de gennemsnitlige niveauer af resultaterne for alle personer uden deltagelse, og at vi lægger den gennemsnitlige effekt af deltagelse for dem, der deltager, dertil.

Vi opstiller regressionen ved at regressere personernes faktiske observation Y_i (fx indkomst) på, om de deltager eller ej, hvilket vi ovenfor angav ved variablen D_i , så:

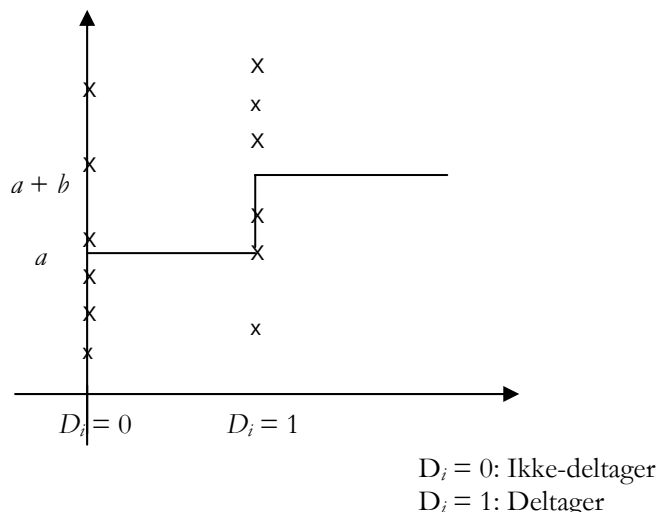
$$Y_i = a + b D_i + e_i \quad (2)$$

Den fælles parameter a for alle individer kan tolkes som det gennemsnitlige resultat for alle individer før deltagelse, og parameteren b kan tolkes som det resultat, personerne i gennemsnit får ud af at deltage. Det svarer til, at der pålægges homogenitet, så resultatet for alle personer vurderes i forhold til gennemsnittet for hele gruppen. Den afvigelse, der er for de enkelte personer i forhold til gennemsnittet, må opfanges i restleddet e_i , hvor heterogeniteten med andre ord bliver repræsenteret. Regressionen er illustreret i figur 5.1.

Figur 5.1

Estimeret sammenhæng mellem indkomst og deltagelse eller ikke-deltagelse.

Faktisk indkomst Y_i



Der er krydser for både personer, der deltager, hvor $D_i = 1$, og for personer, der ikke deltager, hvor $D_i = 0$. Den første nederste del af linjen i figur 5.1 markerer den gennemsnitlige indkomst for de personer, der ikke deltager, som er lig konstantleddet a i vores regression. I forhold til vores tidligere notation kunne vi skrive det som $\text{gns}(Y_{0j} \text{ (ikke-deltager)}) = a$. De personer, der deltager, har herudover også et bidrag fra effekten af deltagelse, som er b . Vurderingen af den faktiske indkomst for deltagerne kan med vores tidligere notation skrives som $\text{gns}(Y_{1i} \text{ (deltager)}) = a + b$. Forskellen fra de enkelte krydser til de gennemsnitlige niveauer svarer til heterogeniteten, der opfanges i restleddet e_i .

Vi er selvfølgelig mest interesserede i at finde et estimat for den homogene deltagereffekt i form af parameteren b , der svarer til springet i den gennemsnitlige faktiske indkomst fra ikke at deltage til at deltage, hvor $D_i = 1$ i figur 5.1. Hvis vi ikke er meget påpasselige med vores regression, kan vi imidlertid godt risikere, at vores estimat for parameteren b bliver skævt. Det kan ske, hvis der er en selektionsskævhed, hvilket vil sige, at ligning (1) ovenfor ikke er opfyldt. Det er imidlertid muligt, at det kan lade sig gøre at finde observerbar information, så vi kan korrigere for

selektionsskævheden. I det følgende beskriver vi det så vidt muligt uden at vise det matematisk.

For at undgå selektionsproblemet er det tilstrækkeligt, at den faktiske indkomst uden deltagelse ikke afhænger af, om personen vælger at deltage eller ej, hvilket svarer til, at ligning (1) gælder (se evt. bevis i Wooldridge (2002) eller Lee (2005)). Vi kan således løse problemet ved at forsøge at finde information, der kan korrigere for systematiske forskelle i indkomsterne uden deltagelse for hhv. deltager- og ikke-deltagergruppe. Det kan således lade sig gøre at estimere ATE og ATET ved en ordinær regression, når der kan skelnes mellem deltagere og ikke-deltagere ud fra observerbare karakteristika. Ud over at estimere et gennemsnit for den faktiske indkomst uden deltagelse (jf. parameteren a ovenfor) tilføjer vi derfor ekstra forklarende variable som fx køn, uddannelsesniveau, civilstand mv., som vi kan benævne x_i . Herved fås den følgende regression for indkomsten uden deltagelse:

$$Y_{0i} = a + c(x_i - \text{gns}(x_i)) + u_i$$

Det at kontrollere for observerbare karakteristika kan således være en måde at undgå, at ikke-deltageres indkomster varierer systematisk mellem deres deltager- og ikke-deltagerresultater, hvor deltagerresultatet selvfølgelig er kontrafaktisk. Vi tror med andre ord på, at kontrollen for de observerbare karakteristika gør, at Y_{0i} ikke længere afhænger af D_i som en slags indikator for den baggrundsinformation, der er med til at forklare selektionsskævheden mellem deltagere og ikke-deltagere. I selve regressionen bevirker det, at der lægges et individuelt bidrag til konstantledet, hvilket i figur 5.1 ses ved, at indkomsterne uden deltagelse kan variere mellem individer, så der er forskellige niveauer for linjen, hvor $D_i = 0$. Ved yderligere at introducere variabelen for deltagelse eller ej i regressionen får vi et udtryk for den faktisk observerede indkomst Y_i , der er lig med Y_{0i} , hvis $D_i = 0$ og Y_{1i} , hvis $D_i = 1$:

$$Y_i = a + b D_i + c(x_i - \text{gns}(x_i)) + u_i$$

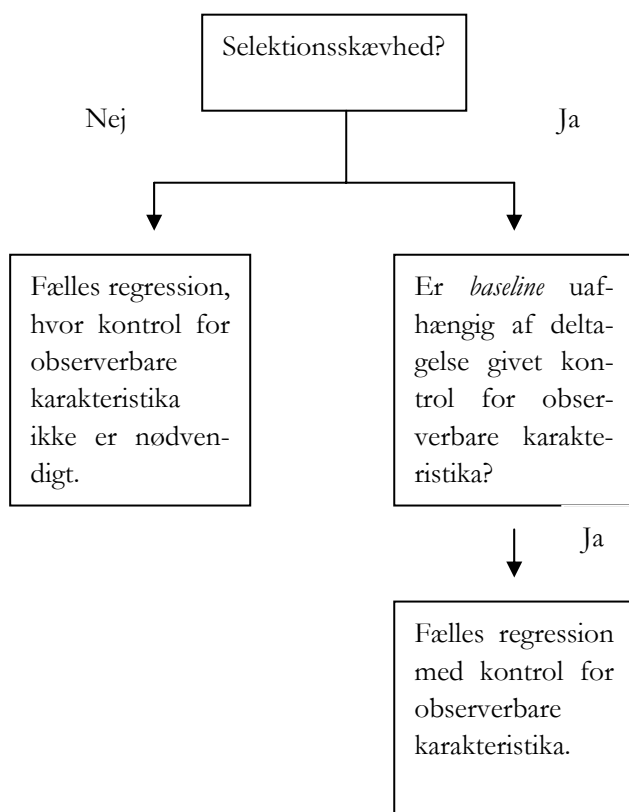
Effekten af at deltage bliver således estimeret ved parameteren b , hvis estimat er en homogen effekt for alle individer. I figur 5.1 svarer det til, at størrelsen af springet fra den nederste linje til den øverste linje er ens for alle individer, selv om de individuelle linjer kan ligge forskelligt.

Regressionen forudsætter, at det er muligt at fange systematisk information tilfredsstillende, så den information, der bliver tilbage i restleddet, er tilfældig. Hvis det er opfyldt, er det derfor lykkedes ved hjælp af udnyttelse af baggrundsinformation at få indkomsterne med og uden deltagelse til at være uafhængige af, om en person deltager eller ej. Hermed vil vi undgå, at de estimerede parametre bliver skæve som følge af selektionsproblemet.

I figur 5.2 illustreres tankegangen med at bruge regressionsmetoden ved hjælp af en oversigt:

Figur 5.2

Oversigt over problemet ved selektionsskævhed for regressioner.



Når vi estimerer en regression, vil vi gerne kunne konkludere, om en indsats virker eller ej. Som vi har set, kræver det, at vi tjekker estimationsarbejdet grundigt, da vi ellers kan risikere at få skæve parameterestimater, jf. figur 5.2. Herudover er det også vigtigt, at vi kan fortolke parameterestimaterne. Estimatet af effekten af deltagelse b i regressionen implicerer her en eksakt sammenvejning af de individuelle effekter, der er forholdsvis enkle at analysere, når der er homogene effekter af deltagelse. Hvis vi ønsker, at metoden skal tage højde for heterogene effekter af deltagelse, vil matching-metoden være at foretrække frem for regressionsmetoden.

MATCHING

Ved matching er der også mulighed for at korrigere for selektionsskævheden ved at betinge på observerbare karakteristika, men det foregår lidt anderledes end ved regressionsmetoden. Når der matches, søger vi for hver deltager efter en tilsvarende person ud fra de udvalgte observerbare karakteristika til dannelsen af en ikke-deltagergruppe. Det kontrafaktiske resultat fremkommer herefter på baggrund af resultaterne for personerne i ikke-deltagergruppen.

Lad os antage, at person i fra deltagergruppen matcher en person j fra ikke-deltagergruppen ud fra veldefinerede observerbare karakteristika. I forhold til vores notation i metode 2.1 kan vi udtrykke det ved:

$$\text{gns}(Y_{0i}(\text{deltager})) = \text{gns}(Y_{0j}(\text{ikke - deltager}))$$

Denne ligning udtrykker, at det forventede kontrafaktiske resultat af ikke at deltage for en deltager, dvs. $\text{gns}(Y_{0i}(\text{deltager}))$, er lig med det gennemsnitlige resultat uden deltagelse for lignende ikke-deltagere. For at se på lignende ikke-deltagere burde der i formlen være betinget på den observerbare information, der gør, at de to grupper, der sammenholdes, bliver sammenlignelige. For at holde det enkelt har vi valgt ikke at vise det i formlen. Det svarer imidlertid blot til, at formlen først bør anvendes, når vi har sorteret alle personer ud fra observerbare karakteristika. Det kan fx være kriterier som køn, alder, indkomstniveau mv. For hver af de

sorterede grupper bruger vi formelen, hvilket svarer til, at vi har betinget på observerbare karakteristika.

Ifølge formelen svarer en deltagers kontrafaktiske resultat i gennemsnit fuldstændigt til det faktisk observerede resultat for en, der ikke deltager. Det er muligt, fordi vi antager, at det er lykkedes at matche så godt, at det faktisk observerede resultat for den tilsvarende person i ikke-deltagergruppen kan benyttes, som om det er deltagerens resultat, hvis deltageren ikke havde deltaget, dvs. som det kontrafaktiske resultat. I praksis kan der selvfølgelig være flere måder at danne sig en forventning om, hvad det kontrafaktiske resultat for deltagergruppen vil være på baggrund af de informationer, der foreligger i ikke-deltagergruppen.

Fra denne matching antages det nu, at der er estimeret kontrafaktiske udfald for alle deltageres resultater, så vi kan beregne et estimat for alle deltager effekter ved at trække det kontrafaktiske resultat fra det faktisk observerede resultat:

$$\hat{\text{Effekt}}_i = Y_{1i}(\text{deltager}) - \text{gns}(Y_{0i}(\text{deltager}))$$

Ved at tage et simpelt gennemsnit af alle deltageres estimerede effekter kan vi igen beregne et estimat for ATET. Fordelen ved matching-metoden er, at der beregnes estimerede effekter for de enkelte individer i stedet for, at de individuelle effekter bliver sammenvejet i en estimeret parameter, hvilket var tilfældet for parameteren b i regressionsmetoden ovenfor. En fordel kan fx være, at det er muligt at tegne fordelingen af de estimerede effekter. Hvis vi som her blot er interesserede i middelværdien, er det som ovenfor tilstrækkeligt, at vi har løst selektionsproblemet ved, at ligning (1) er opfyldt givet de observerbare karakteristika, vi betinger på.

Matching har imidlertid den yderligere fordel, at vi kan skærpe forudsætningen fra ligning (1). Vi kan fx kræve, at hele fordelingen af indkomster uden deltagelse ikke afhænger af, om personen deltager eller ej, når der er kontrolleret for observerbare karakteristika. Hvis det er tilfældet, kan vi nemlig også interessere os for andre mål, som fx medianen. Problemet er, at indkomsterne uden deltagelse er kontrafaktiske for deltagergruppen, så det kan indvendes, at det kan være svært at tjekke, om fordelingerne er ens. Der er dog situationer, hvor det kan sandsynliggøres, at de to fordelinger er ens. Det kan være tilfældet, hvis der er pa-

neldata, hvor personerne følges før og efter iværksættelsen af indsatsen, så det er muligt at undersøge, om indkomsterne svarede til hinanden i deltager- og ikke-deltagergruppe, før indsatsen blev iværksat.

I nogle situationer er der mange observerbare karakteristika, der ønskes at matche på, hvilket kan give et dimensionsproblem, når metoden anvendes. Til afhjælpning af dette er der udviklet en metode inden for matching, der hedder *propensity score*. Det vil imidlertid blive for omfattende at komme ind på den her.

Matching-metoden kan sammenfattende konkluderes at være mere fleksibel end regressionsmetoden. Fordelen ved matching er, at man ikke tvinger en funktionel form ned over resultatligningen. Det er således muligt at håndtere og fortolke heterogene effekter af deltagelse, som kan sammenvejes på forskellige måder, så længe det giver et konsistent estimat. Omvendt vil regressionsmetoden være mest fordelagtig, hvis effekterne af deltagelse er homogene, idet det er en enklere metode.

INSTRUMENT VARIABLE (IV) ESTIMATION

Instrument Variabel estimation, der efterfølgende kaldes IV-estimation, er et alternativ til regressionsmetoden, hvis der er problemer med, at noget information er udeladt, jf. evt. Wooldridge (2003) for en grundlæggende indføring i tankegangen. Til forskel fra regressionsmetoden, hvor det blev antaget, at det var muligt at finde observerbare karakteristika til at korrigere for den udeladte information, kan IV-metoden benyttes, hvis den udeladte information ikke er observerbar. Idéen i IV-metoden er at finde et instrument, der teknisk set er en variabel, der så vidt muligt er korreleret med deltagelsesvariablen D_i uden at være korreleret med restleddet, hvilket uddybes nedenfor. Hvis der *ikke* er observerbare karakteristika, der forklarer, at fx flere personer deltager i et regionalt område frem for i et andet regionalt område, vil IV-metoden således være anvendelig.

Rent metodemæssigt skal man være opmærksom på, at IV-metoden kan komme til at give et skævt parameterestimat, hvis der ikke er tilstrækkeligt mange observationer. Det er især et problem, hvis instrumentet ikke er højt nok korreleret med deltagelsesvariablen. Fundamentalt set er der ikke forskel på at betinge på baggrundsvariabler i regressionsmetoden ovenfor og at finde et godt instrument til IV-metoden.

Det skyldes, at det ikke er mere sandsynligt fx at finde et godt instrument end at finde nogle gode variabler, der kan kontrollere for forskellen mellem at deltage og ikke at deltage. Der er imidlertid enkelte situationer, hvor der kan være en oplagt idé til et instrument, der gør, at IV-metoden kan være at foretrække.

Det er fx oplagt at anvende IV-metoden, hvis der er trukket tilfældigt lod om, hvem der tilbydes at deltage, uden at der er tale om et randomiseret forsøg. Den tilfældige udvælgelse fra lodtrækningen giver herved et godt instrument for deltagelse eller ej. Det kræver imidlertid, at et flertal af dem, der udvælges, rent faktisk vælger at deltage, så instrumentet har en god forklaringskraft som erstatning for deltagelsesvariablen. Endvidere er det oplagt, at instrumentet ikke har problemer med at have en sammenhæng med noget af den uforklarede information i regressionen, da det er udtrukket tilfældigt.

Angrist (1998) har en idé til et oplagt instrument, idet han udnytter, at der blev trukket tilfældigt lod om, hvem der blev indkaldt til at aftjene som soldat i Vietnamkrigen. I sit studie undersøger han, om soldater i Vietnamkrigen har en lavere indkomst livet igennem. De mænd, der blev trukket ud, har som hovedregel aftjent deres militærtjeneste, men det var dog muligt for enkelte at undvige, hvis de var syge eller under uddannelse. Der var også enkelte, der ikke var trukket ud, der meldte sig frivilligt.

I eksemplet trækkes der tilfældigt lod om, hvem der skal deltage. Det er imidlertid vigtigt at sondre mellem, om de kun tænkes at deltage, som på engelsk kaldes *intention to treat*, eller om de faktisk deltager, som kaldes *actual treatment*. Hvis personen trækkes tilfældigt ud, angiver vi det ved en variabel $Z_i = 1$, og hvis personen rent faktisk deltager som $D_i = 1$. Ved hjælp af disse betegnelser vil vi i det følgende præcisere, hvad IV-metoden går ud på.

Vi kan til at begynde med krydstabulere dem, der trækkes ud, og dem, der er soldater i Vietnam, hvilket ses i tabel 5.1:

Tabel 5.1

Korrelationen mellem udtrækning til og aftjening af værnepligt.

Z_i (udtrukket)/ D_i (soldat)	Soldat i Vietnam	Civil
Udtrukket til militær-tjeneste	Forholdsvis mange personer	Relativt få personer
Ikke udtrukket til militær-tjeneste	Relativt få personer	Forholdsvis mange personer

Når vi ser på fællesmængden af alle, der enten er udtrukket og/eller har været soldat i Vietnamkrigen og krydser af, hvor de hører til i tabel 5.1, vil vi forvente, at de fleste krydser vil blive placeret i diagonalen. Det betyder, at udtrækningen til soldat er et godt instrument for deltagelse i Vietnamkrigen, da det er højt korreleret med, om de udtrukne faktisk blev soldater. Vi har således, at D_i afhænger af Z_i , så vi også vil få en signifikant parameter, når vi regresserer D_i på Z_i , som vist i metode 5.1. Samtidig blev de amerikanske mænd udtrukket til krigen ved tilfældig lodtrækning, hvilket gør, at deres indkomster uden deltagelse ikke afhænger af dette instrument, dvs. om de deltager eller ej. Hermed er forudsætningen om, at der ikke er en selektionsskævhed i ligning (1) sandsynlig, og instrumentet vil være ukorreleret med restleddet. Bemærk, at vi tilsvarende krævede, at variabelen for deltagelse D_i skulle være ukorreleret med restleddet for, at der ikke var et selektionsproblem i regressionsmetoden.

Metode 5.1

Instrumentet i IV-metoden.

Før korrektion for observerbare karakteristika i regression (2) kunne variabelen D_i for deltagelse eller ej tænkes at være korreleret med udeladt information, der var samlet i restleddet. I stedet for at kontrollere for observerbare karakteristika består IV-metoden i at finde en variabel, dvs. et instrument Z_i , der er korreleret med D_i uden at være korreleret med restleddet e_i . Den endogene variabel D_i skal således kunne forklares ved instrumentet Z_i i en regression:

$$D_i = c + d Z_i + v_i$$

Hvis parameteren d til instrumentet er signifikant i denne regression, vil der være korrelation mellem D_i og Z_i , som beskrevet i tabel 5.1.

I princippet minder IV-estimatet om estimatet fra regressionsmetoden, idet de begge sammenvægter de individuelle resultater. Det blev tidligere vist, at parameterestimatet for deltagelseeffekten i regressionsmetoden gav et homogent estimat af deltagereffekten. IV-estimatet bør ligeledes fortolkes forsigtigt, hvis der er heterogene deltagelseeffekter. Hvis der er en tilfældigt udtrukket variabel for deltagelse eller ej, som beskrevet ovenfor, vil det imidlertid være muligt (Imbens & Angrist, 1994). Hvis der er heterogene deltagelseeffekter, kan IV-estimatet således under en forudsætning om monotonicitet, se nedenfor, fortolkes som en gennemsnitlig effekt af deltagelse for dem, der kan foranlediges til at deltage, ved at de udtrækkes til deltagelse (Wooldridge, 2002). Denne effekt kaldes LATE, der står for *Local Average Treatment Effect*. LATE defineres som *intention to treat effekten* divideret med forskellen i sandsynlighed for deltagelse, alt efter om personen er udtrukket eller ej:

$$\text{LATE} = \frac{\text{gns}(Y_{1i}(z_i=1)) - \text{gns}(Y_{0i}(z_i=0))}{\text{ssh}(D_i(z_i=1)) - \text{ssh}(D_i(z_i=0))}$$

For at vi kan fortolke IV-estimatet som en LATE, skal der være tilfældig udtrækning til tilbud om deltagelse, og det skal være frivilligt at deltage, hvis man får tilbuddet. Det eneste, der kræves, er, at de, der deltager uden at være udtrukket, også ville have deltaget, hvis de var udtrukket, hvilket er forudsætningen om monotonicitet (Wooldridge, 2002).

Angrist (2004) giver et eksempel fra Israel (Angrist & Lavy, 2002), hvor det er nødvendigt at få et immatrikulationsbevis fra gymnasiet for at få lov til at læse på universitetet, hvilket svarer til kravet i Danmark om at få en god karakter. Totalt set blev der udvalgt 40 gymnasier blandt dem, der havde den laveste andel af elever, der immatrikulerede, på ca. 25 pct. Ved tilfældig udvælgelse blev de 40 gymnasier opdelt i to lige store grupper. Den ene gruppe på 20 gymnasier fik tilbud om at deltage i tilbuddet, så der efter aftale med de enkelte rektorer blev udstedt en belønning på 1.400 dollar til de studerende, det lykkedes at få et immatrikulationsbevis. De resterende 20 gymnasier udgjorde en ikke-deltagergruppe, der ikke fik dette tilbud. Blandt de 20 gymnasier, der modtog tilbud om deltagelse, valgte rektorerne fra 15 ud af de 20 gymnasier at være med. *Intention to treat*-effekten er forskellen i andelen af dem, der immatrikulerer på de 20 gymnasier, der fik tilbud om deltagelse, i forhold til andelen, der immatrikulerer, i ikke-deltagergruppen af gymna-

sier. Det viser sig, at der er 7,5 procentpoint flere elever, der får et immatrikulationsbevis i deltagergruppen. Når der korrigeres for, at det kun var 75 pct. af de tilbudte gymnasier, der deltog, kan den vurderede effekt af deltagelse ud fra LATE beregnes til $7,5/0,75$ svarende til ca. 10 procentpoint. Det er faktisk en betydelig effekt, da der kun er 22 pct., der får et immatrikulationsbevis blandt de studerende, der går på gymnasier fra kontrolgruppen. I dette tilfælde svarer LATE til Blooms estimator (Angrist, 2004).

Det er også tænkeligt, at der er nogle i ikke-deltagergruppen, der ad bagveje får adgang til at deltage. Det kan tænkes, at der er 5 ud af de 20 skoler i ikke-deltagergruppen, som det alligevel lykkedes for at komme med i ordningen. I det tilfælde kan man dividere med sandsynligheden for at deltage givet tilbud om deltagelse fratrasket sandsynligheden for deltagelse givet, at man var udtrukket til ikke-deltagergruppen, som her beregnes til 25 pct. Herved skal vi i stedet for dividere de 7,5 med $0,75-0,25$, hvorved LATE-estimatoren giver en effekt på 15 procentpoint.

QUASI-EKSPERIMENTELLE METODER

Der kan tænkes at være en naturlig variation, som skaber grundlag for at sammenligne dem, der deltager, og dem, der ikke deltager. Et quasi-eksperiment er et naturligt eksperiment, der indtræffer, hvis der sker en eksogen hændelse for en gruppe personer, som ikke sker for en anden gruppe af personer, der så fungerer som kontrolgruppe. Den eksogene hændelse kan ofte være en politisk ændring, der kan tænkes at påvirke adfærden for personer, familier, byer mv.

Der kan fx være grundlag for et quasi-eksperiment, hvis to sammenlignelige kommuner har forskellig praksis på et område. For eksempel kan det tænkes, at der er indført børnepasningsgaranti i den ene kommune, uden at det er tilfældet i den anden kommune, så at den naturlige deltagergruppe er kommunen, hvor der er indført børnepasningsgaranti, og den naturlige ikke-deltagergruppe er kommunen, hvor det ikke er indført (Jensen et al., 2006). Det kan også være, at der er to sammenlignelige grupper inden for en kommune, hvor der er en forskellig praksis. Fx kan man inden for en kommune ønske at undersøge, om børnene får færre sygdomme, hvis der indføres en mere restriktiv hygiej-

nestandard i en børneinstitution. Cobb-Clark & Crossley (2003) refererer et studie, hvor man fandt, at vand kunne være smittebærer for kolera. Her var der en naturlig variation, idet husholdninger i London kunne deles op i, om de fik vand fra den øverste eller den nederste del af Themsen, og hvor de sidstnævnte husholdninger havde relativt flere personer, der blev syge.

Til opstilling af et quasi-eksperiment er der to væsentlige betingelser, der skal være opfyldt. Den første er, at det skal være meget veldefineret, hvad der identificerer forskellen mellem det at være i deltagergruppen eller det at være i ikke-deltagergruppen. Den anden væsentlige betingelse er, at personerne skal allokere tilfældigt til deltagergruppe og ikke-deltagergruppe, så vi undgår problemet med selektionsskævhed, jf. ligning (1). I det omfang, de to betingelser er opfyldt, kan det lade sig gøre at sammenligne de to grupper, så det er muligt at bestemme en kausal effekt af indsatsen for dem, der deltager (ATET).

Der kan være flere måder at måle forskellen mellem deltager- og ikke-deltagergruppe på. Det kan være en simpel sammenligning på et givet tidspunkt eller en sammenligning af den samme gruppe før og efter deltagelse, hvor sammenligning af gennemsnit kan kaldes en *difference estimator*. Det kan også være, at der er observationer for de to grupper over tid, hvilket resulterer i et paneldatasæt. Hvis der er et paneldatasæt til rådighed, kan der benyttes en *difference-in-difference estimator* til at vurdere effekten af indsatsen, hvilket Wooldridge (2003) giver en god grundlæggende gennemgang af.

En anden nært beslægtet målemetode er den såkaldte *regression discontinuity design*. Her udnyttes det, hvis der er en tærskelværdi for, hvornår man er berettiget til deltagelse i et program. Det kan fx være en aldersgrænse, indkomstniveau e.l. Hvis man bliver tilbudt deltagelse under en vis tærskelværdi, kan en vurdering af effekten beregnes ved at se på gennemsnittet af resultatet for dem, der er under tærskelværdien, sammenholdt med dem, der er over tærskelværdien, som så udgør ikke-deltagergruppen. Som eksempel kan man beslutte at lave en indsats for dem, der er i aldersgruppen fra 18 til 25 år. Til sammenligning kan man se på, hvordan det går de 26-årige.

For de fleste af disse målemetoder deler man op i en deltager- og en ikke-deltagergruppe, som herefter sammenlignes. For at undgå selektionsskævheden er det som sagt vigtigt, at grupperne uden indsatsen er så sammenlignelige som muligt. Der skal således hverken være observerba-

re eller uobserverbare karakteristika, der gør, at der er forskel mellem grupperne. Hvis det er den samme gruppe, der sammenlignes før og efter deltagelse, skal man også være opmærksom på, at der kan ske ændringer i individernes karakteristika i løbet af tiden.

Problemerne opstår, hvis de politiske ændringer ikke er eksogene. Det kan være, at politikerne under beslutningen om et politisk tiltag er påvirket af de forhold, der kan beskrives som observerbare karakteristika. Når det politiske tiltag er gennemført, kan dette tænkes at spille tilbage på, hvad der kan observeres. Ex post er det så spørgsmålet, om det er neutralt at betinge på, at de observerbare karakteristika er ens for at sammenligne to grupper.

Det må vurderes fra den ene situation til den anden, hvorvidt den politiske ændring er eksogen eller ej. Selvfølgelig er det sjældent, at der er tale om et rent naturligt eksperiment, hvor politikere beslutter tilfældigt, om det skal være den ene eller den anden politik, der skal gennemføres, eller at politiske ændringer kommer som en overraskelse for målgruppen (befolkningen). Kravet om eksogenitet kan imidlertid tænkes at være mere eller mindre opfyldt.

VALG AF EFFEKTMÅLINGSMETODE

Formålet med dette afsnit er at sammenfatte de væsentlige punkter ved de enkelte metoder inden for effektmåling. Afsnittet er tænkt som en oversigt, der gør det muligt hurtigt at orientere sig, så det ikke er nødvendigt at læse metodeafsnittene igen og igen. I tabel 5.2 ses en sammenhæng mellem de enkelte metoder og de væsentlige informationer, man må se på, når den ene metode vælges frem for den anden.

Tabel 5.2

Oversigt over informationskrav for diverse metoder til effektmåling.

	Tilfældig udvælgelse?	Kan ekstra information korrigere for en selektionsskævhed?	Hvordan sammen- vægter metoden individuelle deltagel- seseffekter?
Randomiserede forsøg	Ja	Nej	Frit
Regression	Nej	Ja	Parameter
Matching	Nej	Ja	Frit
IV-estimation	Ja/Nej	Ja	Parameter
Quasi- eksperimentelle metoder	Nej	Nej	Frit/Parameter

Randomiserede forsøg er baserede på, at der er en tilfældig udvælgelse, hvorved det undgås, at der er et selektionsproblem. Hvis protokollen for programmet følges, skulle der heller ikke opstå problemer med den interne validitet. Der udvælges to grupper, der i princippet ligner hinanden så godt som muligt, hvor vurderingen af, om de ligner hinanden, selvfølgelig kan være baseret på foreliggende information. Når metoden anvendes, er det forudsat, at der er taget højde for selektionsskævheden, hvorfor der svares nej i den anden kolonne under de randomiserede forsøg i tabel 5.2. Der kan frit beregnes effekter på baggrund af resultater for deltager- og ikke-deltagergruppe.

Der er ikke tilfældig udvælgelse ved *regressionsmetoden*, hvorfor der står et nej i denne rubrik i tabel 5.2. I regressionsmetoden kan et evt. selektionsproblem muligvis løses ved at korrigere for observerbare karakteristika. Hvis der er forskellige individuelle effekter af en indsats, vil parameterestimatet i regressionsmetoden være et sammenvejet gennemsnit, hvor vægten for hver person er bestemt af de karakteristika, der er valgt at kontrollere for i regressionen.

Matching minder om regressionsmetoden, ved at denne metode også korrigerer for observerbare karakteristika. Til forskel fra regressionsmetoden beregnes der ved matching individuelle effekter, uden at de ligger sammenvejet i et parameterestimat. Det gør det muligt at undersøge disse effekter mere bredt, end tilfældet er for regressionsmetoden.

IV-estimation minder om regressionsmetoden, hvor der imidlertid findes et instrument for den endogene variabel i stedet for at kontrollere for observerbare karakteristika for at løse selektionsproblemet. Denne metode er særligt fordelagtig, når der er ikke-observerbare karakteristika. Det er en fordel at finde et instrument, hvor der er tilfældig udvælgelse, hvorfor der både er sat ja/nej herved i tabel 5.2.

De *quasi-eksperimentelle metoder* er også stærke, når der er uobserverbare karakteristika. Der er ikke nogen tilfældig udvælgelse, hvorfor der står et nej i denne rubrik i tabel 5.2. Det kan imidlertid godt være tilfældet, at der er en naturlig variation i praksis, der minder om et eksperiment. De quasi-eksperimentelle metoder kan sammenveje effekten som et parameterestimat, hvilket fx er tilfældet for *difference estimatoren* og *difference-in-difference estimatoren*. Det er imidlertid også med de quasi-eksperimentelle metoder muligt at analysere effekter mere frit for grupper, hvor der er en naturlig variation.

BEGRÆNSNINGER VED EFFEKT MÅLINGSMETODER

Selvom der i de skitserede metoder ovenfor er åbnet mulighed for at tage højde for individuel heterogenitet, er der stadig en metodebegrænsning i form af, at denne effekt antages at være konstant for det enkelte individ og ikke udsættes for nogen påvirkning. Fx er det realistisk, at det kan have en gunstig effekt, hvis ens venner også deltager i et kursus, da man i så fald vil kunne hjælpe hinanden. Disse interaktionseffekter tages der som regel ikke højde for i de forskellige metoder. Der tages heller ikke altid højde for, at indsatsen kan have indflydelse på andre grupper. I eksemplet med rapporten fra AKF og Finansministeriet, der så på effekten af de videregående efteruddannelseskurser, kan det tænkes, at en eventuel positiv effekt af et efteruddannelseskursus gør det muligt for arbejdsgiveren at stille øgede krav, der får produktiviteten til at stige og dermed giver basis for stigende lønninger. Disse tilbagespilseffekter tages

heller ikke med i betragtning inden for den teoretiske ramme, som vi beskriver her. Vi antager med andre ord, at der hverken er interaktions-effekter med andre deltagere i programmet eller tilbagespilseffekter på deltagere fra andre grupper, der også påvirkes af effekten.

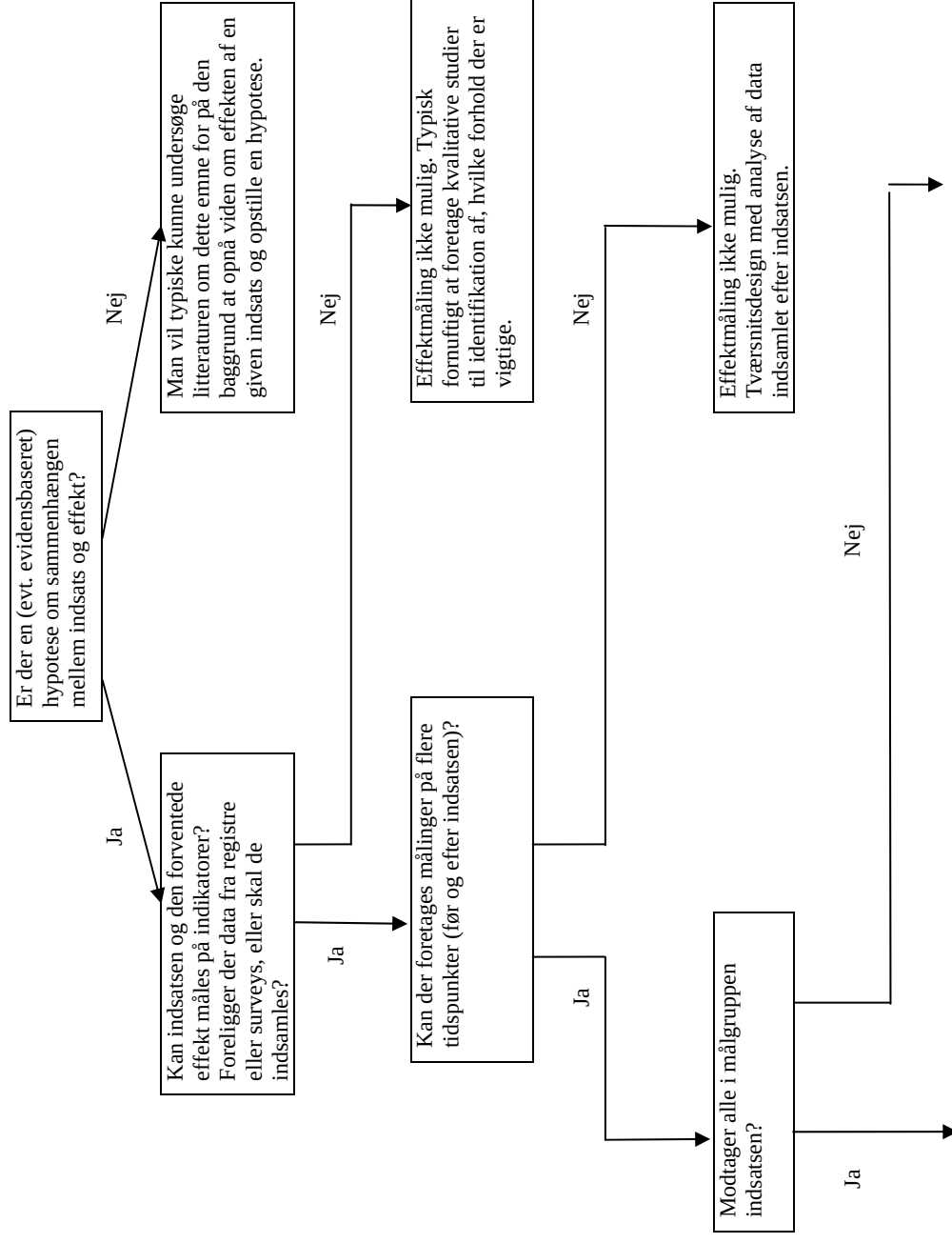
I litteraturen angives disse metodebegrænsninger ofte med betegnelsen SUTVA (*Stable Unit Treatment Value Assumption*). Det står for, at effekten af et program eller en politik kan være varierende for de forskellige personer, men at den vil være konstant for den enkelte person.

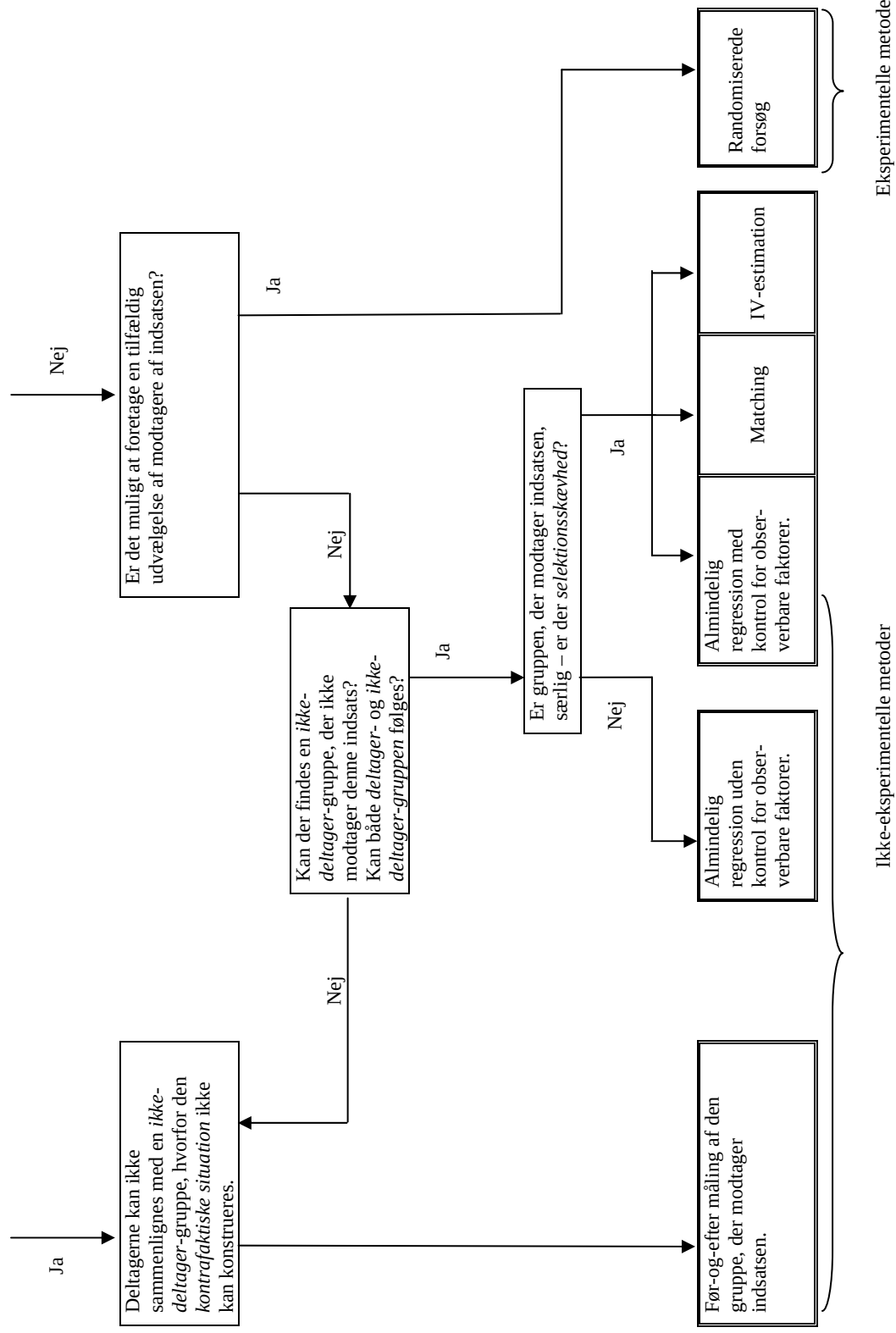
KLASSIFIKATION AF SATSPULJEPROJEKTER

I dette kapitel har vi lavet en oversigt over metoderne ovenfor, der i det følgende skal hjælpe os med at kategorisere satspuljeprojekterne efter, hvilken type information der er til rådighed, og dermed hvilket effektmålingsdesign der egner sig til det enkelte satspuljeprojekt.

Tanken er, at klassifikationsskemaet skal være en hjælp, når læseren sidder med et konkret satspuljeprojekt i tankerne. En række af de centrale begreber, der er blevet anvendt i løbet af gennemgangen af effektmålingsmetoderne i det forrige, er i skemaet kursiverede, så det er nemt at vende tilbage til beskrivelsen af dem. Læserne opfordres til at gennemgå klassifikationsskemaet, når de ønsker at bestemme, hvilke metoder der egner sig til at vurdere effekten for den konkrete indsats.

Klassifikationsskema: Valg af effektmålingsmetode





Selvom de fleste overvejelser over valg af metode er systematiseret i skemaet, findes der også forhold, som ikke er medtaget. Det gælder bl.a. ressourcebetragtninger, idet det for hvert projekt bør overvejes, hvorvidt den viden, der kommer ud af at lave en effektmåling (eller for den sags skyld en anden form for evaluering), står mål med de udgifter, der er forbundet hermed – særligt når det sættes i forhold til, hvad der ellers kunne være opnået med de samme ressourcer.

Herudover inddrager skemaet heller ikke etiske overvejelser i den forstand, at der ikke tages højde for, at det fx i visse situationer kan være uetisk ikke at tilbyde alle en bestemt indsats. Det gælder særligt, hvis denne indsats med stor sikkerhed har en positiv effekt.

Endelig er det for både den eksperimentelle og de ikke-eksperimentelle metoder vigtigt at overveje antallet af observationer, altså hvor mange – typisk enkeltpersoner – en bestemt indsats skal vedrøre. For at kunne lave meningsfulde statistiske beregninger er det således nødvendigt at have et vist antal “enheder”, der både modtager og ikke modtager indsatsen. Kravet til antallet af enheder stiger endvidere med antallet af uafhængige variabler, som det er nødvendigt at kontrollere for.

TO EKSEMPLER PÅ SATSPULJEPROJEKTER

I dette kapitel ser vi på to idéer til satspuljeprojekter for satspuljeforhandlingerne i efteråret 2006. På grundlag af de to projekter diskuteres det, hvordan projekterne er defineret, og der kommes med forslag til, hvad der kan gøre dem mere målrettede til at måle effekter. Det første projekt er indsatsen over for kriminelle unge, som opholder sig på sikrede institutioner, og det andet projekt er en handlingsplan for forebyggelse på ældreområdet – forebyggende hjemmebesøg.

INDSATSEN OVER FOR KRIMINELLE UNGE, SOM OPHOLDER SIG PÅ SIKREDE INSTITUTIONER

På baggrund af Socialministeriets projektskema over indsatsen over for kriminelle unge, der opholder sig på sikrede institutioner, er følgende oplysninger uddraget:

Formål:

- At forbedre indholdet i ungdomssanktionen. På længere sigt er formålet at få kriminelle unge i ungdomssanktionen i gang med en uddannelse, i beskæftigelse og ud af kriminalitet.

Målgruppe:

- Kriminelle unge i varetægtssurrogat og unge i ungdomssanktionen, dog ikke unge, som afsoner straf.

Indsats:

- En grundig udredning af de unge, suppleret med en psykologisk udredning, mens de unge er i varetægtssurrogat, hvilket kan betragtes som en indledende pædagogisk/psykologisk observationsperiode af op til 2 måneders varighed. Herudover tilknyttes en koordinator, når de unge senere modtager en længerevarende behandling i et sikret regi i op til to perioder bestående af hver seks måneder. Der kan således være tale om op til tre faser i behandlingsforløbet.
- Tilknytning af koordinator til den unge har til formål at sikre en sammenhæng i de tre nævnte faser. Herudover vil der ske en opfølgning og inddragelse af relevante parter undervejs i forløbet. Koordinatoren er tillige ansvarlig for at afholde løbende netværksmøder, hvor den unge, forældrene, sagsbehandleren og en repræsentant fra institutionen inddrages og ansvarliggøres. Forsøgsvis ansættes en koordinator på to-tre af de sikrede institutioner.
- En udvikling af metoder/modeller, som kan dokumentere indsatsen, evt. ved at anvende ADAD-systemet.

Operationelle mål:

- Recidiv, altså tilbagefald i forbindelse med kriminalitet, skal reduceres med 10 pct. sammenlignet med unge i ungdomssanktionens første halvandet år og unge, der har været igennem den forbedrede ungdomssanktion.
- 20 pct. flere i ungdomssanktionen skal være i gang med en uddannelse eller i beskæftigelse ved afslutningen af projektperioden (tre år) i forhold til i begyndelsen af projektet.

Der måles på antallet af unge, der har afsluttet ungdomssanktionen, og som er i gang med uddannelse eller i beskæftigelse inden projektstart, og på antallet af unge, der indgår i projektet ved projektets afslutning.

Alle unge i ungdomssanktionen skal have en handlingsplan, som er udarbejdet efter de opstillede kriterier ni måneder efter projektstart. Alle unge og deres forældre skal løbende inddrages i opfølgningen på handlingsplanen.

Dokumentationssystemet skal være tilpasset efter et år og i brug på alle sikrede afdelinger efter tre år.

Effektmåling:

- Effekten af indsatsen til forbedring i ungdomssanktionen måles i forhold til en ikke-deltagergruppe, som består af unge kriminelle under 18 år, der har været idømt og gennemført en ungdomssanktion det første halvandet år, hvor ordningen virkede.
- Effekten af at tilknytte en koordinator til den enkelte unge undersøges i forhold til at fastholde sammenhængen mellem de tre faser og i forhold til de unges recidiv, som sammenlignes med unge, der ikke får tilknyttet en koordinator.

I forhold til ovennævnte oplysninger følger her en vurdering af projektets effektmålingspotentialer:

- målgruppen er velafgrænset
- målene for indsatserne er umiddelbart velbeskrevne
- der er angiveligt tale om to indsatser a) en grundig udredning, b) tilknytning af koordinater på udvalgte institutioner.

Kommentarer til indsatsens del om en grundig udredning

Det er vanskeligt at se, hvad den grundige udredning indebærer og dermed, hvordan den skal få nogen indvirkning på diverse mål for resultaterne, dvs. recidiv, beskæftigelse og uddannelse. Det er derfor vigtigt at afklare, hvad udredningen konkret skal indeholde, og hvordan den skal påvirke behandlingsindsatsen, som igen vil kunne påvirke de nævnte resultater. Udredningen forventes at kortlægge problemer og at inspirere til forbedringer, men det kan ikke forventes, at en bedre udredning i sig selv påvirker resultaterne.

Det er oplagt at udvælge en kontrolgruppe, som består af tidligere behandlede unge, sådan som det foreslås. Da recidiv, uddannelse og beskæftigelse – resultatmålene – formentlig er korrelerede med alder,

køn, etnicitet, opvækstforhold og kriminalitetsadfærd, må der korrigeres herfor. Det betyder, at disse oplysninger må indhentes for begge grupper enten i journaler eller via registeroplysninger. Der må også korrigeres for evt. ændrede behandlingsindsatser over tid, hvilket der derfor må indhentes viden om.

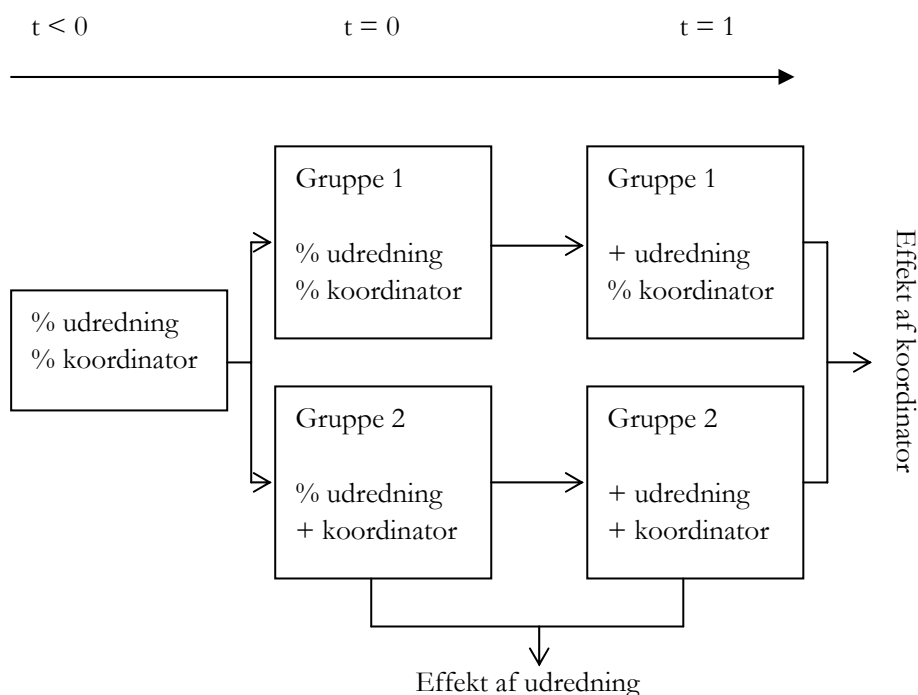
Kommentarer til indsatsens del om tilknytning af koordinator

Med hensyn til tilknytningen af en koordinator i to-tre institutioner er det oplagt at sammenligne unge i institutioner med koordinators med unge i institutioner uden koordinators. Det antages således, at alle unge i de to-tre institutioner får en koordinator, og at ingen i de øvrige får en sådan. Der må endvidere tages højde for, at de unge i de to grupper – behandlingsgruppen og kontrolgruppen – er sammenlignelige på forhold, der kan have betydning for opnåelsen af resultaterne, dvs. alder, køn, etnicitet, opvækstvilkår og kriminalitetsadfærd, ligesom der ikke må være systematiske forskelle i indsatserne over for de to grupper af unge.

De to indsatser, hvor der både indføres en mere grundig udredning og tilknyttes en koordinator, kan tænkes at påvirke samtidigt, så det kan være svært at isolere, hvad effekten af hver del af indsatsen isoleret set er. Det er således muligt, at en grundigere udredning har en større virkning på unge med koordinator end på unge uden koordinator. Det kan fx foregå ved, at koordinatoren påvirker udformningen og indholdet af udredningen rettet mod de unge, som han/hun skal være koordinator for.

Figur 7.1

Effektmåling med eller uden udredning og/eller koordinator.



Figur 7.1 viser en oversigt, hvor effekten af udredning og effekten af tilknytning af koordinator, kan vurderes hver for sig, idet der for effekten af udredningen er kontrolleret for koordinator og for effekten af koordinator er kontrolleret for udredning.

Inden iværksættelsen af indsatsen til tid $t < 0$ er der hverken udredning eller koordinator. I den første fase er der nogle institutioner, hvor de unge får tilknyttet en koordinator, uden at der er foregået en udredning. Andre institutioner får derimod ikke tilknyttet en koordinator, men til gengæld bliver de unge udredte. Det handler herefter om at følge disse to typer af institutioner, hvor der henholdsvis kontrolleres for

koordinator og udredning. Det er således effekten af disse indsatser hver for sig, der måles, idet der netop korrigeres for eventuelle samspilseffekter.

For både *indsats – grundig udredning* - og *indsats – tilknytning af koordinator* er det muligt at følge de unge i registrene, således at der kan anvendes pålidelige oplysninger over en længere periode, hvad angår familieforhold, arbejdsmarkedstilknytning mv.

Hvad angår projektets operationelle mål, vil det være muligt at se, hvor tæt man er kommet på dem ved hjælp af de anvendte indsatser. Hvorvidt andre forhold har medvirket til opnåelse af de opstillede operationelle mål, siger effektmålingen ikke noget om.

HANDLINGSPLAN FOR FOREBYGGELSE PÅ ÆLDREOMRÅDET – FOREBYGGENDE HJEMMEBESØG

På baggrund af Socialministeriets projektskema for handlingsplanen for forebyggelse på ældreområdet er følgende oplysninger uddraget:

Formål:

At styrke kommunernes indsatser på ældreområdet for at opnå

- færre (gen)indlæggelser på sygehuse
- bedre funktionsevne
- reduceret behov for plejehjemsophold
- reduceret dødelighed
- forbedring af ældres trivsel generelt.

Målgruppe:

Ældre over 75 år, som bor i egen bolig.

Indsats:

- Indsamling af viden om ældres sundhedstilstand mv. gennem to forebyggende hjemmebesøg om året.
- En forebyggelsespakke med konkrete initiativer på forebyggelsesområdet:

- a) kost
- b) sociale tilbud
- c) motion
- d) træning
- e) vejledning
- f) aktivering.

Data:

Der foreligger angiveligt følgende tilgængelige data:

- SDU har forskningsdata vedr. forhold for den normale aldring
- data om resultater af træningsindsats
- SFI's ældredatabase om bl.a. funktionsevne med data tilbage i tiden
- Hvidovre-undersøgelsen.

Der planlægges:

- En forundersøgelse (6. februar 2006) i perioden 2006-2007 med det formål at give et kvalificeret bud på, hvilke konkrete initiativer der bedst kan styrke områdeudvælgelse.
- En "forebyggelsespakke" med konkrete tiltag på forebyggelsesområdet. Denne pakke vil kunne offentliggøres i efteråret 2007.

Effektmåling:

Der foreslås gennemført et quasi-eksperimentelt forsøg, hvor forebyggelsesindsatsen sammenlignes mellem forskellige kommuner givet den lokale tilrettelæggelse, og der foreslås korrektion for sociale og demografiske forskelle mv. i kommunernes ældrebefolkninger.

I forhold til ovennævnte oplysninger følger her en vurdering af projektets effektmålingspotentiale:

- Målgruppen er velafgrænset, men også meget heterogen med hensyn til karakteristika, som kan være af betydning for, hvilke indsatser der kan anvendes, og hvilke resultater/effekter der kan forventes opfyldte. Det kunne derfor med fordel overvejes, om målgruppen skulle splittes op efter funktionsevne osv.
- Der er flere indsatstyper ovenfor, som ikke nødvendigvis er lige relevante for alle +75-årige, og som måske kan forventes at påvirke

hinanden. Det må derfor tilstræbes, at hver enkelt person kun udsættes for én indsats i en periode, som forventes at være tilstrækkelig lang til at kunne have nogen forventelig virkning.

- Målene er forskellige, og ligesom med indsatserne er det vigtigt at afgøre, hvilke mål der kan opnås med hvilke indsatser, således at man får nogle enkle årsags-virkningsforhold til effektmålingen. Alene med det nævnte antal indsatser og mål er der 15.625 årsags-virkningsforhold.

Indsats		Mål
Kost	Ja	Færre indlæggelser
	Nej	
Sociale tilbud	Ja	Bedre funktionsevne
	Nej	
Motion	Ja	Behov for plejehjemsophold
	Nej	
Træning	Ja	Reduceret dødelighed
	Nej	
Vejledning	Ja	Trivsel
	Nej	
Aktivering	Ja	
	Nej	

Givet et afgrænset/præciseret årsags-virkningsforhold er det muligt at gennemføre en effektmåling, som:

- Et randomiseret forsøg, hvor der trækkes lod om, hvilke personer der inden for en kommune skal udsættes for indsatsen.
- Matching, hvor personer i kommuner, der hhv. iværksætter og ikke iværksætter indsatsen, sammenlignes, idet der korrigeres for en række baggrundsforhold, dvs. heterogenitet blandt de deltagende personer i de to kommuner.

For begge effektdesign er det vigtigt, at der for de to grupper – deltager- og ikke-deltagergruppe – anvendes de samme indsatser eller indsatskom-

inationer, hvis det ikke er muligt kun at anvende én indsats hhv. ingen indsats.

EFFEKT MÅLING AF 06-SATSPULJEPROJEKTER

I det følgende skema har vi angivet mulighederne for at foretage effektmålingsstudier af indsatsen under satspuljeprojekterne fra 2006. Vurderingen er foretaget ud fra de foreliggende projektbeskrivelser.

Vi har lagt vægt på at foretage en kort beskrivelse af kernepunkterne ved et effektdesign samt en angivelse af, hvorvidt projektet vurderes at høre til i kategorierne:

1. Eksperimentel metode – hvor der i princippet (såfremt myndighed og leverandør er indstillet på det) kan anvendes randomisering af adgang.
2. Ikke-eksperimentel metode, hvor vi gennem forskellige statistiske metoder kan kontrollere for fx selektionseffekter.
3. Uegnet til effektmåling.

Vi har vurderet, at en inddeling af projekter i disse tre overordnede kategorier er den mest hensigtsmæssige måde at klassificere satspuljeprojekter på, da de tre kategorier overordnet angiver, hvilket ambitionsniveau der bør sættes for at belyse tilbuddenes effekter. Særligt forskellen mellem kategori 2 og 3 er central, da denne grænse i praksis vil afgrænse projekterne i en gruppe, hvor det er hensigtsmæssigt og muligt at foretage analyser af projekternes effekter og en anden gruppe, hvor effektmålinger ikke vurderes at være relevante. Den nærmere afgrænsning af,

hvilken rækkevidde kategori 2-projekterne har, kræver en nærmere vurdering af de vilkår og betingelser, der er gennemgået tidligere i manualen, hvilket i nedenstående skema er beskrevet i kategorien effektdesign.

Ser vi på 2006-projekterne, er der en række tilfælde, hvor en randomisering af deltagelsen i princippet er mulig. I de fleste tilfælde dog alligevel ikke-eksperimentelle metoder. Sådanne studier findes især der, hvor der er tale om deltagelse i forskellige typer af afgrænsede indsatser, som fx deltagelse i rådgivning. Der vil i nogle tilfælde kunne afgrænses en ikke-deltagergruppe, og generelt vil disse studier bero på anvendelse af forskellige statistiske metoder til at kontrollere for variation i brugergruppen, herunder til at kontrollere for selektionsskævhed. Der er også tilfælde, hvor en variation i indsatsen mellem kommuner kan udnyttes i et effektstudie.

Der er dog en betydelig del af projekterne, hvor et effektdesign ikke vurderes at være muligt eller relevant. Det gælder særligt en række projekter, der sigter mod vidensopsamling og opkvalificering af medarbejdere etc. Her er det vanskeligt at isolere indsatsen i forhold til brugere. En analyse af fx udviklingen i organisationspersoners kompetencer eller udbredelse af viden/vidensopsamling må vurderes snarere at have karakter af en procesevaluering end af en effektevaluering.

I adskillige tilfælde er de tilgængelige projektbeskrivelser ikke tilstrækkeligt specificerede til, at det er muligt at vurdere potentialet for at opstille et design for effektmåling. Det gælder særligt tilfælde, hvor indsats og/eller forventet effekt ikke er tilstrækkeligt specificeret.

Projekt	Indsats	Målgruppe	Effekt	Effektmetode, beskrivelse	Effektdesign, kategori
Forældreprogrammer	Forældreprogrammer til forældremyndighedsindehaver	Socialt udsatte forældre	Forbedring af børnenes situation. Færre øvrige forebyggelsestiltag	Kontrolgruppe, randomisering principielt muligt forløbsundersøgelse med sammenligning med gruppe, der ikke modtager indsatsen	Eksperimentelt Ikke-eksperimentelt
Negativ social arv	1) Styrkelse af udsatte børns skolegang 2) Støtte til børn i voldsramte familier (behandlings-tilbud)	Børn med socialt udsatte forældre	1) Færre afbrydelser af skoleforløb, færre sociale-/skolemæssige problemer (indikatorbaseret) 2) Bedre 'trivsel'	1) Projektbeskrivelsen må præciseres, for at et muligt effektdesign kan opstilles: Hvilke indsatser er der tale om? 2) Forløbsundersøgelse med sammenligning med gruppe, der ikke modtager indsatsen, randomisering principielt muligt	Eksperimentelt Ikke-eksperimentelt
Forebyggelse af kriminalitet hos unge under 15 år	§38-undersøgelse, forsøgsprojekt i udvalgte kommuner	Kriminelle unge under 15 år	Fald i kriminalitet – på område-niveau og på individniveau	Matching af kommuner og forløbsundersøgelse af individer, randomisering principielt muligt; ellers: sammenligning med kontrolpopulation af unge uden §38-undersøgelse	Eksperimentelt Ikke-eksperimentelt
Projekt Janus	Vidensopsamling og forankring	Behandlere	Øget viden hos sagsbehandlere etc.	Ikke egnet til effektmålingsstudie	Uegnet
Udsatte børn i forældremyndigheds- og samværssager	Tværfaglige møder – kvalificering af socialrådgivere	Indirekte: udsatte børn (og forældre). Direkte: behandlere	Måling af forligsprocent	Før og eftermåling af forligsprocent sammenlignet med kontrolgruppe gennem matching med kommuner uden indsats	Ikke-eksperimentelt

Projekt	Indsats	Målgruppe	Effekt	Effektmetode, beskrivelse	Effektdesign, kategori
Hvervning af plejefamilier for børn og unge med etnisk baggrund	Informationsmateriale	Potentielle (og evt. nuværende) plejeforældre	Flere plejeforældre med etnisk baggrund eller danske plejefamilier med dansk baggrund med viden om etniske børn og unge	Måling af antal plejefamilier, register-/administrativ opgørelse, ikke direkte brugereffekt.	Ikke-eksperimentelt Uegnet
Forsøg med fritidspas til udsatte børn	Tilskud til deltagelse i fritidsaktiviteter for udsatte børn og unge	Udsatte børn	Øget antal børn, der deltager i fritidsaktiviteter	Tælling af antal børn, der deltager, evt. forløbsundersøgelse af børn, der får tilskud, sammenlignet med kontrolgruppe, Randomisering næppe realistisk	Ikke-eksperimentelt
Projekt SAVN	Hjemmeside	Børn af indsatte i fængsler	Brug af hjemmesiden (+ netværksdannelse)	Tælling af brugere af hjemmeside, ikke egnet til effektmålingsstudie	Uegnet
Baglandet Århus	Støttegrupper til unge anbragte, støtte til uddannelse og arbejde	Anbragte børn og unge samt tidligere anbragte	Ikke specifikke målsætninger – 'afhjælpe sociale arv' netværk, modvirke udstødning	Ikke egnet til effektmålingsstudie	Uegnet
Baglandet Vesterbro	Støttegrupper til unge anbragte, generel driftsstøtte	Anbragte børn og unge, samt tidligere anbragte	Ikke specifikke målsætninger – 'afhjælpe sociale arv' netværk, modvirke udstødning	Ikke egnet til effektmålingsstudie	Uegnet
Børns voksenserhverner	Projektledere	Børn med lille voksenkontakt	Styrkelse af frivilligt arbejde (voksenserhverner)	Ikke egnet til effektmålingsstudie	Uegnet

Projekt	Indsats	Målgruppe	Effekt	Effektmetode, beskrivelse	Effektdesign, kategori
Børnerådets Børneinfo samt digitalt panel	Informations-virksomhed			Opgørelse af kontakter. Ikke egnet til effektmålingsstudie	Uegnet
Undersøgelse af vilkårene for børn på asyl-centre	Forskning			Ikke egnet til effektmålingsstudie	Uegnet
Nye og nemmere veje	Koordinerende sagsbehandlere	Handicap-pede	Meget bredt formulerede effekter – mainstreaming, frit valg, selvbestemmelse	Brugertilfredshedsundersøgelse. Ikke muligt at isolere effekter af indsats	Uegnet
Styrkelse af retssikkerhed	Lovforpligtelse for kommuner til at give information om indsats, baggrund, formål til brugere	Brugere under serviceloven	Styrkelse af borgerens retssikkerhed og partsbeføjelser/viden om grundlag for indsats	Ikke egnet til effektmålingsstudie	Uegnet
Behandling af mental retardering	Forskning			Ikke egnet til effektmålingsstudie	Uegnet
Rådgivning om tilgængelighed til byggeri	Rådgivning i forbindelse med byggeri			Ikke egnet til effektmålingsstudie	Uegnet
Undersøgelse af diabetikers kost og kostudgifter	Uklar projektformulering, undersøgelse vs. faktisk rådgivning			Ikke egnet til effektmålingsstudie	Uegnet
Botilbud til unge hjerne-skadede	Botilbud samt metodeudvikling	Unge med hjerneska-de	Forbedring af indlæring og sociale kompetencer	Forløbsundersøgelse, effektmålingsstudie med kontrolgruppe (de, der ikke gives tilbud), ikke randomiseret, statistisk kontrol for selektionseffekter gennem regression eller matching	Ikke-eksperimentelt

Projekt	Indsats	Målgruppe	Effekt	Effektmetode, beskrivelse	Effektdesign, kategori
Center for tegnsprog	Udvikling af undervisningsmetoder m.m.	Døve m.fl.	Forbedring af kommunikationskompetencer for målgruppen	Ikke egnet til effektmålingsstudie	Uegnet
Bofællesskab til handicappede (Landsforeningen Ligeværd)	Botilbud	Handicappede unge med indlæringsvan- skeligheder	Forbedring af livskvalitet, indlæring m.m.	Forløbsstudie, kontrolgruppe af unge, der ikke modtager tilbud, randomisering næppe realiserbar, kontrol gennem statistiske metoder	Ikke-eksperimentelt
Retssikkerhed, sjældne handicap	Styrkelse af vidensdannelse og udbredelse i kommunalt regi	Direkte: organisatorisk målgruppe	Ikke specifikke effekter	Ikke egnet til effektmålingsstudie	Uegnet
Center for autisme	Klubtilbud	Autister	Bedre adgang til tilbud	Brugertilfreds- hedsundersøgelse, forløbsundersøgelse af betydning af klubtilbud, statistisk kontrol af dem, der ikke deltager i tilbud	Ikke-eksperimentelt
Nærdemokrati på handicapområdet (DSI)	Oprettelse af lokalafdelinger	Direkte: organisatorisk målgruppe	Styrkelse af nærdemokrati	Ikke egnet til effektmålingsstudie	Uegnet
Plejeboliggarranti	Garanti for plejebolig inden for to måneder	Ældre plejekrævende	Kortere ventetid	Effektmålingsstudier af hurtigere henvisningspraksis, forløbsstudie, statistisk kontrol for ventetider	Ikke-eksperimentelt
Genoptræning	Styrkelse af kommunikation og samarbejde	Organisatorisk	Brede målsætninger – kortlægning af kommunikation, samarbejdsmodel for tværsektorielt samarbejde	Ikke egnet til effektmålingsstudie, da projektet kun indirekte fokuserer på virkning for brugere	Uegnet

Projekt	Indsats	Målgruppe	Effekt	Effektmetode, beskrivelse	Effektdesign, kategori
Styrket indsats på frivillighedsområdet	Proces i samarbejde med frivillige organisationer, konkrete projekter ikke specificeret	Ensomme svage ældre	Diffuse målsætninger	Ikke egnet til effektmålingsstudie pga. diffuse målsætninger	Uegnet
Styrket indsats over for ældre med anden etnisk baggrund end dansk	Aktiviteter og tiltag over for ældre med anden etnisk baggrund, konkrete tilbud ikke specificeret	Ældre med anden etnisk baggrund	Brede målsætninger, ikke afgrænset i forhold til konkret indsats	Ikke egnet til effektmålingsstudie ud fra projektbeskrivelse	Uegnet
Pensionsportalen	Oprettelse af internetportal	Alle borgere	Øget viden om pensionsforhold	Ikke egnet til effektmålingsstudie	Uegnet
Underøgelse af madservice	Viden/forskning			Ikke egentlig til effektmålingsstudie	Uegnet
Visitering af frihedskæmpere fra udland til danske plejehjem	Visitering til danske plejehjem, dækning af udgifter for kommuner	Tidligere frihedskæmpere bosiddende i udlandet	Forbedring af livskvalitet	Effektmålingsstudie i princip muligt, før og eftermåling sammenholdt med kontrolgruppe, der ikke gives tilbud	Ikke-eksperimentelt
Metoder i samarbejde mellem pårørende og personale på plejehjem	Erfaringsopsamling, undersøgelse, dialog, udviklings- og afprøvningsforløb	Personale på plejehjem	Indirekte effekt af indsatsen for brugerne	Ikke egnet til effektmålingsstudie	Uegnet
Ældreråd	Samarbejde mellem ældre-råd og frivillige organisationer i pilotkommuner – intra-kommunale netværk, øget samarbejde, debatforum, skærpe interesse om ældrepolitik	Medlemmer af ældreråd, indirekte: ældre medborgere	Indirekte effekt af indsatsen for brugerne	Ikke egnet til effektmålingsstudie	Uegnet

Projekt	Indsats	Målgruppe	Effekt	Effektmetode, beskrivelse	Effektdesign, kategori
Videreførelse af projekter under kvinde-krisecenter-pulje	Videreførelse af tilbud, herunder et udvidet antal pladser	Kvinder og familier i krise		Forløbsunder-søgelse, næppe muligt at opstille kontrol-gruppe, men evt. sammenlig-ning mellem centre i forhold til metoder etc.	Uegnet
Videreførelse af psykiatriaf-talen	Botilbud	Psykisk syge børn og unge, dobbeltdi-agnoser	Livskvalitet, stabilisering m.m.	Effektmålings-studie af botilbud, randomise-ring principielt muligt, ellers kontrolgruppe, der ikke tilbydes botilbud, eller variation mellem botilbud, for-løbsstudie	Eksperimentelt Ikke-eksperi-mentelt
Beskæftigel-sesfremmen-de initiativer samt integra-tion af børn og unge	Vejledning, nye aktive-ringsformer, fritidsjob, faglige tiltag, aktiviteter, lektiehjælp m.m.		Øget be-skæftigelse, børn. Fx øgede sprog-kundskaber	Evt. områdestu-dier af områder med forstærket indsats, vanske-ligt at isolere indsats	Ikke-eksperi-mentelt
Udvidelse af SKP-ordning	Udvidelse af SKP-ordning til hjemløse og misbruger-gruppen	Hjemløse og misbru-gere	Fastholdelse i kontakt med myndighe-der, praktisk støtte, styr-kelse af livskvalitet etc.	Effektmåling – fx af betydning af iværksatte tiltag, varighed af kontakt m.m. kræver øget dokumentation af indsats, næppe muligt at anvende kon-trolgruppe, men statistisk kontrol af variation i karakteristika og indsats	Ikke-eksperi-mentelt
Social aktive-ring	Social indsats, forebyggelse behandling,	Socialt udsatte kontant-hjælps-modtagere og førtids-pensionis-ter	Øgede jobchancer	Effektmåling – registerbaseret langtidssstudie af indsats i forhold til beskæftigel-seschancer, statistisk kontrol	Ikke-eksperi-mentelt

Projekt	Indsats	Målgruppe	Effekt	Effektmetode, beskrivelse	Effektdesign, kategori
Permanentgørelse af alkoholteam VFC	Viden/forskning			Ikke egnet til effektmålingsstudie	Uegnet
Permanentgørelse af pro-team VFC	Viden/forskning			Ikke egnet til effektmålingsstudie	Uegnet
VFC – seksuelle overgreb	Vidensindsamling, metodeudvikling			Ikke egnet til effektmålingsstudie	Uegnet
Fremme af kønsligestilling – etniske minoriteter	Indsamling og formidling af viden, konkrete initiativer, men disse er ikke specificeret	Etniske minoriteter	Flere kvinder med etnisk minoritetsbaggrund i foreningslivet	Undersøgelse af andel af kvinder med etnisk minoritetsbaggrund i foreninger, indirekte kobling mellem indsats og effekt, 'makroeffekter'	Ikke-eksperimentelt
Sundhedsfremmeindsats	Modelprojekter, erfaringsindsamling, informationstiltag, ukonkrete tiltag	Udsatte grupper, førtidspensionister, kontanthjælpsmodtagere, arbejdsløse arbejdere	Bedre sundhed i målgruppen	Meget brede målgrupper, ikke klar specifikation af kobling mellem indsats og effekt, ved evt. specificering er det muligt at opstille forløbsstudie på deltagere i indsats	Ikke-eksperimentelt
God social praksis	Websted oprettes	Myndigheder, leverandører	Vidensdeling	Ikke egnet til effektmålingsstudie	Uegnet
Rehabilitering af unge kriminelle	Arbejdsmarkedsrelaterede projekter – samarbejdet mellem kommunale koordinationsudvalg og KIF, mentorordning. Indsats for unge kriminelle med etnisk baggrund (uspecificeret indsats)	Unge kriminelle	Andel i beskæftigelse lavere tilbagefaldsrate	Forløbsstudie af deltagere, sammenligning med kontrolgruppe, der ikke har fået indsatsen, kræver præcisering af målgruppe/afgrænsning af indsats, statistisk kontrol af selektion	Ikke-eksperimentelt

Projekt	Indsats	Målgruppe	Effekt	Effektmetode, beskrivelse	Effektdesign, kategori
Oplysningskampagne om prostitution	Informationskampagner	Unge generelt. Prostitutionskunder	Begrænse efterspørgsel af køb af seksuelle ydelser, holdningsændringer	Måling før og efter kampagne, fx blandt unge på ungdomsudannelser etc. Vanskeligt at isolere kampagnens intervention i holdningsdannelse	Ikke-eksperimentelt
Døgnværested for grønlandere i København	Stoffrit værested	Udsatte grønlandere	Deltagelse i aktiviteter, samvær	Forløbsstudier af brugere, livskvalitet og tilfredshed, vanskeligt at opstille kontrolgruppe/kontrolere for selektion, statistisk kontrol for variation i brugergruppe	Ikke-eksperimentelt
Danmarks Blindebiblioteks digitalisering	Digitalisering af ydelser/bøger/udlån	Blinde	Øget adgang	Brugertilfredshedsundersøgelse. Brug af ydelser, før og efter	Uegnet
Landsforeningen Bedre Psykiatri	Drifts- og udviklingsstøtte, meget brede formuleringer om indsats, døgnåben, webrådgivning, rådgivning af børnefamilier	Pyskisk syge	Styrket rådgivning	Kræver specificering af indsats for at afdække potentiale for effektstudie	Uegnet
Krisecenter for mænd	Botilbud, støtte, rådgivning	Mænd i krise	Øge chancer for at undgå deroute	Forløbsundersøgelse af brugere, randomisering næppe mulig, statistisk kontrol for variation i brugergruppe	Ikke-eksperimentelt

Projekt	Indsats	Målgruppe	Effekt	Effektmetode, beskrivelse	Effektdesign, kategori
Botræning – Reden i Odense	Botilbud, bostøtte	Prostituerede kvinder i misbrug	Bedre behandlingschancer, styrkede kompetencer, bedre livskvalitet	Forløbsundersøgelse af brugere, randomisering næppe muligt, statistisk kontrol for variation i brugergruppe, evt. kontrolgruppe af kvinder, der ikke tilbydes indsatsen	Ikke-eksperimentelt
Forum for brugerindflydelse og metodeudvikling	Metodeudvikling, Brugerforum	Organisation. Handicappede/socialt udsatte	Øget brugerindflydelse	Kræver specificering af indsats/mål for at afdække potentiale for effektmålingsstudie	Uegnet
LAP – driftstilskud	Socialrådgiver – juridisk rådgivning	Psykiatrici-brugere		Brugerundersøgelse tilfredshed med rådgivning, næppe muligt at opstille kontrolgruppe uden rådgivning	Uegnet
Projekt Udenfor	Gadeplansarbejde, maduddeling, omsorg, kontakt	Hjemløse/andre udsatte		Vanskeligt at afgrænse og fastholde brugergruppe mht. forløbsstudier, tællinger af kontakter, brugertilfredshed	Ikke-eksperimentelt
Kirkens Korshær, natcafé i Stengade	Natophold, varmestue, kontakt, omsorg	Hjemløse/andre udsatte		Vanskeligt at afgrænse og fastholde brugergruppe mht. forløbsstudier, tællinger af kontakter, brugertilfredshed	Ikke-eksperimentelt

Projekt	Indsats	Målgruppe	Effekt	Effektmetode, beskrivelse	Effektdesign, kategori
Mødrehjælpen – sårbare mødre	Rådgivning, mentorordning, gruppeforløb	Sårbare enlige mødre	Øge uddannelse og beskæftigelseschancer	Forløbsstudier af deltagere, næppe muligt at opstille kontrolgruppe, statistisk kontrol for indsats	Ikke-eksperimentelt
De grønlandske huse	Styrke de grønlandske huse	Udsatte grønlandere	Mindske tilgangen til de dårligst stillede grupper	Kræver specificering af indsats/mål for at afdække potentiale for effekt-målingsstudie	Ikke-eksperimentelt
Clean House	Nye afdelinger	Tidligere indsatte stofmisbrugere, der er blevet stoffrie	Sikre stoffrihed	Forløbsstudier af ophold i botilbud, væresteder muligt, sammenligning med kontrolgruppe af løsladte, der ikke modtager lignende tilbud	Ikke-eksperimentelt
Reden i Århus	Driftsmidler	Prostituerede		Kræver specificering af indsats/mål for at afdække potentiale for effektstudie	Ikke-eksperimentelt
Forældreforeningen – Vi har mistet et barn	Driftsmidler	Forældre, der har mistet et barn fra 1 år og op efter		Kræver specificering af indsats/mål for at afdække potentiale for effektstudie	Uegnet
Foreningen Grønlandske Børn – et godt liv i Danmark	Samtalegrupper (terapi), m.m.	Grønlandske børn og unge fra udsatte familier	Ikke specificeret	Forløbsstudie principielt muligt	Ikke-eksperimentelt
Kristen rådgivning for incestofre og seksuelt misbrugte	Driftsstøtte, rådgivning, supervision	Seksuelt krænkede	Bedre behandling/rådgivning	Før og eftermåling af brugerens gavn af tilbuddet, før og efter opkvalificering principielt mulig, dog indirekte effekt for brugerne	Ikke-eksperimentelt

Projekt	Indsats	Målgruppe	Effekt	Effektmetode, beskrivelse	Effektdesign, kategori
Støttecenter mod incest	Driftsstøtte behandling/rådgivning	Seksuelt krænkede		Måling af effekter af indsats principielt muligt – før og eftermåling af brugere	Ikke-eksperimentelt
Thora Center	Driftsstøtte, behandling/rådgivning	Seksuelt krænkede		Måling af effekter af indsats principielt muligt – før og eftermåling af brugere	Ikke-eksperimentelt
Frejacentret	Driftsstøtte, behandling/rådgivning	Seksuelt krænkede		Måling af effekter af indsats principielt muligt – før og eftermåling af brugere	Ikke-eksperimentelt
Reden, Odense	Driftsmidler	Prostituerede		Kræver specificering af indsats/mål for at afdække potentiale for effektmålingsstudie	Ikke-eksperimentelt
Center for Selvmordsforskning	Driftsmidler	Selvmonds-truede		Kræver specificering af indsats/mål for at afdække potentiale for effektmålingsstudie	Ikke-eksperimentelt
Dansk Blindsamfund	Konsulentordning	Blinde		Kræver specificering af indsats/mål for at afdække potentiale for effektmålingsstudie	Ikke-eksperimentelt
Kvalitetssikring på anbringelsesområdet, København	Udvikling af manual	Organisation indirekte: anbragte børn og unge		Indirekte virkning for brugerne. Kræver specificering af indsats/mål for at afdække potentiale for effektmålingsstudie	Uegnet

IMPLEMENTERING AF EFFEKTÅLING PÅ DET SOCIALE OMRÅDE

Efter at have set på forskellige effektmålingsmetoder flyttes fokus nu til, hvordan man kan implementere disse. Udgangspunktet er, at effektmålinger rummer store muligheder på det sociale område, men erfaringerne fra andre lande viser, at der kan være forskellige barrierer, som hæmmer implementeringen heraf. I det nedenstående vil der, på baggrund af erfaringerne fra andre lande, fremhæves en række forhold på socialområdet, som kan fungere som barrierer for implementeringen af effektmålinger.

- uenighed om dokumentation: metodediskussionen
- fagkultur og socialisering blandt praktikere
- samspil mellem forskning og praksis
- den politiske udfordring.

BARRIERER FOR IMPLEMENTERING

Som det er skildret i de foregående kapitler, er dokumentationen af effekten af givne tiltag motivationen bag effektmålinger. I skarp kontrast til de randomiserede kontrollerede forsøg, som har været ”guldstandard” inden for effektmåling, har den fremherskende metode til dokumentation af effekter af tiltag inden for det sociale område været kvalitative

interview, som i visse tilfælde er blevet parret med spørgeskemaundersøgelser blandt brugerne og socialarbejderne (Macdonald, 2000). Mens kvalitative interviews kan bidrage til en væsentlig indsigt i brugerne/klienternes individuelle situation og dermed stor forståelse for naturen af specifikke problemer på det sociale område, rummer de det grundlæggende problem, at man kun følger bestemte individer eller en bestemt gruppe, mens den kontrafaktiske situation typisk ikke konstrueres. Indførelsen af effektmåling og de medfølgende metoder, der netop søger at konstruere den kontrafaktiske situation, synes derfor at være et væsentligt skridt på vejen i forhold til at fastlægge effekten af bestemte tiltag.

Til trods for de tilsyneladende oplagte kvaliteter ved effektmålingsmetoderne, der er beskrevet i de foregående kapitler, er disse dog langt fra ukontroversielle. Det gælder særligt i forhold til anvendelsen på det sociale område, og der er fra såvel praktikere som forskere blevet fremført flere argumenter for, at denne tilgang er uegnet til anvendelse i denne kontekst.

Uenighed om dokumentation: metodediskussionen

En grundlæggende kritik vedrører det synspunkt, at det sociale område er for komplekst til at kunne afdækkes ved hjælp af effektmåling. Bag denne kritik ligger i høj grad en modvilje mod at "kvantificere" fænomener inden for det sociale område og dermed definere forskellige målbare parametre, som resultatet af en indsats kan vurderes på baggrund af. I forhold til fx sundhedsforskning, hvor man ofte har klare succeskriterier (grader af helbredelse, længere levetid osv.), hævdes det, at indgreb på det sociale område er mere komplekse og de mulige udfald derfor også mere diffuse og udefinerbare, idet de drejer sig om fx højere livskvalitet eller større modstandsdygtighed. Webb (2001) gør i den forbindelse gældende, at fokus på stringens i effektmålingsoptikken skaber en forkærlighed for målbare fænomener, der ikke nødvendigvis er de centrale i forhold til en given problemstilling. Der er med andre ord risiko for, at man kun forsøger at forbedre forhold, i forhold til hvilke der kan dokumenteres en effekt, og ikke forhold, som er af virkelig betydning for den givne problemstilling. Der er med andre ord risiko for, at man kun forsøger at forbedre forhold, hvor man kan dokumentere effekten, og ikke de forhold, som virkelig har betydning for den givne problemstilling.

Det må konstateres, at de foretrukne effektmålingsmetoder, sådan som de er beskrevet ovenfor, ligger fjernt fra de traditionelle meto-

der anvendt inden for det sociale område. Det skaber problemer i forhold til at implementere effektmålinger på området. På mange måder er debatten om forskellige metoders berettigelse – som er temmelig generel inden for flere samfundsvidenskabelige discipliner – præget af dogmer og ofte ukonstruktiv. De forskellige metoder er oftest bedst, når de supplerer hinanden. Den stærkeste undersøgelse af, hvorvidt et tiltag har en effekt eller ej, vil således ofte omfatte kvalitative forstudier af klienter/brugere og deres kontekst, som så danner udgangspunkt for en senere effektmåling fx i form af et randomiseret kontrolleret forsøg.

Fagkultur og socialisering blandt praktikere

Effektmålinger kritiseres også mere grundlæggende for at være baseret på en urealistisk antagelse om, hvordan socialarbejdere fungerer. Webb (2001) mener således, at den underliggende formodning bag denne tilgang er, at socialarbejdere problemfrit kan afgrænse et givent problem og definere en målsætning for løsning af det. På baggrund af den information, socialarbejderen opnår, er det så forventningen, at de relevante tiltag til løsningen af dette problem ”objektivt” kan identificeres. Webb (2001) kritiserer dette for at være udtryk for en naiv ”objektivisme” og fremhæver, at der ikke findes ét overordnet tiltag, som er det rette, men snarere at enhver socialarbejder er påvirket af egne kulturelle referencer i sit syn på det ”rette tiltag”.

Det er formodentlig de færreste fortalere for effektmålinger, som vil hævde, at man med denne tilgang foreskriver ét tiltag som det rette. Mange vil nok snarere hævde, at tilgangen bidrager med viden om, ”hvad der virker” i forhold til en given problemstilling. Det er så op til socialarbejderne at vurdere, om dette tiltag kan formodes at virke i forhold til en specifik situation (Mullen & Streiner, 2006).

Det er generelt vigtigt at være opmærksom på, at socialarbejderne i et vist omfang kan være påvirket af en *motive bias*, som er et udtryk for, at de dispositioner, som man selv har, spiller ind på, hvilke tiltag og behandlinger man opfatter som de mest hensigtsmæssige. Selvom der findes håndgribelige beviser herpå, vil socialarbejderne således i visse situationer undlade at drage konklusioner, som går imod deres grundlæggende opfattelse af fænomenet.

Socialarbejdernes dispositioner i forhold til bestemte tiltag er afgørende for muligheden for at implementere effektmålinger. I forhold til dette problem må den uddannelse, som medarbejderne på det sociale

område gennemgår, anses for at have en væsentlig socialiserende effekt. Det er her, at socialarbejdernes opfattelse af, hvad der udgør godt socialt arbejde, er blevet formet. Kulturen på uddannelserne inden for det sociale område har generelt været orienteret i retning af mere kvalitative tilgange til undersøgelse af effekten af sociale tiltag, hvilket betyder, at man ikke er vant til at tænke i operationelle indikatorer og i det hele taget produktionen af kvantitative data. Socialarbejderne undervises typisk heller ikke i, hvordan man konkret udfører effektmålinger (Macdonald, 1999). Det betyder, at effektmålinger bliver vanskelige at udføre for socialarbejderne, hvilket formodentligt også fremmer modviljen imod disse, idet man er afhængig af eksterne parter til at hjælpe med udførelsen og derved i et vist omfang kan miste autonomi over eget arbejde.

Grundlæggende synes det at være en afgørende betingelse for implementeringen af effektmålinger, at man indgår i en dialog med uddannelsesinstitutioner om, hvad indholdet bør være i uddannelserne på det sociale område. Denne dialog vil kunne bidrage med bedre færdigheder i forhold til de metoder, som er essentielle for at udføre effektmålinger, og kan medføre en større åbenhed over for alternative tilgange til den traditionelle kvalitative.

Samspil mellem forskning og praksis

En væsentlig betingelse for at lave effektmålinger er en tæt kobling mellem forskning og praksis. En af grundtankerne er således, at den videnskabelige forskning, gennem valide undersøgelser af forskellige tiltags effekt, skal bidrage med nyttig viden til både politiske beslutningstagere og praktikere om, hvilke indsatser der er mest hensigtsmæssige i forskellige situationer. På det sociale område har der tidligere været en tendens til, at praksis og forskning stort set har været løsrevet fra hinanden (Marsh & Fisher, 2005). Der har således manglet en mekanisme til at sikre, at forskellige praktiske problemer blev sat på forskningsdagsordenen, hvorfor forskningen derfor heller ikke har formået at sætte sit præg på praktikernes daglige aktiviteter. Marsh & Fisher (2005) går endnu videre til at sige, at der på det sociale område nærmest findes et omvendt forhold mellem forskning og praksis, således at den gruppe akademikere, der forsker i socialt arbejde, er længst mulig fra praksis, hvilket selvsagt er u hensigtsmæssigt i forhold til at gøre forskningen relevant for praktikerne.

I modsætning til på det sociale område er forskning og praksis integreret på tværs af de forskellige sektorer på sundhedsområdet, hvilket er med til at sikre implementeringen af effektmålinger. Sundhedsområdet er således et eksempel på et felt, hvor forskning og praksis hænger tæt sammen, idet næsten alle læger, der forsker, også fungerer som praktikere. Det medfører, at praktiske spørgsmål og problemer former forskningen. Herudover findes der inden for sundhedsområdet ligeledes formaliserede kommunikationskanaler, særligt tidsskrifter etc., som betyder, at resultaterne af forskningen spredes. Dermed sikrer man et bredt kendskab til forskningsresultaterne blandt praktikere.

Generelt synes der at være behov for et bedre samspil mellem den akademiske socialforskning og det sociale arbejde, der praktiseres af socialarbejderen. Ved at sikre, at forskningen resulterer i forskellige materialer og vejledninger, som er relevante for praktikerne, skabes fundamentet for at kunne udføre effektmålinger på det sociale område.

Den politiske udfordring

Et af formålene med effektmålinger har været en "afpolitisering" af de forskellige sektorområder i den forstand, at ideologiske prioriteringer afløses af, "hvad der virker" (Davies et. Al. 2000). Dette nyttefokus virker umiddelbart appellerende, men erfaringer fra England viser imidlertid, at det ofte bliver ved hensigten, og at de gode intentioner ikke nødvendigvis bliver omsat til handling (Macdonald, 2000).

For at sikre troværdigheden bag effektmålinger er det derfor væsentligt, at politikerne tager implementeringen af disse alvorligt. Der findes flere eksempler på det modsatte, hvor man fra politisk hold har klare præferencer i forhold til bestemte programmer, som effektmålinger så kun har været et middel til at legitimere. Hermed undergraves selve idéen bag effektmålinger, og fundamentet for at implementere disse forsvinder. Praktikerne opfatter hermed ikke effektmålinger som et legitimt forsøg på at fastlægge den bedst mulige politik, men snarere som en måde at gennemtrumfe ideologisk prægede beslutninger på (Nutley & Webb, 2000). Heri ligger formodentlig også kimen til en del af den modvilje over for effektmåling, som har eksisteret blandt mange praktikere.

Ud over at udvise troværdighed i relation til implementeringen af effektmålinger er det ligeledes vigtigt, at politikerne udviser tålmodighed i forhold til at vurdere effekten af bestemte tiltag. På mange måder er det naturligt og forståeligt, at politikerne ønsker at ændre et program,

som tilsyneladende ikke virker, men i flere tilfælde er der således tale om tiltag og programmer, hvor effekten først forventes at kunne konstateres efter et stykke tid. Det er derfor vigtigt, at man accepterer, at det tager tid at udføre disse forsøg.

Det er således særdeles vigtigt i forhold til at udnytte potentialet i effektmålinger, at politikerne "binder deres egne hænder" og udviser troværdighed og tålmodighed for herved at give effektmålinger gunstige vilkår.

ANBEFALINGER

Generelt findes der store potentielle gevinster ved at indføre effektmålinger for såvel politikere som praktikere. Det er derfor nærliggende at lade tankegangen om "hvad der virker" være ledemotivet i socialt arbejde. Et afgørende forhold i relation til at indføre effektmålinger bliver i høj grad at få klargjort principperne bag denne tilgang for de berørte faggrupper. I den forbindelse er det væsentligt at fastslå, at effektmåling ikke betyder, at den traditionelle tilgang til studiet af effekten af forskellige tiltag, som praktikerne tidligere har anvendt, er blevet overflødig. Det er tværtimod meningen, at de to tilgange skal supplere hinanden. Effektmålinger kan således bidrage med viden om, hvad der tidligere har fungeret i forhold til bestemte problemstillinger, men det er stadig socialarbejderen, som skal afgøre, hvorvidt disse tiltag kan forventes at være anvendelige i forhold til en given bruger/klient (Mullen & Streiner, 2006). Socialarbejderens ekspertise er altså stadig afgørende i forhold til de anvendte tiltag over for de enkelte brugere/klienter.

Et andet afgørende forhold i relation til implementeringen af effektmåling er udbredelsen af kendskab til metoderne. Det er væsentligt, at man på de sociale uddannelser inkorporerer metoderne i pensum og dermed udbreder kendskabet hertil. Det vil for det første gøre socialarbejderne i stand til at udføre effektmålinger, hvilket er særdeles afgørende i forhold til at kunne opbygge en database om, "hvad der virker" i forhold til bestemte problemstillinger. For det andet vil undervisning i effektmålingsmetoder give socialarbejderne et fundament for at kunne vurdere gyldigheden af tidligere studier. Herved bliver man i stand til at afgøre, om resultaterne af tidligere studier kan anvendes på en given problemstilling og i det hele taget sondre mellem gode og mindre gode

effektmålinger (Mullen & Streiner, 2006). Dette er væsentligt som et middel til at overvinde en del af den skepsis over for effektmåling, der har været blandt faggrupperne på det sociale område. I forhold til at gøre de socialarbejdere, som allerede arbejder i felten, opmærksomme på styrkerne ved effektmålinger, er det centralt at tilbyde forskellige former for efteruddannelse, som kan hjælpe til en opdatering af de tidligere praksisser.

Som det tredje og sidste forhold, der kan bidrage til en større udbredelse af effektmåling på det sociale område, skal forskernes rolle fremhæves. Det er i den forbindelse vigtigt, som allerede nævnt, at der skabes en bedre sammenhæng mellem forskning og praksis. Det betyder helt konkret, at man fra forskernes side skal være mere opmærksomme på, hvad der er relevant i forhold til de problemstillinger, socialarbejderne møder i felten. Herudover er det særdeles vigtigt, at resultaterne af forskningen gøres tilgængelige for socialarbejderne fx i form af såkaldte *practice guidelines*, der typisk har form af notater, som kort og præcist fremstiller de væsentligste resultater fra et forskningsprojekt.

Effektmålinger rummer altså en lang række muligheder, men der eksisterer omvendt også en række barrierer, som gør det svært at implementere disse. Disse barrierer er dog ikke umulige at overvinde, men det kræver i høj grad en konstruktiv dialog mellem forskerne og de praktikerne, som skal udføre effektmålingerne.

LITTERATUR

- Angrist, J.D. (1998): Estimating the Labor Market Impact of Voluntary Military Service Using Social Security Data on Military Applicants. *Econometrica*, vol. 66, No. 2.
- Angrist, J.D. (2004): American Education Research Changes Tack. *Oxford Review of Economic Policy*, vol. 20 No. 2.
- Angrist, J.D., Imbens, G.W. & Rubin, D.B. (2004): Identification of Causal Effects Using Instrumental Variables. *Journal of the American Statistical Association*, vol. 91.
- Angrist, J.D. & Lavy, V. (2002): *The Effect of High School Matriculation Awards: Evidence from Randomized Trials*. NBER Working Paper 9389.
- Clausen, J., Larson, A., Rosholm, M. & Skipper, L. (2006): *Effekten og oplevet udbytte af deltagelse i voksen-, efter- og videreuddannelse på individniveau*. København: AKF-Forlaget.
- Cobb-Clark, D.A. & Crossley, T. (2003): Econometrics for Evaluations: An Introduction to Recent Developments. *The Economic Record* 79.
- Davies, H.T.O. et al. (2000): Introducing evidence-based policy and practice in public services. I Davies, Huw T.O. (ed.): *What works?* The Policy Press.

- Davies, P., Newcomer, K. & Soydan, H. (2006): Government as Structural Context for Evaluation. I Shaw, I.F., Greene, J.C. & Mark, M.M. (eds): *The SAGE Handbook of Evaluation*.
- Fabricius, N. et al. (2005): *Fra hjemløshed til fast bolig*. København: Socialforskningsinstituttet.
- Imbens, G.W. & Angrist, J.D. (1994): Identification and Estimation of Local Average Treatment Effects. *Econometrica*, vol. 62, No. 2.
- Jensen, P., Simonsen, M. & Smith, N. (2006): Parental Leave Policies. Work in progress. Working Paper.
- Konnerup, M. (2006): *Effekter og evidensbaseret politik og praksis*. Slides fra husmødet i Socialministeriet d. 30. marts 2006.
- Kusek, J.Z. & Rist, R.C. (2004): *A Handbook for Development Practitioners: Ten Steps to a Results-Based Monitoring and Evaluation System*. Washington, D.C.: The World Bank.
- Lee, M-J. (2005): *Micro-econometrics for policy, program, and treatment effects*. Oxford University Press.
- Macdonald, G. (1999): Evidence-based Social Care: Wheels of the runaway. *Public money & management*, januar-marts.
- Macdonald, G. (2000): Social care: rhetoric and reality. I Davies, H.T.O. et al. (eds.): *What works?* The Policy Press.
- Marsh, P. & Fisher, M. (2005): *Using knowledge in social care report 12: Developing the evidence base for social work and social care practice*, Social Care Institute for Excellence.
- Mullen, E.J. & Streiner, D.L. (2006): The Evidence for and Against Evidence-Based Practice. I Roberts, A.R. & Yeager, K.R.: *Foundations of Evidence-Based Social Work Practice*. Oxford University Press.
- Nutley, S. & Webb, J. (2000): Evidence and the policy process. I Davies, H.T.O. et al. (eds.): *What works?* The Policy Press.
- Tsemberis, S., Gulcur, L. & Nakae, M. (2004): Housing first, consumer choice and harm reduction for homeless individuals with dual diagnosis. *American Journal of Public Health* 94 (4).
- Webb, S.A. (2001): Some Considerations on the Validity of Evidence-based Practice in Social Work, *British Journal of Social Work*, vol. 31.
- Wooldridge, J. M. (2002): *Econometric Analysis of Cross Section and Panel Data*. The MIT Press.

Wooldridge, J. M. (2003): *Introductory Econometric: A Modern Approach*, 2nd Edition. Thomson South-Western.

SFI-RAPPORTER SIDEN 2006

SFI-rapporter kan købes eller downloades gratis fra www.sfi.dk. Enkelte rapporter er kun udkommet som netpublikationer, hvilket vil fremgå af listen nedenfor.

- 06:01 Egelund, T.: *Sammenbrud i anbringelser*. 2006. 79 s. ISBN 87-7487-802-6. Kr. 65,00
- 06:02 Holt, H., Geerdsen, L.P., Christensen, G., Klitgaard, C. & Lind, M.L.: *Det kønsopdelte arbejdsmarked. En kvantitativ og kvalitativ bety-
ning*. 2006. 250 s., ISBN 87-7487-804-2. Kr. 228,00.
- 06:03 Rosdahl, A.: *Kommunale aktiveringsprojekter med produktion*. 2006. 51
s. ISBN 87-7487-805-0. Kr. 50,00.
- 06:04 Christensen, E.: *Opvækst med særlig risiko. Indkredsning af børn med
behov for en tidlig forebyggende indsats*. 2006. 92 s. ISBN 87-7487-
806-9. Kr. 85,00.
- 06:05 Jørgensen, M.S., Holt, H., Hohnen, P. & Schimmel, G.: *Job på
særlige vilkår. Overblik over viden på området*. 2006. 83 s. ISBN 87-
7487-807-7. Kr. 75,00.
- 06:06 Rasmussen, M.: *Kontanthjælpsmodtageres gæld. Eftergivelse af offentlig
gæld*. 2006. 68 s. ISBN 87-7487-808-5. Kr. 55,00.
- 06:07 Møller, S.S. & Rosdahl, A.: *Indvandrere i job. Marginalisering og be-
skæftigelse blandt ikke-vestlige indvandrere og efterkommere*. 2006. 171 s.
ISBN 87-7487-809-3. Kr. 160,00.

- 06:08 Bengtsson, S. & Kristensen, L.K.: *Særforsorgens udlægning*. 2006. 96 s. ISBN 87-7487-810-7. Kr. 100,00.
- 06:09 Larsen, M.: *Fastholdelse og rekruttering af ældre. Arbejdspladsers indsats*. 2006. 101 s. ISBN 87-7487-813-1. Kr. 100,00.
- 06:10 Hestbæk, A.-D., Lindemann, A., Nielsen, V.L. & Christoffersen, M.N.: *Nye regler – ny praksis. Ændringerne i servicelovens børneregler 2001. Afslutningsrapport*. 2006. 265 s. ISBN 87-91247-80-2. Rapporten er udgivet af Styrelsen for Social Service. Kontakt denne eller send bestilling pr. e-mail til bestilling@servicestyrelsen.dk.
- 06:11 Olsen, H.: *Guide til gode spørgeskemaer. En manual*. 2006. 100 s. ISBN 87-7487-812-3. Kr. 100,00.
- 06:12 Bonke, J.: *Ludomani i Danmark. Udbredelsen af pengespil og problemspillere*. 2006. 79 s. ISBN 87-7487-811-5. Kr. 85,00.
- 06:13 Müller, M.M.: *Arbejds miljø og indvandrere. Erfaringer i forhold til rekruttering og fastholdelse*. 2006. 92 s. ISBN 87-7487-816-6. Kr. 90,00.
- 06:14 Hansen, H.: *Time Series of APW-Calculations - Module for Great Britain 1991-2004*. 2006. 83 s. ISBN 87-7487-815-8. Netpublikation.
- 06:15 Clausen, J., Heinesen, E. og Hussain, M.A.: *De nye kommuners rammevilkår for beskæftigelsesindsatsen*. 2006. 106 s. ISBN 87-7487-824-7. Netpublikation.
- 06:16 Christensen, G. og Christensen, S.: *Etniske minoriteter, frivilligt socialt arbejde og integration. Afdækning af muligheder og perspektiver*. 2006. 220 s. ISBN 87-7487-817-4. Netpublikation.
- 06:17 Schimmel, G.: *LO-dokumentation nr. 2/2006. Barrierer for kvinder i fagligt arbejde. En kvalitativ undersøgelse af årsagerne til kvinders lavere repræsentation i LO-fagbevægelsen*. 2006. 120 s. ISBN-10: 87-7735-770-1, ISBN-13: 978-87-7735-770-1. Kr. 20,00. Rapporten er udgivet af Landsorganisationen i Danmark.
- 06:18 Boje, T.P.: *Frivillighed og nonprofit i Danmark. Omfang, organisation, økonomi og beskæftigelse*. 2006. 275 s. ISBN 87-7487-821-2. Kr. 250,00.
- 06:19 Boje, T.P., Fridberg, T. og Ibsen, B. (redaktion): *Den frivillige sektor i Danmark. Omfang og betydning*. 2006. 172 s. ISBN 87-7487-822-0. Kr. 160,00.

- 06:20 Geerdsen, P.P., og Geerdsen, L.: *Fra aktivering til beskæftigelse. En gennemgang af aktiveringsindsatsen i det danske dagpengesystem.* 2006. 72 s. ISBN 87-7487-818-2. Kr. 70,00
- 06:21 Jespersen, C.: *Socialt udsatte børn i dagtilbud.* 2006. 108 s. ISBN 87-7487-835-2. Kr. 100,00
- 06:22 Christensen, V.T.: *Uført? Betydningen af nedsat hørelse for arbejdsmarkedstilknytning og arbejdsliv.* 2006. 254 s. ISBN 87-7487-823-9. Kr. 248,00.
- 06:23 Jensen, T.G., Schmidt, G., Jareno, K.N. & Roselius, M.: *Indsatser mod æresrelateret vold.* 2006. 185 s. ISBN: 87-7487-825-5. Netpublikation.
- 06:24 Miiler, M.M., Høgelund, J. & Geerdsen, P.P.: *Handicap & beskæftigelse. Udviklingen mellem 2002 og 2005.* 2006. 128 s. ISBN 87-7487-826-3. Kr. 110,00.
- 06:25 Christensen, E. & Andersen, K.V.: *Livsvilkår for børn med familie på danske asylcentre.* 2006. 120 s. ISBN: 87-7487-827-1. Kr. 120,00.
- 06:26 Rostgaard, T.: *Oplysning om demens. En evaluering af Socialministeriets pulje til oplysning om demens.* 2007. 65 s. ISBN: 978-87-7487-829-2. Kr. 60,00.
- 06:27 Bengtsson, S. & Nemli, A.: *Oplevelsen af MST. Forældres, unges og terapeuters erfaringer med Multisystemisk Terapi.* 2006. 136 s. ISBN: 87-7487-830-1. Kr. 140,00.
- 06:28 Stigaard, M.V., Sørensen, M.F., Winter, S.C., Friisberg, N. & Henriksen, A.C.: *Kommunernes beskæftigelsesindsats.* 2006. 113 s. ISBN 87-7487-832-8. Kr. 90,00.
- 06:29 Madsen, M.B., Mortensøn, M.D. & Rosdahl, A.: *Arbejdsmarkedsparat eller ej? En kvalitativ undersøgelse af visitationen af kontanthjælpsmodtagere i ti kommuner.* 2006. 109 s. ISBN 87-7487-833-6. Kr. 100,00.
- 06:30 Rosdahl, A. & Petersen, K.N.: *Modtagere af kontanthjælp. En litteraturoversigt om kontanthjælpsmodtagere og den offentlige indsats for at hjælpe dem.* 2006. 87 s. ISBN 87-7487-834-4. Kr. 65,00.
- 06:31 Deding, M & Jakobsen, V.: *Indvandreres arbejdsliv og familieliv.* 2006. 101 s. ISBN 87-7487-836-0. Netpublikation.
- 06:32 Deding, M., Lausten, M. & Andersen, A.R.: *Børnefamiliers balance mellem familie- og arbejdsliv.* 2006. 139 s. ISBN 87-7487-837-9. Netpublikation.

- 06:33 Hansen, H.: *Time Series of APW-Calculations - Module for Denmark 1994-2005*. 2006. 121 s. ISBN 87-7487-838-7. Netpublikation.
- 06:34 Christensen, E.: *Uledsagede asylansøgerbørn*. 2006. 64 s. ISBN 87-7487-840-9. Kr. 65,00.
- 06:35 Christensen, V.T.: *Hard of Hearing? Hearing problems and working life*. 2006. 49 s. ISBN 87-7487-823-9. Engelsk sammenfatning af rapporten *Ubørt?* Netpublikation.
- 07:01 Damgaard, B. & Boll, J.: *Opfølgning på sygedagpenge – Del I. Kommuner, lægers, og virksomheders erfaringer med de nye regler*. 2007. 116 s. ISBN 978-87-7487-842-1. Kr. 100,00.
- 07:02 Bach, Henning B. og Petersen, Kirstine N.: *Kontanthjælpsmodtagerne i 2006. En surveyundersøgelse af matchkategorier, arbejde og økonomi*. 2007. 146 s. ISBN 978-87-7487-843-8. Kr. 110,00.
- 07:03 Sivertsen, M.: *Hvordan virker indsatsen mod negativ social arv? Gennemgang og analyse af 54 projektevalueringer*. 2007. 55 s. ISBN 978-87-7487-844-5. Kr. 60,00.
- 07:04 Jespersen, S.T., Junge, M., Munk, M.D. & Olsen, P.: *Brain drain eller brain gain? Vandringer af højtuddannede til og fra Danmark*. 2007. 64 s. ISBN 978-87-7487-846-9. Netpublikation.
- 07:05 Benjaminsen, L.: *Storbypuljen – Indsatser for socialt udsatte. Ideer og erfaringer*. 2007. 47 s. ISBN 978-87-7487-847-6. Kr. 60,00. Pjece.
- 07:06 Müller, M. M., Havn, L., Holt, H., Jensen, S.: *Virksomheders sociale engagement. Årbog 2006*. 2007. ISBN: 978-87-7487-848-3. Kr. 180,00.
- 07:07 Madsen, M. B., Filges, T., Hohnen, Jensen S., Nærvig Petersen, Kirstine: *Vil De gerne have et arbejde?* 2007. 194 s. ISBN 978-87-7487-849-0. Kr. 175,00.
- 07:08 Nielsen, C., Benjaminsen, L., Dinesen P. T., Bonke, J.: *Effektmåling*. 2007. ISBN: 978-87-7487-850-6. Netpublikation.

EFFEKTmåLING

Effektmåling vinder større og større indpas på det sociale område. Derfor har Socialministeriet bedt SFI udarbejde denne manual, der kan fungere som et redskab til at vurdere, hvilken form for effektmåling der er mest frugtbar i forskellige projekter. Især er projekter med bevilling fra satspuljen i fokus.

Manualen gennemgår forskellige metoder til at effektmåle. Metoderne stilles op i et overskueligt skema, der guider praktikere i at vælge, hvilken metode der passer bedst til netop deres projekt. Samtidig rummer manualen en oversigt over alle satspuljeprojekter i 2006 og angiver, hvilken effektmålingsmetode der kan være relevant for hvert projekt.

Effektmåling diskuteres ivrigt på det sociale område, og manualen perspektiverer forskellige standpunkter inden for denne diskussion. Formålet med manualen er, at den skal hjælpe med at forbedre dialogen mellem forskere, politikere og praktikere på det sociale område.