



Europa-
Kommissionen

UDFORMNING OG IVÆRKSÆTTELSE AF KONTRAFAKTISKE KONSEKVENSANALYSER

EN PRAKTISK VEJLEDNING FOR ESF-
FORVALTNINGSMYNDIGHEDER

EUROPA-KOMMISSIONEN

Generaldirektoratet for Beskæftigelse, Sociale Anliggender, Arbejdsmarkedsforhold og Inklusion
Direktorat G — Fonde: programmering og gennemførelse
Kontor G5 — Bedre regulering

Kontakt: Linda Adamaite

E-mail: EMPL-G5-UNIT@ec.europa.eu eller Linda.Adamaite@ec.europa.eu

Europa-Kommissionen
B-1049 Bruxelles

UDFORMNING OG IVÆRKSÆTTELSE AF KONTRAFAKTISKE KONSEKVENSANALYSER

EN PRAKTISK VEJLEDNING FOR ESF-
FORVALTNINGSMYNDIGHEDER

Manuskript afsluttet i september 2021

2021-udgaven af "Udformning og iværksættelse af kontrafaktiske konsekvensanalyser — En praktisk vejledning for ESF-forvaltningsmyndigheder" er udarbejdet af Jochen Kluge (Humboldt-Universität zu Berlin) og Andrea Naldini og Marco Pompili (Ismeri Europa). Dette er en opdateret version af 2013-udgaven af vejledningen (ISBN 978-92-79-28238-6, DOI 10.2767/94454) udarbejdet af Stephen Morris (Policy Evaluation and Research Unit, Manchester Metropolitan University), Herta Tödting-Schönhofer (Metis GmbH, Wien) og Michael Wiseman (George Washington Institute of Public Policy).

Dette dokument er udarbejdet på vegne af Europa-Kommissionen. Det afspejler dog udelukkende forfatterens synspunkter. Europa-Kommissionen er ikke ansvarlig for eventuelle følger af videreanvendelsen af denne publikation. Yderligere information om EU findes på internettet (<http://www.europa.eu>).

Luxembourg: Den Europæiske Unions Publikationskontor, 2022

© Den Europæiske Union, 2022



Politikken for videreanvendelse af Europa-Kommissionens dokumenter gennemføres i henhold til Kommissionens afgørelse 2011/833/EU af 12. december 2011 om videreanvendelse af Kommissionens dokumenter (EUT L 330 af 14.12.2011, s. 39). Medmindre andet er angivet, er videreanvendelse af dette dokument tilladt under en Creative Commons Kreditering 4.0 International (CC-BY 4.0)-licens (<https://creativecommons.org/licenses/by/4.0/>). Det betyder, at videreanvendelse er tilladt med passende kildeangivelse og angivelse af eventuelle ændringer.

Ved enhver anvendelse eller gengivelse af elementer, der ikke ejes af Den Europæiske Union, kan det være nødvendigt at indhente tilladelse direkte fra de respektive rettighedshavere.

Omslag: © Shutterstock, 2021

PDF-udgave

ISBN 978-92-76-40722-5

doi: 10.2767/143803

KE-02-21-992-DA-N

Indhold

INDLEDNING: BAGGRUND OG FORMÅL MED VEJLEDNINGEN	5
KAPITEL 1. KONCEPT OG TILGANGE	11
1.1. HVAD BETYDER KONTRAFAKTISK?	11
1.2. HVORFOR ER KONTRAFAKTISKE EVALUERINGER VIGTIGE?	12
1.3. HVORFOR ER KONTRAFAKTISKE EVALUERINGER TEKNISK UDFORDRENDE?	13
1.4. EN OVERSIGT OVER FORSKELLIGE DESIGNS OG METODER TIL KONTRAFAKTISKE KONSEKVENSANALYSER	15
1.5. HVORDAN KAN KONTRAFAKTISK KONSEKVENSANALYSE INTEGRERES I EN BREDERE EVALUERINGSRAMME?	15
KAPITEL 2. PRAKTISKE OVERVEJELSER I UDARBEJDELSEN AF EN KONTRAFAKTISK KONSEKVENSANALYSE	23
2.1. UDVÆLGELSE AF INTERVENTIONER TIL KONSEKVENSANALYSE	26
2.1.1. <i>Prioritering af interventioner til konsekvensanalyse</i>	28
2.1.2. <i>Udvælgelse af interventioner, der egner sig til en kontrafaktisk tilgang</i>	29
2.2. EVALUERINGSSPØRGSMÅL OG OUTCOMEVARIABLER	34
2.2.1. <i>Hvad er interventionens mål og formål?</i>	34
2.2.2. <i>Hvad er formålet med evalueringen?</i>	34
2.3. DATA TIL IDENTIFIKATION AF KONTROLGRUPPEN OG MÅLING AF OUTCOMEVARIABLERNE	39
2.3.1. <i>Er de relevante data til rådighed, eller kan de stilles til rådighed?</i>	39
2.3.2. <i>Hvordan identificeres den "behandlede" gruppe?</i>	51
2.3.3. <i>Faktorer, der skal tages i betragtning ved identificering af en kontrolgruppe</i>	53
2.3.4. <i>Hvilke typer spørgsmål vedrørende data skal overvejes i evalueringsordningen?</i>	57
2.3.5. <i>Hvad er de vigtigste begrænsninger i forbindelse med analyse af data og resultater?</i>	58
2.3.6. <i>En tjekliste til forberedelse af en kontrafaktisk konsekvensanalyse og vurdering af, om den bør gennemføres</i>	61
2.4. HVILKEN METODE SKAL ANVENDES TIL DEN KONTRAFAKTISKE KONSEKVENSANALYSE?	62
2.5. TIDSPLAN OG BUDGET	63
2.5.1. <i>Hvilke ressourcer er til rådighed?</i>	63
2.5.2. <i>Hvornår skal interventionen evalueres?</i>	68
2.6. GENNEMFØRELSE AF DEN KONTRAFAKTISKE KONSEKVENSANALYSE	70
2.6.1. <i>Udvælgelse af evaluator</i>	70
2.6.2. <i>Tilsyn med den kontrafaktiske konsekvensanalyse</i>	72
2.6.3. <i>Rapportering</i>	72
2.6.4. <i>Anvendelse af resultaterne</i>	74
KAPITEL 3. VALG AF DEN RETTE METODE TIL KONTRAFAKTISK KONSEKVENSANALYSE	77
3.1. RANDOMISERING — DEN EKSPERIMENTELLE TILGANG	77
3.2. IKKERANDOMISEREDE ELLER KVASIEKSPERIMENTELLE DESIGNS	80
3.2.1. <i>Mål- og kontrolgrupper uden randomisering</i>	80
3.2.2. <i>Propensity score matching (matchning efter sandsynlighed)</i>	82
3.2.3. <i>Difference-in-differences</i>	84
3.2.4. <i>Regressionsdiskontinuitetsdesign</i>	87
3.2.5. <i>Instrumentvariabler</i>	90
KAPITEL 4. KONTRAFAKTISKE KONSEKVENSANALYSER — DE NÆSTE SKRIDT	96

4.1. BEDRE FORSTÅELSE BLANDT INTERESSEENTERNE	96
4.2. KAPACITETSUDVIKLING	97
4.3. HÅNTERING AF JURIDISKE HINDRINGER	100
4.4. UDVIKLING I RETNING AF MERE FREMADRETTEDE TILGANGE	100
4.5. UDVIDELSE AF ANVENDELSESOMRÅDET FOR KONTRAFAKTISKE KONSEKVENSANALYSER	101
GLOSSARER	106
AKRONYMER	106
DEFINITIONER	107
BILAG	115
BILAG 1. YDERLIGERE OPLYSNINGER	115
BILAG 2. FORSLAG TIL KURSUS OM KONTRAFAKTISK KONSEKVENSANALYSE	118
BILAG 3. KONTRAFAKTISKE KONSEKVENSANALYSER — EKSEMPLER NÆVNT I VEJLEDNINGEN	119

TEKSTBOKSE

TEKSTBOKS 1 ET EKSEMPEL PÅ EN KONTRAFAKTISK KONSEKVENSANALYSE TIL SAMMENLIGNING AF FORDELE OG OMKOSTNINGER	22
TEKSTBOKS 2 KONTRAFAKTISK KONSEKVENSANALYSE I EN BREDERE EVALUERINGSRAMME	25
TEKSTBOKS 3 SPØRGSMÅL TIL UDVÆLGELSE AF INTERVENTIONER TIL EN KONTRAFAKTISK KONSEKVENSANALYSE.....	27
TEKSTBOKS 4 DE MEST ALMINDELIGE INTERVENTIONSTYPER OG MÅLGRUPPER TIL KONTRAFAKTISKE KONSEKVENSANALYSER FOR ESF	30
TEKSTBOKS 5 DEFINITION AF KONTROLGRUPPER.....	33
TEKSTBOKS 6 EKSEMPLER PÅ DATA TIL KONTRAFAKTISKE KONSEKVENSANALYSER	41
TEKSTBOKS 7 EKSEMPLER PÅ INTEGREREDE DATABASER TIL KONTRAFAKTISKE KONSEKVENSANALYSER	42
TEKSTBOKS 8 EU'S REGELSÆT FOR BEHANDLING AF PERSONOPLYSNINGER.....	46
TEKSTBOKS 9 BESKYTTELSE OG UDVEKSLING AF DATA.....	51
TEKSTBOKS 10 POLITISKE SPØRGSMÅL VEDRØRENDE ET PRAKTIKFORLØB	56
TEKSTBOKS 11 FORTOLKNING AF NETTOEFFEKTERNE	56
TEKSTBOKS 12 USIKKERHEDER I FORTOLKNINGEN AF RESULTATERNE	60
TEKSTBOKS 13 POLENS ERFARINGER MED EVALUERINGSKONFERENCER.....	76
TEKSTBOKS 14 EKSEMPEL PÅ ET RANDOMISERET FORSØG PÅ ET ESF-PROJEKT FOR UNGE	79
TEKSTBOKS 15 ET EKSEMPEL PÅ EN EVALUERING BASERET PÅ EN MATCHNINGSMETODE	84
TEKSTBOKS 16 ET EKSEMPEL PÅ EN EVALUERING BASERET PÅ EN DIFFERENCE-IN-DIFFERENCES-METODE	86
TEKSTBOKS 17 EKSEMPEL PÅ EN EVALUERING BASERET PÅ REGRESSIONSDISKONTINUITET	89
TEKSTBOKS 18 EKSEMPEL PÅ EN EVALUERING BASERET PÅ INSTRUMENTVARIABLER	93
TEKSTBOKS 19 EKSEMPEL PÅ ET PROJEKT, DER HAR TIL FORMÅL AT STYRKE EN KULTUR OG KAPACITET FOR KONTRAFAKTISKE KONSEKVENSANALYSER.....	99
TEKSTBOKS 20 EKSEMPLER PÅ EVALUERINGER PÅ UDDANNELSESOMRÅDET	102
TEKSTBOKS 21 ET EKSEMPEL PÅ VURDERING AF VIRKNINGERNE PÅ "BLØDE OUTCOMES" I TYSKLAND	104

FIGURER

FIGUR 1 FORSKELLIGE OPGAVER OG EVALUERINGSTYPER.....	17
FIGUR 2 ILLUSTRATION AF DEN LOGISKE MODEL ELLER "FORANDRINGSTEORI"	18
FIGUR 3 RÆKKEFØLGE AF DE VIGTIGSTE TRIN I EN KONTRAFAKTISK KONSEKVENSANALYSE	24
FIGUR 4 MINDSTE PÅVISELIGE VIRKNING VED FORSKELLIGE STIKPRØVESTØRRELSER.....	59
FIGUR 5 FORENKLET TIDSPLAN FOR RESULTATERNE AF ET PRAKTIKFORLØB	69
FIGUR 6 RANDOMISERET KONTROL MED TO GRUPPER	78
FIGUR 7 STILISERET KVASIEKSPERIMENTELT DESIGN MED BEHANDLINGS- OG KONTROLGRUPPER	81
FIGUR 8 ILLUSTRATION AF PROPENSITY SCORE-METODEN	83
FIGUR 9 ILLUSTRATION AF DIFFERENCE-IN-DIFFERENCES-METODEN.....	86
FIGUR 10 ILLUSTRATION AF REGRESSIONSDISKONTINUITETSMETODEN.....	88
FIGUR 11 ILLUSTRATION AF METODEN MED INSTRUMENTVARIABLER.....	91

TABELLER

TABEL 1 ANBEFALET INDHOLD I EN EVALUERINGSORDNING	25
TABEL 2 DATATYPER OG KILDER	41
TABEL 3 OVERBLIK OVER DE PRIMÆRE OMKOSTNINGER TIL EN KONTRAFAKTISK KONSEKVENSANALYSE	67
TABEL 4 GRUNDLÆGGENDE OPLYSNINGER, DER SKAL MEDTAGES I EN RAPPORT OM DEN KONTRAFAKTISKE KONSEKVENSANALYSE.....	74
TABEL 5 SAMMENLIGNING AF HOVEDTRÆKKENE I DE PRIMÆRE METODER TIL KONTRAFAKTISK KONSEKVENSANALYSE.....	94
TABEL 6 KARAKTERISTIKA FOR DE KONTRAFAKTISKE KONSEKVENSANALYSER, DER ER NÆVNT I VEJLEDNINGEN SOM EKSEMPLER	119

Indledning: Baggrund og formål med vejledningen

Programmeringsperioden 2021-2027 lagde ud med de dramatiske begivenheder med covid-19-pandemien og den efterfølgende økonomiske krise. Den Europæiske Union (EU) øgede sine finansielle og politiske bestræbelser på at støtte genopretningen af de nationale økonomier og øge beskæftigelsen. Inden for disse rammer spiller Den Europæiske Socialfond plus (ESF+) en central rolle med hensyn til at yde omfattende bistand til arbejdsløse med fokus på unge og kvinder samt støtte foranstaltninger mod børnefattigdom og fremme bedre uddannelse og social inklusion af svagere sociale grupper i hele EU. De forskellige mål, der forfølges med ESF+, og behovet for, at den skal opnå hurtige resultater inden for beskæftigelse og social inklusion, kræver en effektiv tildeling af ressourcer. Evidensbaserede tilgange til politikudformning bliver stadig vigtigere, og evaluering er et grundlæggende instrument til at styre den offentlige politik.

En vejledning til forvaltningsmyndighederne med erfaringerne fra 2014-2020

I perioden 2014-2020 støttede Europa-Kommissionen (Kommissionen) medlemsstaternes og forvaltningsmyndighedernes evalueringskapacitet og fremmede brugen af kontrafaktiske konsekvensanalyser. De positive resultater, der konstateres i kontrafaktiske konsekvensanalyser, er håndgribelige tegn på, at ESF har skabt andre resultater, som ellers ikke ville være opnået. Et betydeligt antal kontrafaktiske konsekvensanalyser blev først planlagt i evalueringsplanerne og derefter udført under gennemførelsen af ESF-programmerne. I mange tilfælde stødte man i forbindelse med kontrafaktiske konsekvensanalyser på vanskeligheder under forberedelsen og gennemførelsen, eller de blev ikke brugt og indgik ikke i en systematisk evalueringsramme.

Denne vejledning henvender sig til forvaltningsmyndigheder og andre organer, der er ansvarlige for gennemførelsen af ESF+-finansierede interventioner og programmer, og har til formål at støtte planlægning, udformning og iværksættelse af kontrafaktiske konsekvensanalyser. Den tager hensyn til tidligere erfaringer og giver praktisk vejledning om nogle af de centrale spørgsmål, der skal tages i betragtning i forbindelse med udarbejdelsen af en kontrafaktisk konsekvensanalyse. Vejledningen ajourfører de tidligere retningslinjer for 2014-2020 og lægger større vægt på de problemer, der er opstået i forbindelse med den praktiske gennemførelse af en kontrafaktisk konsekvensanalyse. Ikke desto mindre drøftes de metodologiske aspekter, som, hvor det er muligt, forenkles og kombineres med ajourførte eksempler fra ESF-evalueringerne for 2014-2020.

Kontrafaktiske konsekvensanalyser tager fat på de afgørende spørgsmål, der gør det muligt at træffe evidensbaserede politiske beslutninger: Hvad er årsagsvirkningerne af interventionerne, og "hvad virker?". De søger at dokumentere, om ændringerne i deltagernes forhold rent faktisk kan tilskrives de ESF-finansierede interventioner, og de deraf følgende resultater af interventionerne. Når de udføres rigtigt,

En kontrafaktisk konsekvensanalyse handler om, "hvad der virker?"

giver de kontrafaktiske konsekvensanalyser dokumentation for en interventions "nettoeffekt"¹ eller virkning, hvilket gør det muligt for de politiske beslutningstagere at udelukke alternative forklaringer på ændringer i deltagernes forhold eller bedrifter, der kan være blevet observeret. Når kontrafaktiske konsekvensanalyser estimerer, at der er en nettoeffekt, og hvor stor den er, er dette nødvendigvis forbundet med en vis usikkerhed, afhængigt af den metodologiske nøjagtighed og de tilgængelige oplysninger. Den type dokumentation, der leveres af kontrafaktiske konsekvensanalyser, gør det muligt for de politiske beslutningstagere at vurdere effektiviteten af interventionerne, foretage sammenligninger mellem interventioner og vurdere deres relative resultater. Den giver desuden vigtige input til cost-benefit-analyser eller omkostningseffektivitetsanalyser.

Denne vejledning offentliggøres i en tid med hidtil usete udfordringer for ESF+. I betragtning af den enorme stigning i EU's ressourcer til investeringer og beskæftigelse som følge af NextGenerationEU-pakken er det afgørende, at de politiske beslutningstagere måler og forstår virkningerne af de interventioner, de er ansvarlige for. Der skal afsættes offentlige midler til mere produktive og effektive interventioner for at fremskynde genopretningen og mindske de sociale ubalancer. Det påhviler derfor de ansvarlige for udbetalingen af ESF+-midler at begrunde deres valg ved at påvise, at deres interventioner er effektive og skaber værdi for borgerne. Den bedste måde at opnå dette på er ved at gennemføre et større antal kontrafaktiske konsekvensanalyser af høj kvalitet.

ESF+ er det vigtigste europæiske instrument til støtte for beskæftigelse og social inklusion. I programmeringsperioden 2014-2020 brugte ESF næsten 125 mia. EUR på aktive arbejdsmarkedspolitikker, uddannelsespolitikker og politikker for social inklusion, der blev gennemført gennem operationelle programmer (OP'er) i de 28 medlemsstater. Som fastsat i forordning (EF) nr. 2013/1303 om fælles bestemmelser foretages der evalueringer "for at forbedre kvaliteten af udformningen og gennemførelsen af programmer og for at vurdere deres effektivitet og virkninger".

I programmeringsperioden 2021-2027 vil resultaterne fortsat blive undersøgt². Dette vil kræve en styrkelse af de nuværende overvågnings- og evalueringssystemer og -kapaciteter, herunder dataindsamlingsordningerne. Evalueringsplanerne vil fortsat være obligatoriske, og der vil blive lagt yderligere vægt på konsekvensanalyse. Da der findes en række forskellige metoder til at opfange virkningerne af de ESF+-støttede operationer, skal forvaltningsmyndighederne beslutte, hvilken metode eller kombination af metoder der er bedst egnet til at opfylde de lovgivningsmæssige krav.

Resultatorientering og evaluering af høj kvalitet

¹ Nettoeffekten, eller virkningen, er forskellen mellem den samlede effekt (eller bruttoeffekt) og det, der ville være opnået i fravær af interventionen. Nettoeffekten kan også være negativ, hvis interventionen er mindre effektiv end markedsdynamikken. Dette er forklaret nærmere på side 9.

² Europa-Parlamentets og Rådets forordning (EU) 2021/1060 af 24. juni 2021 om fælles bestemmelser for Den Europæiske Fond for Regionaludvikling, Den Europæiske Socialfond Plus, Samhørighedsfonden, Fonden for Retfærdig Omstilling og Den Europæiske Hav-, Fiskeri- og Akvakulturfond og om finansielle regler for nævnte fonde og for Asyl-, Migrations- og Integrationsfonden, Fonden for Intern Sikkerhed og instrumentet for finansiel støtte til grænseforvaltning og visumpolitik.

En omhyggelig kvantificering af virkningerne af interventionerne omfatter også den kontrafaktiske situation.

Fokus på gode resultater og resultatorientering er et vigtigt element i de nye forordninger. Evalueringsstrategier og -teknikker af høj kvalitet er afgørende for at høste væsentlig viden, der viser, hvilke interventioner der "virker", og hvilke der ikke gør. Det er vigtigt at styrke evalueringernes kvalitet og udvikle pålidelig dokumentation for merværdi.

I princippet er udgangspunktet for indsamling af dokumentation for effektiviteten af politiske interventioner ligetil. Det kræver svar på følgende:

- Hvilket problem skal løses?
- Hvilke instrumenter skal bruges til at løse det?

En formel, der knytter instrumenter sammen med resultater.

For at kunne evaluere en finansieringsordning eller et finansieringsinstrument er det afgørende at have klare og målbare indikatorer for både de anvendte input og problemrelaterede output og resultater. Det er almindeligt at fastsætte mål for både output og resultater og at sammenligne de faktiske resultater med målene. Overvågning anvendes til at spore input, output og resultater over tid samt sikre feedback fra ledelsen. Den underliggende interventionsteori peger ofte i retning af foreløbige resultater, som også kan blive fokus for overvågningen. Det er dog ikke så simpelt at komme derfra og til at finde ud af, om en bestemt intervention virker.

Behovet for resultater ...

Der er behov for at supplere den eksisterende evalueringspraksis med tilgange, der skaber langt stærkere beviser for interventionernes nettoeffekter eller virkninger. At måle resultaterne handler om at sikre ansvarlighed for de anvendte midler. Kontrafaktiske konsekvensanalyser besvarer det grundlæggende spørgsmål om, hvorvidt en intervention er effektiv. De forsøger at fastslå en årsagssammenhæng mellem interventioner og resultater, men der kan være behov for yderligere teoribaserede og procesbaserede evalueringsmetoder for at identificere de underliggende årsagsmekanismer og bidrage til at sikre, at de virkninger, der opnås på ét sted, giver et evidensgrundlag for at kopiere interventionerne i politikkerne andre steder.

... og dokumentation for nettoeffekter

I programmeringsperioden 2014-2020 vedtog alle medlemsstater og forvaltningsmyndigheder en evalueringsplan, der beskriver evalueringsmål, aktiviteter, organisatoriske elementer og produkter. Evalueringsplanerne omfatter flere typer evalueringer, såsom overordnede løbende evalueringer af programmerne samt tematiske evalueringer, der har til formål at besvare specifikke evalueringsspørgsmål eller fokusere på specifikke akser eller investeringsprioriteter i et OP. En analyse af evalueringsplanerne for 177 ESF-programmer pr. udgangen af 2018³ viste, at der forventes ca. 132 kontrafaktiske konsekvensanalyser i denne programmeringsperiode — et tal, der tyder på fremskridt og større fokus

Medlemsstaternes erfaringer med kontrafaktiske konsekvensanalyser

³ Se [Ismeri Europa — Ecorys — Institute for Employment Studies, 2019](#).

på de kontrafaktiske metoder i programmeringsperioden 2014-2020 sammenlignet med 2007-2013⁴.

Ifølge Evaluation Helpdesk-projektdatabasen⁵ blev der gennemført 1 795 evalueringer frem til juni 2021, hvoraf 1 001 vedrørte ESF/UBI eller ESF/UBI og EFRU-programmer (henholdsvis 675 og 326). Ud af 1 001 evalueringer, der dækkede ESF, var 323 konsekvensanalyser, og 106 af dem anvendte en kontrafaktisk tilgang: 82 ud af 234 konsekvensanalyser vedrørte ESF/UBI-programmer, og 24 ud af 89 konsekvensanalyser vedrørte ESF/UBI-programmer og EFRU-programmer⁶.

Brugen af kontrafaktiske konsekvensanalyser er stigende

I de første år af den nuværende programmeringsperiode fokuserede de kontrafaktiske evalueringer af ESF/UBI-programmerne og/eller EFRU-programmerne på den foregående programmeringsperiode, mens næsten alle kontrafaktiske konsekvensanalyser siden 2018 har vurderet virkningerne af interventionerne i perioden 2014-2020⁷. 20 ud af 27 medlemsstater foretog kontrafaktiske evalueringer af programmeringen, herunder ESF, hvilket også i dette tilfælde afspejlede en mere udbredt anvendelse af denne tilgang i forhold til tidligere.

Selv om de metodologiske tilgange til kontrafaktiske konsekvensanalyser varierer fra medlemsstat til medlemsstat, blev den kontrafaktiske tilgang oftest anvendt til at vurdere virkningen af aktive arbejdsmarkedspolitikker (uddannelse, incitament, hjælp til jobsøgning, arbejdserfaring), og i meget mindre grad på andre områder såsom interventioner til støtte for uddannelsesinstitutioner og studerende og politikker vedrørende sociale spørgsmål, såsom interventioner til bekæmpelse af fattigdom eller social udstødelse.

Metodisk er *propensity score matching* (matchning efter sandsynlighed) den hyppigst anvendte metode, mens andre metoder er mindre udbredte. Kontrafaktiske konsekvensanalyser fokuserer ofte på virkninger på kort eller mellemlang sigt (6 eller 12 måneder), mens langsigtede virkninger vurderes mindre hyppigt.

Den analyse af kontrafaktiske konsekvensanalyser, der blev foretaget i Evaluation Helpdesk-projektet, viste begrænsninger som f.eks.: personer i kontrolgrupper, der ikke i tilstrækkelig grad svarer til dem i støttegrupperne, lav pålidelighed af data, f.eks. når de stammer fra efterfølgende undersøgelser, interview, der bygger på respondenternes evne til at huske hændelser nøjagtigt, eller en lille stikprøvestørrelse, der giver statistisk insignifikante resultater osv. I nogle tilfælde indeholder evalueringsrapporterne ikke tilstrækkelige oplysninger om metodetrin, -valg og -begrænsninger, som er nødvendige for at give andre end dem, der har iværksat dem, oplysninger om virkningerne af

Men der er stadig udbredte begrænsninger

⁴ Se [Bratu C. et al., 2014](#).

⁵ Helpdeskprojektet, som finansieres af GD REGIO og GD EMPL, indsamler oplysninger om de evalueringer, der er blevet offentliggjort online siden den 1. januar 2015 på forvaltningsmyndighedernes websteder. Et netværk af nationale eksperter identificerer evalueringerne, hvorefter de sammenfattes og vurderes med hensyn til resultaternes kvalitet og pålidelighed.

https://ec.europa.eu/regional_policy/en/policy/evaluations/member-states/.

⁶ 90 ud af 106 blev undersøgt og sammenfattet af helpdeskprojektet.

⁷ 29 ud af 33 kontrafaktiske konsekvensanalyser udført frem til 2018 omhandler den foregående programmeringsperiode.

de undersøgte foranstaltninger og således styrke deres viden om virkninger.

Sammenfattende kræver gennemførelsen af kontrafaktiske konsekvensanalyser på trods af de seneste års fremskridt teknisk ekspertise og politisk vilje. Denne vejledning slår til lyd for kontrafaktiske konsekvensanalyser og beskriver nogle af de overvejelser, som forvaltningsmyndighederne skal gøre sig, hvis de skal kunne gennemføre dem. Ud over de praktiske aspekter af kontrafaktiske konsekvensanalyser er der også fokus på de bredere spørgsmål, der eventuelt skal tages op for at sikre konsekvensanalyser af højere kvalitet. Vejledningen indeholder en række grundlæggende anbefalinger, som forhåbentlig er nyttige for forvaltningsmyndighederne, men et af hovedbudskaberne er, at udførelsen af kontrafaktiske konsekvensanalyser kræver omhyggelig planlægning af udarbejdelsen af data (navnlig administrative data), klare mål osv. for at undgå potentielle problemer i gennemførelsesfasen eller evalueringer af lav kvalitet.

En vejledning for praktikere

Vejledningen indeholder også eksempler på evalueringer og praksis til hjælp for læseren. Selv om de evalueringer, der er indsamlet i Evaluation Helpdesk-projektet, er den vigtigste informationskilde til at identificere de eksempler og den praksis, der præsenteres i denne vejledning, blev der også anvendt andre kilder, navnlig de evalueringer, der er foretaget af Det Fælles Forskningscenter (Centre for Research on Impact Evaluation — CRIE), og i mere begrænset omfang akademiske publikationer. Hvor det er muligt, omhandler eksemplerne og erfaringerne i denne vejledning Den Europæiske Socialfonds interventioner.

Vejledningen er opdelt i fire kapitler.

Kapitel 1 beskriver kontrafaktiske konsekvensanalyser, og hvorfor de er vigtige. Det giver en indledende oversigt over metoder til kontrafaktiske konsekvensanalyser og forklarer sondringen mellem eksperimentelle og kvasiekperimentelle metoder. Den beskriver også de tekniske og praktiske udfordringer, der skal tackles i forbindelse med gennemførelsen af en kontrafaktisk konsekvensanalyse. Her inddrages generelt de typer politiske spørgsmål, som kontrafaktiske konsekvensanalyser kan behandle, og forholdet mellem kontrafaktiske konsekvensanalyser og andre evalueringsmetoder (f.eks. teoribaserede tilgange, procesevaluering og effektivitetsanalyse).

En struktur, der afspejler trinnene i en evaluering

Kapitel 2 ser på en række spørgsmål, som forvaltningsmyndighederne bør overveje, når de udformer og gennemfører kontrafaktiske konsekvensanalyser. Denne vejledning beskriver nogle af de centrale udfordringer, som man ofte møder i forbindelse med udarbejdelse af kontrafaktiske konsekvensanalyser, og indeholder nogle anbefalinger til, hvordan disse kan tackles. Spørgsmålene kan vejlede dem, der ønsker at iværksætte kontrafaktiske konsekvensanalyser af ESF-finansierede interventioner.

Kontrafaktiske konsekvensanalyser giver solid dokumentation for midlernes effektivitet. Det gør de imidlertid kun, hvis de planlægges godt og gennemføres korrekt. Med henblik herpå skal forvaltningsmyndighederne fokusere på visse centrale spørgsmål, når de iværksætter en evaluering. Hvordan forvaltningsmyndighederne

præcist griber disse spørgsmål an, og i hvilken rækkefølge, vil være dikteret af praktiske forhold og institutionelle ordninger på stedet i medlemsstaten. I denne vejledning fremhæves nogle af disse vigtige spørgsmål, og forvaltningsmyndighederne opfordres til at gennemgå dem. De vigtigste emner er opstillet i henhold til seks trin i evalueringsprocessen: 1) udvælgelse af den eller de operationer, der skal vurderes, 2) identifikation af evalueringsspørgsmålene og outcomevariablerne, 3) analyse af data med henblik på at identificere kontrolgruppen og måle outcomevariablerne, 4) valg af metode til kontrafaktisk konsekvensanalyse, 5) fastlæggelse af tidsplanen og budgettet og 6) gennemførelse af evalueringen. Der lægges særlig vægt på datatilgængelighed og -beskyttelse, som ofte kan stå i vejen for muligheden for at udføre en kontrafaktisk konsekvensanalyse.

Kapitel 3 omhandler trin 4 ovenfor (valg af metode til kontrafaktisk konsekvensanalyse) og fokuserer på hovedtrækkene ved de primære metoder, der anvendes til kontrafaktiske analyser. Der redegøres navnlig for fem metoder: den eksperimentelle metode eller randomisering, propensity score matching (matchning efter sandsynlighed), difference-in-differences, regressionsdiskontinuitetsdesign og instrumentvariabler. Dette afsnit i vejledningen har ikke til formål at lære læserne, hvordan man bruger metoderne, men blot at beskrive hver metodes styrker og svagheder, og i hvilke situationer én metode er mere velegnet end en anden.

Kapitel 4 omhandler bredere spørgsmål om den fremtidige udvikling. Disse omfatter behovet for at udvikle kapacitet til at udføre vellykkede kontrafaktiske konsekvensanalyser, både inden for forvaltningsmyndighederne (politiske beslutningstagere og embedsmænd) og i medlemsstaternes forskningsmiljøer, akademiske kredse og konsulentvirksomheder. I dette afsnit behandles også behovet for at imødegå juridiske hindringer for dataadgang og videreudvikle de kontrafaktiske konsekvensanalyser af ESF+-programmer til mere avancerede design for at dække politikken for social inklusion og uddannelsespolitikken, vurderingen af virkningerne for "bløde outcomes" og gennemførelsen af metaevalueringsmetoder.

Sammenfattende kan det siges, at denne vejledning: 1) slår til lyd for kontrafaktiske konsekvensanalyser, 2) identificerer de vigtige skridt hen imod en vellykket gennemførelse af kontrafaktiske konsekvensanalyser og 3) betragter kontrafaktiske konsekvensanalyser som en væsentlig del af ESF+-landskabet. Det endelige mål er at øge ESF+-programmets bidrag til EU-borgernes velfærd.

Forfatterne ønsker at anerkende den værdifulde støtte og bistand, som medlemmerne af kontor G5 i GD for Beskæftigelse, Bedre Regulering, har ydet i forbindelse med udarbejdelsen af denne rapport, navnlig Linda Adamaite, Jeannette Monier og Maria José Cueto Faus. Forfatterne vil også gerne takke Andrea Pisano og Ernesto Belisario for deres bidrag til afsnittet om databeskyttelse samt Paweł Hess for hans støtte og bidrag.

Kapitel 1. Koncept og tilgange

I dette kapitel behandles grundlæggende spørgsmål om arten af og formålet med den kontrafaktiske tilgang til vurdering af interventioners årsagsvirkninger. Det beskriver navnlig essensen af den kontrafaktiske konsekvensanalyse, især for så vidt som den vedrører de typer operationer, der medfinansieres gennem ESF+. Det undersøger også forholdet mellem kontrafaktiske tilgange og andre evalueringsmetoder og drøfter, hvorfor kontrafaktiske konsekvensanalyser er vigtige — navnlig på nuværende tidspunkt. De politiske spørgsmål, som kontrafaktiske konsekvensanalyser kan dække, undersøges, og der gives en kort oversigt over nogle af de vigtigste kontrafaktiske metoder, der er relevante for evalueringen af ESF+-finansierede interventioner, inden de beskrives nærmere i kapitel 3.

1.1. Hvad betyder kontrafaktisk?

Kontrafaktiske konsekvensanalyser søger at identificere interventionernes nettoeffekter eller virkninger. Deres særlige kendetegn er, at de har til formål at støtte påstande om, at en given intervention giver et specifikt resultat. Det vil sige, at det konkrete resultat udelukkende kan tilskrives interventionen. Kontrafaktiske konsekvensanalyser opnår dette ved at isolere interventionen og udelukke alternative forklaringer på det observerede resultat.

Muligheden for at udelukke alternative forklaringer er baseret på ideen om "kontrafaktisk", dvs. svaret på spørgsmålet "Hvad ville der være sket — med hensyn til de ønskede outcomes — hvis interventionen ikke var blevet gennemført?". For at forstå begrebet kontrafaktisk og forklare det helt enkelt kan man se på eksemplet med en arbejdsløs person, der deltager i et praktikforløb, der har til formål at hjælpe vedkommende i beskæftigelse. Med henblik på at fastslå praktikforløbets effekt for personen indgår der i den kontrafaktiske tilgang to potentielle resultater⁸. Det første er praktikantens beskæftigelsesstatus efter at have deltaget i praktikforløbet. Dette er det observerede resultat for praktikanten. Det andet potentielle resultat er denne praktikants beskæftigelsesstatus, hvis han eller hun ikke havde deltaget i praktikforløbet, alt andet lige. Under disse omstændigheder betegnes dette andet resultat som det kontrafaktiske resultat. Virkningen af praktikforløbet for den enkelte praktikant bestemmes som forskellen mellem de observerede og de kontrafaktiske resultater. Dette er den kausale nettoeffekt eller virkning af praktikforløbet for den enkelte. Den eneste forskel mellem de omstændigheder eller betingelser, der gav anledning til de observerede og kontrafaktiske resultater, er personens deltagelse i praktikforløbet. Derfor skal enhver forskel mellem de to resultater være praktikforløbets indvirkning på personens beskæftigelsesstatus.

I virkeligheden observerer vi ikke, og vi kan ikke observere, kontrafaktiske resultater for personer, der eksponeres for en intervention. Hovedformålet med kontrafaktiske konsekvensanalyser er således at tilvejebringe overbevisende skøn over kontrafaktiske resultater for grupper af personer eller virksomheder, der er genstand for ESF-medfinansierede interventioner. Virkningerne udtrykkes således f.eks. i form af forskelle i midler eller forhold

Det kontrafaktiske resultat

Definition af behandlingsgrupper og ...

⁸ Der findes en mere detaljeret diskussion af årsagsmodellen med "potentielle outcomes" i [Holland P., 1986](#).

mellem gennemsnitlige observerede og "estimerede" kontrafaktiske værdier. I de fleste tilfælde søges der med kontrafaktiske konsekvensanalyser at sammenligne resultaterne af en intervention (en foranstaltning eller et instrument) for disse enheder (personer, SMV'er osv.), der har været omfattet af den, med en gruppe, der ikke har været omfattet af interventionen. Inden for rammerne af kontrafaktiske konsekvensanalyser skelnes der mellem "behandlingsgruppen" og "kontrolgruppen", som i alle henseender (bortset fra de behandlinger, der modtages) så vidt muligt bør svare til den behandlede gruppe. Det er fra kontrolgruppen, at der opnås skøn over kontrafaktiske resultater, idet der lægges særlig vægt på forskelle i karakteristika — observerede og ikke observerede — mellem de to grupper. Det er også muligt at sammenligne en række forskellige behandlinger ved at eksponere relevante enheder for en række behandlingsvarianter (f.eks. andre ESF-finansierede behandlinger eller interventioner, der finansieres gennem andre kilder), danne en række behandlingsgrupper og sammenligne resultater mellem dem og/eller resultater for en ikkebehandlet kontrolgruppe.

Hvis kontrolgruppen ikke modtager behandling, er evalueringsspørgsmålet "Hvad er virkningen af at modtage interventionen i forhold til ikke at modtage nogen hjælp eller støtte?". Modsat gælder det, at hvis resultaterne af at modtage den pågældende behandling sammenlignes med resultaterne af at modtage en anden behandling, er evalueringsspørgsmålet: "Hvad er virkningen af at modtage interventionen i forhold til et veldefineret alternativ?". En kontrafaktisk konsekvensanalyse kan i mange tilfælde udformes, så den giver svar på et af disse grundlæggende spørgsmål. Valget af, hvilket spørgsmål der skal behandles, afhænger af de politiske beslutningstageres prioriteter og de praktiske begrænsninger i designet.

... kontrolgrupper

I tilfælde, hvor der foretages en sammenligning mellem to forskellige behandlinger, bør der være en klart defineret kontrast mellem dem, hvilket giver mening i forbindelse med politikudformning.

1.2. Hvorfor er kontrafaktiske evalueringer vigtige?

Kontrafaktiske konsekvensanalyser søger at identificere interventionernes nettoeffekter eller virkninger. De giver skøn over virkningernes omfang, deres tegn (positivt eller negativt) og statistiske mål for usikkerhed. De bidrager til at validere eller afvise den formodede årsagssammenhæng mellem interventionen og resultaterne, som ligger til grund for udformningen af interventionen. Disse målte virkninger kan anvendes til at forfølge en række mål: at udvise gennemsigtighed og ansvarlighed i forbindelse med anvendelsen af skatteydernes midler, at træffe politiske beslutninger på grundlag af klar dokumentation, at lære af resultater på tværs af regioner i samme land eller på tværs af medlemsstaterne.

De ansvarlige for interventionerne og for at sikre, at deres programmer fortsat kan tiltrække finansiering, vil have stor interesse i at få udført kontrafaktiske konsekvensanalyser for at vise, at deres programmer giver valuta for pengene og målbare fordele for deltagerne og samfundet som helhed. Dokumentation fra kontrafaktiske konsekvensanalyser vil være af særlig interesse for dem, der er ansvarlige for ressourcetildelingen. Forvaltningsmyndighederne vil være ivrige efter at vise, at deres programmer rent faktisk "virker". For at gøre dette på en overbevisende måde skal de have kontrafaktiske konsekvensanalyser af høj kvalitet.

Gennemsigtighed og ansvarlighed

Det er disse ting, der gør, at kontrafaktiske konsekvensanalyser kan give vigtige oplysninger til de politiske beslutningstagere, hvis opgave det er at afsætte ressourcer til forskellige interventioner. Beslutningerne om finansiering af potentielle interventioner træffes i en situation med ressourcebegrænsninger. I den forbindelse har beslutningstagerne brug for solid dokumentation for programmets virkninger og omkostningseffektivitet, så de kan udnytte de disponible ressourcer bedst muligt. Ved vurderingen af interventionernes relative effektivitet kan de nettovirkninger, som de kontrafaktiske konsekvensanalyser estimerer, vises mere detaljeret ved at undersøge en interventions omkostningseffektivitet eller foretage en fuldstændig cost-benefit-analyse.

Støtte til beslutninger om ressourcefordeling

En sammenligning af nettoeffekterne af lignende operationer, der er gennemført i forskellige områder, kan også sikre mere vigtig viden. Hvis der i kontrafaktiske konsekvensanalyser opdages betydelige forskelle i effekterne, betyder det sandsynligvis, at kvaliteten af operationerne varierer betydeligt, eller at nogle operationer passer bedre i deres socioøkonomiske kontekst end andre. Disse sammenligninger — generelt kaldet "metaevalueringer" eller "metaanalyser" — gør det muligt at opnå en højere generalisering af de kontrafaktiske konsekvensanalyseres resultater og udgør vigtige test for politiske foranstaltninger (f.eks. uddannelse og integrerede tjenester osv.)⁹. For at foretage disse sammenligninger i henhold til videnskabelige kriterier er det nødvendigt med et tilstrækkeligt antal og en systemisk gennemførelse af kontrafaktiske konsekvensanalyser. Derfor er god koordinering på nationalt plan eller EU-plan en grundlæggende forudsætning for at foretage disse sammenligninger hyppigere og gøre dem mere anvendelige i forbindelse med politiske beslutninger.

Sammenligning af aktiviteter i forskellige regioner eller medlemsstater

1.3. Hvorfor er kontrafaktiske evalueringer teknisk udfordrende?

Der findes en række tilgange, der kan beskrives som "upålidelige" forsøg på at vurdere en interventions virkninger. Disse diskuteres her for at illustrere kompleksiteten i kontrafaktiske konsekvensanalyser, og der henvises ikke til den faktiske evalueringspraksis.

Et skridt videre end de forenklede tilgange

For det første kan en politisk beslutningstager ønske at evaluere virkningen af et praktikforløb for arbejdsløse ved at sammenligne indkomst for praktikanter efter uddannelsen med indkomsten for alle arbejdsløse, der ikke deltog. Den politiske beslutningstager tilskriver derefter praktikforløbet den konstaterede indkomstforskel mellem deltagere og ikkedeltagere.

Dette er ikke en gyldig strategi til at identificere praktikforløbets indvirkning på indkomsten, fordi andre personer på mange måder kan være forskellige fra praktikanterne, og disse forskelle kan påvirke resultaterne — en sådan tilgang udelukker derfor ikke alternative forklaringer på eventuelle konstaterede indkomstforskelle. For eksempel kan praktikanter have bedre naturlige evner end andre. Med andre ord kan arbejdsløse med bedre evner melde sig frivilligt til at deltage i praktikforløbet. Disse evner påvirker således ikke kun beslutningen om at deltage, men også resultaterne — arbejdsløse med bedre naturlige evner har større sandsynlighed for at få en højere indkomst end mindre dygtige personer. Som følge heraf kan enhver

⁹ Se f.eks. [Card D., Kluve J. og Weber A., 2017.](#)

konstateret forskel i indkomst mellem behandlings- og kontrolgrupper skyldes både praktikforløbet og/eller iboende forskelle i evner, og man ville således "overvurdere" virkningen af interventionen.

Hvis evner ikke kan måles, og forskelle i naturlige evner mellem de to grupper ikke kan tages i betragtning i vurderingen af virkningerne ved at sammenligne outcomes mellem de to grupper, siges den anslåede virkning af praktikforløbet at lide under **selektionsbias**. For at løse dette problem forsøger evaluatorene at indsamle så mange oplysninger som muligt om vigtige faktorer, der påvirker beslutningen om at deltage og de deraf følgende outcomes. Disse data anvendes til at udvælge en valid kontrolgruppe blandt ikkedeltagende personer — dvs. en gruppe, der så vidt muligt svarer til behandlingsgruppen med hensyn til disse faktorer — og foretage statistiske analyser under hensyntagen til disse faktorer. I den forbindelse anlægger evaluatorene ofte den antagelse, at udvælgelsen i programmet bestemmes af disse observerbare faktorer. Denne "identificerende antagelse" kan ikke uden videre valideres generelt, og evaluatoren skal på overbevisende måde argumentere for, at antagelsen er plausibel i den givne kontekst på grundlag af kendskab til institutionelle faktorer og adfærdsteori.

Modvirkning af selektionsbias

En anden "upålidelig" tilgang kan være, at den politiske beslutningstager observerer indkomsten for praktikanter før og efter praktikforløbet og tilskriver interventionen ændringen fra før til efter. Denne tilgang forudsætter i det væsentlige, at gennemsnitsindkomsten i mangel af interventionen forbliver uændret.

Før og efter ændringen

Igen er dette i næsten alle tilfælde ikke en gyldig strategi for måling af praktikforløbet indvirkning på indkomsten, medmindre antagelsen om tidsmæssig stabilitet med rimelighed kan anvendes. Dette skyldes, at praktikanternes indkomst uundgåeligt vil ændre sig over tid, uden at det på nogen måde har forbindelse til praktikforløbet. Det er f.eks. almindeligt at se, at praktikanternes indkomst dykker forud for deltagelsen, til dels på grund af forbigående faktorer. I mange tilfælde vil den stige igen, uanset om interventionen med praktikforløbet gennemføres¹⁰. En tilgang, der går ud på at måle virkningen af et praktikforløb ud fra forskellen mellem indkomst umiddelbart før forløbet start og den efterfølgende indkomst, er upålidelig, fordi den ikke tager hensyn til, at indkomsten i mange tilfælde alligevel ville være steget.

For at tilpasse sådanne konstruktioner er det nødvendigt med et mål for det kontrafaktiske — dvs. et mål for, hvordan indkomsten ville have ændret sig for praktikanter uden interventionen. Et sådant kontrafaktisk resultat kan f.eks. opnås fra en omhyggeligt matchet kontrolgruppe, der ikke er eksponeret for interventionen, og hvis indkomst observeres på samme tidspunkt som praktikanternes indkomst. Derefter anvendes ofte antagelsen om en fælles tendens ("common trends assumption"), ifølge hvilken indkomstudviklingen for praktikanterne og kontrolgruppen ville have været den samme, hvis interventionen ikke havde fundet sted.

Begrænsningerne i disse "upålidelige" tilgange betyder, at det er relevant at se efter mere overbevisende evalueringsmetoder. Som anført ovenfor er mere overbevisende metoder imidlertid mere teknisk udfordrende at gennemføre. Næste afsnit i dette kapitel indeholder en kort beskrivelse af

¹⁰ Dette mønster kaldes populært "Ashenfelter Dip" efter den økonom, der beskrev det først. Se [Ashenfelter O., 1978](#).

nogle af de specifikke tilgange til kontrafaktisk konsekvensanalyse, som sandsynligvis vil være mest relevante i en ESF-sammenhæng.

1.4. En oversigt over forskellige designs og metoder til kontrafaktiske konsekvensanalyser

Når der fokuseres på virkningerne af en intervention på deltagerne, estimeres de kontrafaktiske resultater normalt ved hjælp af data indsamlet fra grupper af ikkedeltagere, som svarer til dem, der deltager i den intervention, der evalueres. Tabel 1 i slutningen af dette kapitel giver en kort oversigt over tilgange, nogle af deres fordele og begrænsninger og de vigtigste typer data, der skal bruges til dem.

Hovedsondringen i kontrafaktisk konsekvensanalyse er mellem eksperimentelle og kvasiekperimentelle evalueringdesigns. Den eksperimentelle tilgang betegnes almindeligvis som "kontrollerede randomiserede forsøg" eller "RCT" ("Randomised Controlled Trial") og undertiden også som "sociale eksperimenter".

Den eksperimentelle tilgang anses for at være den gyldne standard blandt metoderne til kontrafaktisk konsekvensanalyse til evaluering af virkningerne af interventioner, der kan testes og manipuleres over relativt korte tidshorisonter, og er i de fleste tilfælde den ideelle metode. Et godt konsekvensanalysedesign opnår skøn over kontrafaktiske resultater, der er uden bias. I mange anvendelser kan en eksperimentel tilgang anses for at give sådanne skøn uden bias. Når man drøfter tilgange til kontrafaktisk konsekvensanalyse, er det ofte ønskeligt først at skitsere den eksperimentelle tilgang. Dette skyldes, at kvasiekperimentelle metoder i det væsentlige søger at efterligne det eksperimentelle ideal.

Randomiseret design — den gyldne standard

I diskussioner om kontrafaktiske konsekvensanalyser design beskrives hver tilgang i hovedtræk så enkelt som muligt for at præcisere de underliggende principper. I virkeligheden kan anvendelsen af disse metoder være betydeligt mere kompleks, og problemer som manglende opfyldelse — dvs. når personer, der er blevet udvalgt til at deltage i interventionen, ikke deltog — kan bidrage væsentligt til de eksisterende udfordringer.

Der findes en lang række tilgange, som i det væsentlige søger at efterligne randomisering. Disse betegnes som **kvasiekperimentelle**. Det er ikke muligt at gennemgå dem alle inden for rammerne af denne vejledning eller at give en fuldstændig og detaljeret teknisk redegørelse for hver enkelt. De følgende kvasiekperimentelle metoder er imidlertid dem, der generelt med størst sandsynlighed vil blive anvendt inden for rammerne af ESF: 1. propensity score matching, 2. difference-in-differences, 3. regressionsdiskontinuitet, 4. instrumentvariabler. I Tabel 5 i kapitel 3 findes en oversigt over de vigtigste tilgange og deres relative fordele. De er fremstillet på en noget forenklet måde for at fremhæve de centrale principper i hver enkelt tilgang. Yderligere oplysninger om kvasiekperimentelle metoder findes i bilag 1.

Ikkerandomiserede eller kvasiekperimentelle designs

1.5. Hvordan kan kontrafaktisk konsekvensanalyse integreres i en bredere evalueringssammenhæng?

Kontrafaktiske evalueringer behandler visse typer spørgsmål om interventionernes årsagsvirkninger. Disse tilgange er begrænset i forhold til, i hvor høj grad de kan behandle andre spørgsmål vedrørende en intervention. Det er nyttigt at skelne mellem evalueringsspørgsmål vedrørende **årsagsforklaring** og spørgsmål vedrørende **årsagsbeskrivelse**. Kontrafaktiske konsekvensanalyser har til formål at beskrive konsekvenserne af en intervention. Sådanne metoder er mindre velegnede til at **forklare** de mekanismer og kontekster, gennem hvilke årsagssammenhænge opstår. Denne sondring er vigtig, da den bidrager til at afklare de kontrafaktiske konsekvensanalysers særlige rolle¹¹.

Årsagsforklaring og -beskrivelse

En veldesignet kontrafaktisk konsekvensanalyse vil informere den politiske beslutningstager om, hvorvidt en intervention har ført til ændringer i de resultater, den er designet til at påvirke. Den vil give dokumentation for omfanget af eventuelle virkninger eller eventuel effekt, vise den politiske beslutningstager, om virkningen var positiv eller negativ, men er også forbundet med en vis usikkerhed. Hvad de kontrafaktiske konsekvensanalyser gør mindre godt, er at give en forklaring på, hvorfor og hvordan de virkninger, der måles gennem den kontrafaktiske konsekvensanalyse, opstod. Omvendt er det ofte vanskeligt på grundlag af en kontrafaktisk konsekvensanalyse at afgøre, hvorfor en intervention ikke havde nogen virkning, hvis dette viser sig at være tilfældet.

Hvad kontrafaktiske konsekvensanalyser kan fortælle de politiske beslutningstagere, og hvad de ikke kan

Inden for de fleste politiske beslutningstagende organer har de interessenter, der stiller spørgsmål om årsagsbeskrivelse og årsagsforklaring, ofte forskellige interesser og perspektiver. Programledere og praktikere har tendens til at fokusere på spørgsmål om årsagsforklaring. De ansvarlige for at fordele ressourcerne og de ledende beslutningstagere med ansvar for budgetfastsættelsen har tendens til at fokusere på spørgsmål om årsagsbeskrivelse. I praksis kan sondringen mellem årsagsforklaring og årsagsbeskrivelse være udvisket. I nogle tilfælde kan kontrafaktiske konsekvensanalyser give en forklaring på, hvorfor der blev konstateret visse virkninger, f.eks. ved at undersøge interventionernes indvirkning på vigtige undergrupper. Det er imidlertid vigtigt nøje at overveje de typer spørgsmål, som interessenterne har vedrørende en intervention, og at vælge den rigtige tilgang til at få besvaret hvert enkelt. I de tilfælde, hvor det primære spørgsmål er, om en intervention virker, er en kontrafaktisk konsekvensanalyse under mange omstændigheder hensigtsmæssig. I de tilfælde, hvor det primære spørgsmål er, hvordan en intervention virker, er teoribaserede evalueringsmetoder og procesevalueringsmetoder mere velegnede.

Disse forskellige niveauer af spørgsmål og formål er sammenfattet i figuren nedenfor.

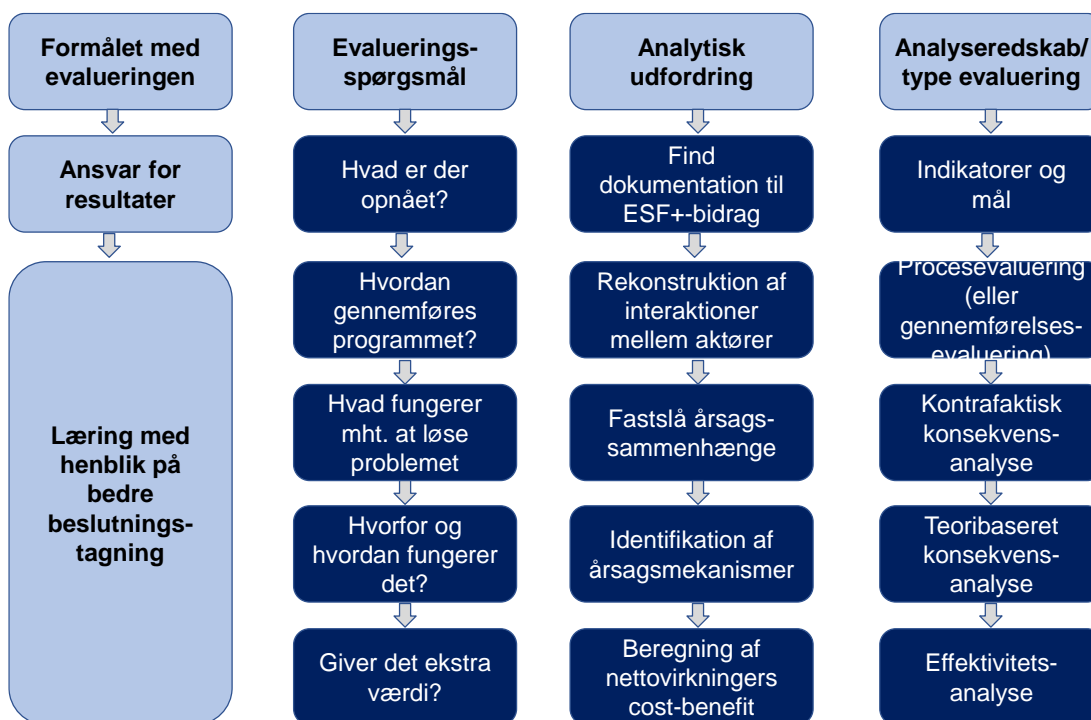
Efter denne diskussion kan det konkluderes, at kontrafaktiske konsekvensanalyser skal ligge inden for rammerne af evalueringsplanen. Denne evalueringsplan skal omfatte forskellige former for evaluering, der tager sigte på at besvare forskellige spørgsmål for forskellige politiske interessenter. I praksis vil en evalueringsplan sjældent, om nogensinde, indeholde en kontrafaktisk konsekvensanalyse uden en procesvurdering.

Der anvendes en bred vifte af tilgange til evaluering, og de tjener en række forskellige formål. Det afgørende spørgsmål er, hvordan disse tilgange kan

¹¹ Se Shadish W.R., Cook T.D. og Campbell D.T., 2002, og [Stern E. et al., 2012](#).

kombineres på nyttige måder for at fremme politisk læring. At kombinere forskellige typer evaluering rigtigt — med forskellige formål inden for programmeringscyklussen — er den reelle udfordring på dette område. Som det er blevet drøftet, og som det fremgår af den næste figur, supplerer kontrafaktisk konsekvensanalyse, procesevaluering og teoribaserede tilgange hinanden.

Figur 1 Forskellige opgaver og evalueringstyper



Kilde: Tilpasset fra [Martini A., 2009](#).

En solid evalueringsstrategi bør omfatte følgende elementer:

- teoribaseret konsekvensanalyse
- procesevaluering (eller gennemførelsesevaluering)
- kontrafaktisk konsekvensanalyse
- effektivitetsanalyse.

Kombinering af evalueringstyper

I denne vejledning vil vi kun se nærmere på kontrafaktiske tilgange til konsekvensanalyse. I forbindelse med kontrafaktisk konsekvensanalyse er teoribaserede tilgange et middel til at forstå den hensigt, der ligger til grund for designet af en intervention.

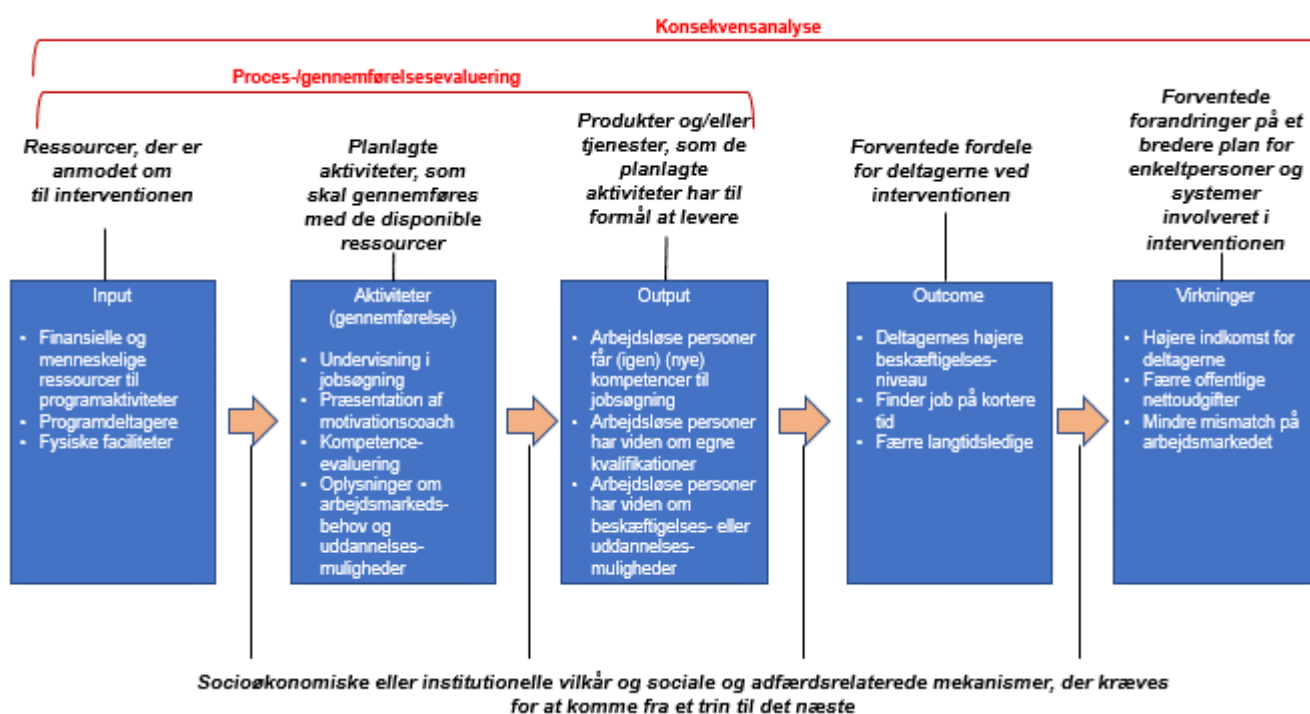
Teoribaserede evalueringer anvendes under visse omstændigheder til at beskrive den planlagte gennemførelse af interventionen og til at teste, om den ændring i resultaterne, der forudses i interventionsteorien, kan observeres. I den forstand kan de teoribaserede tilgange anvendes til at vurdere virkningen i forbindelse med besvarelsen af spørgsmålet om, "hvordan" en virkning er blevet skabt, og kan anvendes til at undersøge en interventions indvirkning, hvis det ikke er muligt at gennemføre kontrafaktiske konsekvensanalyser. Den næste figur viser et stiliseret eksempel på forandringsteori ("Theory of Change") for en intervention, der skal forbedre arbejdsformidlingstjenester. En teoribaseret evaluering undersøger, om dokumentationen bekræfter trinnene i årsagsrækkefølgen, fra input til virkning, som foreslået i

Teoribaseret evaluering henviser til en "forandringsteori"

forandringsteorien, og under hvilke betingelser og gennem hvilke sociale mekanismer dette var muligt. En detaljeret redegørelse for anvendelsen af teoribaserede tilgange til at bestemme virkningen ligger uden for denne vejledning.

I forbindelse med kontrafaktisk konsekvensanalyse vurderes i en teoribaseret evaluering den måde, hvorpå en intervention planlægges og designes, og hvordan den skal fungere. Tilgangen indebærer i det væsentlige samarbejde med interessenterne i en intervention om at udvikle en fælles redegørelse for en interventions underliggende "forandringsteori", som vises forenklet i den næste figur. Alle interventioner er udtryk for en programlogik, der knytter input og aktiviteter sammen med output, foreløbige resultater og derefter langsigtede resultater. Formuleringen af en forandringsteori er derfor også vigtigt input til kontrafaktiske konsekvensanalyser. Den letter rekonstruktionen af gennemførelsesprocessen og dens mulige indflydelse på output og outcomes samt identifikationen af de mest repræsentative outcomevariabler, der skal kontrolleres i den kontrafaktiske analyse.

Figur 2 Illustration af den logiske model eller "forandringsteori"



Kilde: Tilpasset fra [W.K. Kellogg Foundation, 2004](#) og [Bredgaard T., 2015](#).

Den teoribaserede evaluering kan kædes sammen med kontrafaktiske konsekvensanalyser på en række nyttige måder. En klart formuleret forandringsteori (eller interventionslogik) kan danne grundlag for designet af en kontrafaktisk konsekvensanalyse. Blandt andre aspekter kan en veldefineret forandringsteori fortælle designeren af en konsekvensanalyse følgende:

- Hvilke resultater er vigtige og kræver måling?
- Vil de sandsynlige virkninger af interventionen være negative eller positive, og hvor store vil de være?

... bidrager til den kontrafaktiske konsekvensanalyse

- Hvem er den tilsigtede målgruppe, og hvordan kan en kontrolgruppe udvælges?
- Hvor lang tid kan det tage, før programmets virkninger og resultater bliver til virkelighed?
- Hvilke data kan være nødvendige for at måle deltagelsen i interventionen?
- Hvor plausibel er kontrolgruppen som mål for den kontrafaktiske situation?

Udarbejdelse af en forandringsteori kan også bidrage til at opdage uventede virkninger, man kan tage højde for i designet af den kontrafaktiske konsekvensanalyse. Til en vis grad kan en klart formuleret forandringsteori også hjælpe evaluatoren med at fortolke resultaterne af en kontrafaktisk konsekvensanalyse. Med hensyn til fortolkning kan en **procesevaluering** imidlertid også være meget informativ.

Procesevaluering

Procesevaluering i forbindelse med kontrafaktiske konsekvensanalyser har to formål. Det første er at vurdere "fidelitet", mens det andet er at vurdere forskellen mellem behandlingsdeltagernes og kontroldeltagernes erfaringer.

Ved vurderingen af "fidelitet" undersøges det, i hvilket omfang en intervention blev gennemført i overensstemmelse med dens design. I en procesevaluering tages der hensyn til, hvilke tjenester der rent faktisk blev stillet til rådighed for deltagerne i interventionen. Svarer de til det, der var tilsigtet med teorien bag interventionen? Hvilken variation er der i leveringen fra sted til sted, hvis der observeres variationer? De fleste interventioner har både en styrings- og en effektlogik:

Vurdering af fidelitet

- Styringslogikken vedrører, hvordan gennemførelsesorganerne forventes at reagere på programmernes regler og incitamenter.
- Effektlogikken vedrører, hvordan de personer, der er mål for interventionen, forventes at reagere i betragtning af, hvad der rent faktisk leveres.

Fideliteten i procesanalysen afspejler, hvad der rent faktisk blev opnået ved en intervention, og hvad der dermed rent faktisk bidrager til de observerede virkninger. Den giver også vigtig feedback til projektledelsen.

Vurderingen af forskellen er særlig vigtig i forbindelse med den kontrafaktiske evaluering. Det er almindeligt, at der, som det er gjort for en stor del af denne vejledning, fokuseres på interventionens effekt. Men før indvirkning på resultaterne kommer indvirkning på input, nemlig forskellen i muligheder mellem behandlings- og kontrolgrupperne, som interventionen rent faktisk opnår. I princippet kan hver enkelt kontrafaktisk konsekvensanalyse "vendes på hovedet", og behandlingsgruppen anvendes som kontrol til at vurdere resultatet for personer i det, der før blev kaldt kontrolgruppen. Det betyder, at man skal vide lige så meget om, hvad kontrolgruppen har opnået som, hvad behandlingsgruppen har, fordi det er for forskellen

Forskel mellem behandlings- og kontrolgruppe

mellem input i behandling og kontrol, at den kontrafaktiske konsekvensanalyse viser forskelle i resultaterne.

Hvis vi igen vender tilbage til praktikforløbet, kan man forestille sig to helt forskellige udgangspunkter. På den ene side gennemføres praktikforløbet i en generel kontekst, hvor der ikke udbydes noget lignende. Personerne i kontrolgruppen klarer sig simpelthen uden. Men en anden mulighed er, at der findes alternativer. Der kan f.eks. være adgang til praktik fra virksomheder, der er specialiseret i erhvervsfaglig forberedelse. Hvis dette er tilfældet, skal procesanalysen så vidt muligt omfatte en vurdering af forskellen i deltagelsen i praktikforløb mellem behandlings- og kontrolgruppen, og ikke blot antage, at alle dimensioner af behandlingen ligger uden for kontrolgruppens rækkevidde.

Selv om procesevalueringer kan iværksættes helt uafhængigt af andre former for evaluering, gør deres betydning for både ledelsen og for den kontrafaktiske konsekvensanalyse det afgørende, at proces- og konsekvensanalysen planlægges sammen.

Kontrafaktisk konsekvensanalyse kræver en procesevaluering

En god procesanalyse kan bidrage til at sikre fidelitet, og procesevalueringer kan give en årsagsforklaring på en intervention. Uden en veludvalgt procesevaluering er det ofte vanskeligt at fortolke resultaterne fra en kontrafaktisk konsekvensanalyse fuldt ud eller at vurdere de omkostninger, der er nødvendige for at foretage en cost-benefit-analyse, når først der foreligger skøn over virkningerne.

Som nævnt ovenfor kan procesevaluering desuden bidrage til fortolkningen af resultaterne af konsekvensanalyser ved at vise, i hvilken sammenhæng en intervention er gennemført. Forståelse af sammenhængen er vigtig, fordi den angiver betingelserne for, at interventionen lykkes, og en forståelse af, i hvilket omfang den kan have lignende virkninger, hvis den gennemføres andre steder, inden for andre geografiske områder eller på andre tidspunkter. Dette er især vigtigt for at drøfte muligheden for at overføre politiske strategier og fremhæve god praksis inden for tværnational læring og udveksling. Procesanalyser bidrager til tilliden til, hvad der kaldes evalueringsresultaternes **eksterne validitet**.

I de fleste anvendelser omfatter effektivitetsanalyser enten en vurdering af omkostningseffektiviteten eller en fuldstændig cost-benefit-analyse.

Beregning af omkostningseffektivitetsgrad

Omkostningseffektivitetsanalysen omfatter en sammenligning af omkostningerne ved interventionen med dens effekt eller virkninger, som er blevet fastlagt i en kontrafaktisk konsekvensanalyse. Kort sagt beregnes omkostningseffektivitetsgraden ved at dividere virkningen af en intervention — udtrykt enten i måleenheder eller standardiserede enheder — med nettoomkostningerne ved at levere interventionen pr. behandlet enhed.

Omkostningseffektivitetsgraden for et praktikforløb, der har til formål at hjælpe arbejdsløse med at finde arbejde, kan afsløre, hvilke midler der skal bruges pr. deltager for at flytte den pågældende deltager fra arbejdsløshed til arbejdsmarkedet.

Cost-benefit-analyse til sammenligning af fordele med nettoomkostninger

Omkostningseffektivitetsgraden er et vigtigt mål for dem, der er ansvarlige for at fordele ressourcerne på tværs af programmerne. Nøgletal fra en række forskellige interventioner gør det muligt for de ansvarlige for ressourcefordelingen at foretage relative vurderinger af, hvilke interventioner der giver mest valuta for pengene.

I stedet for at udtrykke interventionens effekt i enten en måleenhed eller standardiserede enheder forsøger en **cost-benefit-analyse** at omsætte de skøn over virkningerne, der er opnået ved en kontrafaktisk konsekvensanalyse, til penge og sammenligne dem med en interventions nettoomkostninger. Formålet med cost-benefit-analysen er at fastslå, om de værdisatte fordele ved et program overstiger dets nettoomkostninger. En cost-benefit-analyse af et typisk ESF-uddannelsesprogram vil sammenligne interventionens fordele for deltagerne, regeringen og samfundet mere generelt med nettoomkostningerne ved interventionen. For deltagerne fastslås fordelene ved programmet (normalt forbedret beskæftigelsesegnethed og øget indtjening) ved hjælp af en kontrafaktisk konsekvensanalyse. Herfra fratrækkes værdien af den skat, der betales af deltagerne, og andre omkostninger i forbindelse med beskæftigelse for at nå frem til nettofordelen. Set fra regeringens synspunkt vil fordelene ved interventionen hidrøre fra yderligere skatteindtægter og reducerede velfærdssydelsers, mens staten vil bære størstedelen af omkostningerne ved interventionen. Omkostningerne for samfundet som helhed fremkommer ved at sammenlægge fordelene for deltagerne og det offentlige og trække summen af deltagerne og den offentlige sektors omkostninger fra disse.

Omkostningseffektivitetsanalyser og cost-benefit-analyser er stadig ikke særlig udbredte blandt ESF-evalueringerne. Disse analyser er imidlertid meget nyttige til at afgøre, om en intervention bør finansieres igen i fremtiden, eller til at identificere den mest effektive intervention blandt en række ensartede interventioner (se eksemplet i nedenstående tekstboks).

Konsekvensvurderinger fra en kontrafaktisk konsekvensanalyse er et centralt element i både omkostningseffektivitets- og cost-benefit-analyser. For førstnævnte er de et mål for effektiviteten, mens de for sidstnævnte udgør en vigtig kilde til estimering af de værdisatte fordele. Det er også klart, at både omkostningseffektivitetsundersøgelser og cost-benefit-analyser kræver, at der indsamles nøjagtige omkostningsdata, som kan danne grundlag for beregningen af nettoomkostninger. Sådanne aktiviteter betegnes normalt som en omkostningsanalyse. I nogle komplekse evalueringer med blandede metoder integreres omkostningsanalyser ofte i procesevalueringen, hvor researchinstrumenterne kan tilpasses med henblik på at indsamle vigtige omkostningsdata.

Tekstboks 1 Et eksempel på en kontrafaktisk konsekvensanalyse til sammenligning af fordele og omkostninger

Et eksempel på anvendelsen af en kontrafaktisk tilgang til at vurdere fordele og omkostninger findes i [Bazzoli M. et al., 2018](#). Undersøgelsen fokuserer på erhvervsuddannelsesprogrammer, der blev gennemført i den italienske selvstyrende provins Trento i 2010-2011, og som omfattede mere end 300 timers uddannelsesaktiviteter. To grupper af kurser blev evalueret: projekter finansieret af provinssmidler (PVR) og projekter finansieret af Den Europæiske Socialfond (ESF) med deltagelse af henholdsvis 954 og 205 deltagere.

De vigtigste trin i cost-benefit-analysen var følgende: 1) Forfatterne vurderede virkningen af kurserne på sandsynligheden for, at deltagerne fandt et job i løbet af de første tre år efter starten af deres kursus, ved hjælp af propensity score matching¹². 2) Kursernes indvirkning på bruttolønnen frem til udgangen af 2013 blev anslået. 3) Forfatterne anslog værdien af yderligere skatteindtægter (som følge af virkningerne på indkomsten) og de besparelser i de offentlige midler, der fulgte af faldet i antallet af modtagere af arbejdsløshedsunderstøttelse. 4) Fordelene og omkostningerne ved kurserne blev sammenlignet.

Der blev anvendt flere administrative datasæt: a) overvågningsdata vedrørende deltagerne og deres karakteristika, b) data fra de offentlige arbejdsformidlinger (Centri per l'impiego eller jobcentre) over arbejdsløse for at identificere kontrolgruppen, c) data fra COB-databasen, som er arkiveret over virksomhedernes obligatoriske anmeldelse af ansættelseskontrakter, der sendes til de offentlige arbejdsformidlinger, og som anvendes til at identificere beskæftigelsesstatus for behandlingsgrupper og kontrolgrupper, både før og efter deltagelsen i kurset, d) data fra fortegnelserne over skatteindtægter med henblik på at beregne den enkeltes indkomst og e) data fra det italienske socialsikringsinstitut (INPS) for oplysninger om arbejdsløshedsunderstøttelse, som de enkelte personer modtager.

Efter 36 måneder var sandsynligheden for at være i beskæftigelse i den behandlede gruppe i PVR-kurserne ca. 5 procentpoint højere end i kontrolgruppen, mens virkningen af ESF-kurserne var meget højere, omkring 28 procentpoint. I de tre år efter interventionen tjente personer, der deltog i PVR-kurserne, årligt i gennemsnit 2 250 EUR mere end kontrolgruppen, mens personer, der deltog i ESF-kurserne, tjente 4 106 EUR mere end kontrolgruppen. Dataene gjorde det også muligt for forfatterne at vurdere fordelene for de offentlige forvaltninger i form af øgede skatteindtægter og færre udgifter til sociale ydelser. Personer i PVR-kursusgruppen betalte 126 EUR mere i indkomstskat end kontrolgruppen for hvert af de tre år, mens skønnet for personer i ESF-kursusgruppen beløb sig til 318 EUR om året. Med hensyn til lavere arbejdsløshedsydelse betalt af det offentlige var virkningen af kurserne ubetydelig, sandsynligvis fordi ydelseernes pengemæssige værdi afhang af, hvor længe de havde været i beskæftigelse, inden de blev arbejdsløse, og mange deltagere var unge med ringe erhvervs erfaring.

Udgifterne til interventionerne beløb sig til 4 800 EUR pr. deltager i PVR-kurserne og 14 500 EUR i ESF-kurserne. Forfatterne sammenlignede også omkostninger og fordele på individuelt plan for begge typer kurser: Det blev konstateret, at omkostningerne i gennemsnit var større end fordelene i perioden 2010-2013¹³.

¹² Mere specifikt anvendte forfatterne blokering med regressionsjusteringsestimatoren.

¹³ En anden "lignende" øvelse findes i [Lammers M. and Kok L., 2021](#).

Kapitel 2. Praktiske overvejelser i udarbejdelsen af en kontrafaktisk konsekvensanalyse

I dette kapitel gennemgås de praktiske spørgsmål, der skal tages i betragtning i forbindelse med forberedelsen af en evaluering. Det bør anvendes ved planlægning af evalueringsaktiviteter, når det besluttes, hvilke interventioner der skal underkastes en kontrafaktisk konsekvensanalyse, og hvilke centrale spørgsmål der skal tages op ved udformningen af en kontrafaktisk konsekvensanalyse.

Udarbejdelse af en kontrafaktisk konsekvensanalyse

Udgangspunktet antages at være en situation, hvor en programleder i en forvaltningsmyndighed (eller en leder af et bemyndiget organ med ansvar for gennemførelsen af en ESF-intervention) overvejer, hvilke interventioner der skal evalueres, og hvilke strategier der er hensigtsmæssige for at udføre en kontrafaktisk konsekvensanalyse. Det antages også, at embedsmænd i en forvaltningsmyndighed ikke selv foretager evalueringer, men i stedet udliciterer eller bestiller evalueringstjenester fra eksterne eksperter. Selv om den kontrafaktiske konsekvensanalyse vil blive gennemført af en underleverandør, skal forvaltningsmyndigheden (eller det bemyndigede organ) planlægge og forberede konsekvensanalysen, inden den iværksætter den.

Evalueringsstrategien, herunder de forskellige typer evalueringer, der er beskrevet i det foregående kapitel, skal fastlægges i evalueringsplanen.

Evalueringsplaner er obligatoriske for alle programmer og skal godkendes af overvågningsudvalget senest et år efter programmets godkendelse¹⁴. Planerne skal udarbejdes i begyndelsen af programmeringsperioden og omfatte ordninger for evalueringsprocessen (styring af evalueringerne og forbindelsen mellem evaluering og overvågning), faktiske evalueringsaktiviteter (f.eks. en vejledende liste over evalueringer, der skal gennemføres, hver evaluerings omfang, de vigtigste spørgsmål, nødvendige data, potentiel anvendelse, vejledende tidsplan, ledelsesstruktur), tidsplan for evalueringerne, det samlede budget og opbygning af evalueringskapacitet.

Evalueringsplanerne er normalt af generel karakter, mens planlægning af en kontrafaktisk konsekvensanalyse kræver en mere detaljeret forberedelse. Ideelt set bør dette ske, når evalueringsplanen udarbejdes, og nogle detaljer kan blive tilføjet senere. Forvaltningsmyndigheden/det bemyndigede organ skal dog være opmærksom på, at det sjældent er let at skabe de forbindelser med interessenterne og få tilrettelagt de andre ting, der er nødvendige for interventionsrelateret dataindsamling, og at der er behov for planlægning i god tid.

Denne vejledning fokuserer på, hvordan man kan udvikle en evalueringsordning for specifikke interventioner, der er relevante for en

Udarbejdelse af en evalueringsordning

¹⁴ Evalueringsplaner er påkrævet i henhold til artikel 44, stk. 4 og 6, i Europa-Parlamentets og Rådets forordning (EU) 2021/1060. En plan kan omfatte de planlagte evalueringer af mere end ét program, men alle programmer skal indgå i en evalueringsplan.

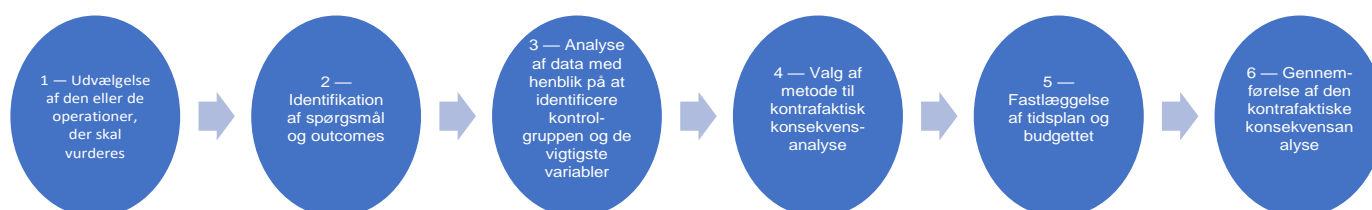
kontrafaktisk konsekvensanalyse. Denne ordning bør være en del af evalueringsplanen eller kan alternativt etableres som et operationelt skridt i forlængelse af en evalueringsplan.

*for specifikke
interventioner*

Ikke alle ESF-finansierede interventioner kan gøres til genstand for kontrafaktisk evaluering. De politiske beslutningstagere skal vælge, hvorhen de skal rette deres opmærksomhed. Det vil være nødvendigt at gennemføre en proces, hvor interventioner udvælges til konsekvensanalyse. I denne vejledning peges der på nogle aspekter, som forvaltningsmyndighederne skal tage hensyn til, når de udvælger passende interventioner. Desuden er hovedformålet med denne vejledning at hjælpe de ansvarlige for iværksættelse af kontrafaktiske konsekvensanalyser med at tænke over nogle af de udfordringer, der er forbundet med udarbejdelsen af en vellykket konsekvensanalyse, og i den forbindelse udarbejde evalueringsordninger for de forskellige kontrafaktiske konsekvensanalyser, som de overvejer.

I denne vejledning antages det, at forvaltningsmyndighederne efter udvælgelsen af interventionerne til kontrafaktisk konsekvensanalyse skal udarbejde en evalueringsordning for hver udvalgt intervention. Her anvendes udtrykket "ordning" til at skelne denne aktivitet fra de formelle "evalueringer", der kræves i henhold til forordning (EU) 2021/1060 om fælles bestemmelser for programmeringsperioden 2021-2027. Udtrykket "ordning" henviser navnlig til det sæt standardiserede aktiviteter, der er nødvendige for at definere og gennemføre en kontrafaktisk konsekvensanalyse, som skal udarbejdes, inden den lanceres, som vist i nedenstående figur.

Figur 3 Rækkefølge af de vigtigste trin i en kontrafaktisk konsekvensanalyse



Disse ordninger vil danne grundlag for iværksættelse af kontrafaktiske konsekvensanalyser og lægge grunden til, at leverandørerne kan foretage en grundig og brugbar undersøgelse. Resten af kapitlet omhandler de spørgsmål, der skal behandles i evalueringsplanlægningen. Evalueringsordningerne skal være skræddersyet til de særlige omstændigheder, der gælder for interventionen. Det er umuligt at sige noget om, hvad disse særlige omstændigheder vil være. Derfor diskuterer denne vejledning spørgsmål, som a) bør behandles i ordninger, eller b) bør kunne bidrage til overvejelserne om de udfordringer, som ordningerne skal tackle.

*Spørgsmål, der skal indgå
i
evalueringplanlægningen*

Efter at have gennemgået nogle af de spørgsmål, der skal overvejes ved vurderingen af, hvilke interventioner der kan gøres til genstand for en kontrafaktisk konsekvensanalyse, og om det endda er muligt at foretage en kontrafaktisk konsekvensanalyse med de tilgængelige typer data, er der nu fokus på nogle af de centrale spørgsmål, der skal tages i betragtning i forbindelse med udviklingen af en evalueringsordning. Dette skal gennemføres inden iværksættelse af en kontrafaktisk

konsekvensanalyse — eller en bredere evalueringsundersøgelse — for at kunne udarbejde et kommissorium og udpege en leverandør. Det vigtigste indhold i en sådan evalueringsordning er anført i nedenstående tabel.

Tabel 1 Anbefalet indhold i en evalueringsordning

De vigtigste trin i udarbejdelsen og gennemførelsen af en kontrafaktisk konsekvensanalyse	Indhold
1. Udvælgelse af den eller de operationer, der skal vurderes	<ul style="list-style-type: none"> – de ESF-operationer, eller typer af operationer (om muligt som led i en på forhånd fastsat typologi), der skal med den kontrafaktiske konsekvensanalyse – kort beskrivelse af de udvalgte operationers arbejdslogik (mål, de vigtigste udvælgelseskriterier og målgruppe, bistandstyper, gennemførelsesmodel, omtrentlige datoer for aktivering og afslutning, vejledende budget)
2. Identifikation af evalueringsspørgsmålene og outcomevariablerne	<ul style="list-style-type: none"> – evalueringsspørgsmål – operationernes forventede funktion ("forandringsteori") og outcomevariabler (f.eks. beskæftigelsesstatus, ændringer i indkomst, fattigdomsstatus, gennemsnitlig eksamenskarakter osv.)
3. Analyse af data med henblik på at identificere kontrolgruppen og måle outcomevariablerne	<ul style="list-style-type: none"> – administrative data (f.eks. arbejdsløshedsregistre, skatteregistre, administrative forsikringsdata, skoleregistre eller database over studerende osv.) eller andre data (f.eks. undersøgelser, big data), der skal anvendes - variabler i databasen eller undersøgelsen, der skal anvendes i den kontrafaktiske konsekvensanalyse - tidsserier for de nødvendige variabler - vigtigste regler og spørgsmål vedrørende dataadgang (direkte adgang for forvaltningsmyndigheden, behov for aftaler med andre forvaltninger, regler om beskyttelse af privatlivets fred og begrænsninger)
4. Valg af metode til kontrafaktisk konsekvensanalyse	<ul style="list-style-type: none"> – eventuel metode til kontrafaktisk konsekvensanalyse, der skal anvendes (dette kan præciseres senere, men anvendelsen af randomiserede kontrolforsøg eller andre metoder bør angives på forhånd for at gøre evalueringsprocessen mere konsekvent)
5. Fastlæggelse af tidsplan og budgettet	<ul style="list-style-type: none"> – tidsplan for kontrafaktiske konsekvensanalyser, vigtigste delmål: a) beslutning om eksperimentelle eller kvasiekperimentelle metoder, b) detaljerede evalueringsspørgsmål, c) udarbejdelse af kommissorium, d) identifikation af evaluator, e) forberedelse af data, f) dataanalyse, g) foreløbige og endelige rapporter, h) validering af resultaterne, i) formidling af resultater og erfaringer – det budget, der er til rådighed for den kontrafaktiske konsekvensanalyse
6. Gennemførelse af evalueringen	<ul style="list-style-type: none"> – udvælgelse af evaluator – tilsyn med gennemførelsen af den kontrafaktiske konsekvensanalyse – rapportering om den kontrafaktiske konsekvensanalyse – formidling af resultater (de vigtigste interessenter, der skal inddrages, vigtigste værktøjer)

Tekstboks 2 Kontrafaktisk konsekvensanalyse i en bredere evalueringsramme

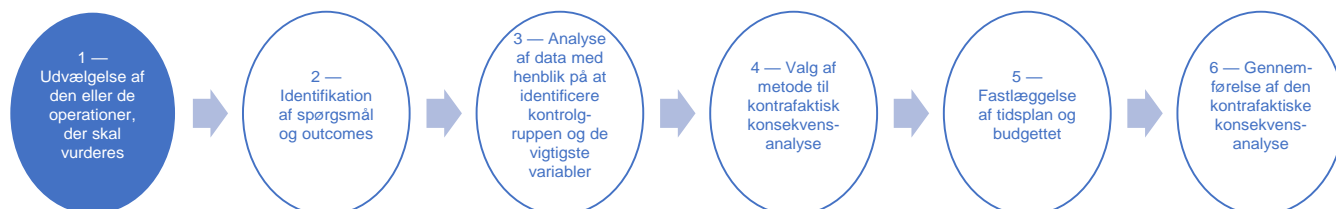
Mange kontrafaktiske konsekvensanalyser for ESF-finansierede interventioner, der gennemføres på tværs af medlemsstaterne, indgår i bredere evalueringsrammer:

- I Tyskland indgår den kontrafaktiske evaluering af programmet for integration af langtidsledige på arbejdsmarkedet finansieret under IP 9.i under OP Bund ESF 2014-2020 som led i en bredere evalueringsstrategi, der omfatter årlige foreløbige rapporter, som også behandler emner i forbindelse

med gennemførelsen af interventionerne. Fra 2017 til 2021 blev der udarbejdet fire rapporter¹⁵. Det samme gælder evalueringen af ESF-foranstaltninger til støtte for integration af langtidsledige i Baden-Württemberg, hvor den kontrafaktiske analyse fulgte en mere kvalitativ analyse med fokus på deltagernes vurdering af interventionerne¹⁶.

- I Piemonte (Italien) var den kontrafaktiske evaluering, der vurderede de beskæftigelsesmæssige virkninger af beskæftigelsesvouchers for sårbare personer finansieret under IP 8.i og 9.i, et skridt i en mere omfattende løbende evaluering. I to indledende rapporter blev der set nærmere på gennemførelsesspørgsmål og deltagernes subjektive opfattelse (april 2018 og februar 2019), mens yderligere to rapporter i juli 2019 og ved udgangen af 2020 fokuserede på beskæftigelsesmæssige virkninger ved hjælp af en kontrafaktisk tilgang¹⁷.
- I Marche (Italien) er den kontrafaktiske konsekvensanalyse, der undersøger virkningerne af interventioner for langtidsledige i 2020, en konsekvensanalyse med tematisk fokus efter en mere generel konsekvensanalyse i 2019 af ESF-operationer rettet mod arbejdsløse (Placement-rapporten)¹⁸.
- I Polen bestilte det operationelle program om viden, uddannelse og vækst 2014-2020 en række evalueringer (otte rapporter¹⁹) mellem udgangen af 2015 og maj 2020 med henblik på at analysere støtten fra ESF og UBI til unge fra forskellige perspektiver og under anvendelse af forskellige analysemetoder (f.eks. kvalitative analyser gennem undersøgelser og interviews, makromodeller, feltaktiviteter og en kontrafaktisk tilgang).

2.1. Udvælgelse af interventioner til konsekvensanalyse



Udvælgelsen af interventioner til konsekvensanalyse omfatter tre primære trin:

Kriterier for udvælgelse af interventioner

1. De strategiske spørgsmål skal fastlægges.
2. Når de strategiske prioriteter er klare, skal det vurderes, om de enkelte interventioner er i overensstemmelse med de grundlæggende krav i en kontrafaktisk tilgang, og i hvilket omfang de er innovative og/eller yder et væsentligt bidrag til videnbasen.
3. Tilgængeligheden eller den potentielle tilgængelighed af de typer data, der er nødvendige for at gennemføre en kontrafaktisk konsekvensanalyse, skal præciseres. Dette tredje spørgsmål har hidtil vist sig at være en

¹⁵ Se [Boockmann B. et al., 2017](#), [Boockmann B. et al., 2018](#), [Boockmann B. et al., 2019](#) og [Boockmann B. et al., 2021](#). I 2019 og 2021 blev der gennemført en kontrafaktisk konsekvensanalyse.

¹⁶ Se [Hunger K. og Sattler K., 2017](#) og [Scheller F. og Seidel K., 2020](#). I 2020 blev der gennemført en kontrafaktisk konsekvensanalyse.

¹⁷ Se [Pomatto G., 2017](#) — [Pomatto G., 2019](#) — [Poy S., 2019](#), [Poy S., 2020](#). I 2019 og 2020 blev der gennemført en kontrafaktisk konsekvensanalyse.

¹⁸ Se [Pompili M., Giorgetti I., 2020](#) — [Pompili M., Giorgetti I., 2020a](#)

¹⁹ Se [Instytut Badań Strukturalnych — Imapp — IQS, 2015](#), [Baran J. et al., 2016](#), [Baran J. et al., 2017](#), [Baran J. et al., 2018](#), [Baran J. et al., 2018a](#), [Palczyńska M. et al., 2019](#), [Kalinowski H., 2020](#), [Kalinowski H. et al., 2020](#). I 2017 og 2020 blev der gennemført én kontrafaktisk konsekvensanalyse.

væsentlig hindring for at foretage kontrafaktiske evalueringer af ESF-interventioner og kræver særlig opmærksomhed.

Tekstboks 3 Spørgsmål til udvælgelse af interventioner til en kontrafaktisk konsekvensanalyse

Kontrafaktiske konsekvensanalyser egner sig ikke til alle interventioner, og det er generelt ikke omkostningseffektivt at gennemføre en analyse for dem alle. Forvaltningsmyndighederne skal beslutte, hvordan ressourcerne skal fordeles for at opnå den største fordel. Evalueringsplanen forventes at afspejle disse valg, og ved planlægningen af kontrafaktiske konsekvensanalyser bør der tages hensyn til tre hovedelementer: a) strategiske prioriteter, b) gennemførligheden af en kontrafaktisk konsekvensanalyse og c) tilgængeligheden af de nødvendige data.

Evalueringsstrategien påvirkes af omfang, politikudvikling, uklare områder og behovet for viden. Forvaltningsmyndighederne bør stille følgende spørgsmål:

- Gør de store beløb, der afsættes til denne intervention, det særlig vigtigt at begrunde udgifterne? Disse interventioner er forholdsvis lette at identificere, fordi de modtager størstedelen af de midler, der er afsat til hvert specifikt mål (som defineret i artikel 2 i ESF+-forordningen 2021/1057).
- Gennemføres foranstaltningen som led i en reformproces, og vil resultaterne af evalueringen sandsynligvis bidrage til en kritisk gennemgang af indsatsen? Disse interventioner er knyttet til de seneste reformer af arbejdsmarkeds-, uddannelses- eller socialpolitikken. Selv om de måske ikke modtager store beløb, er de ikke desto mindre afgørende for, at reformen bliver en succes.
- Er interventionen innovativ og testes gennem et pilotprojekt eller et forsøg, inden den opskales? Disse interventioner modtager måske ikke enorme ressourcer, men kræver en tidlig vurdering for at afgøre, om de skal fortsætte og udvides, eller om forsøget skal afsluttes.
- Fokuserer interventionen på områder, der kræver yderligere dokumentation for effektivitet? Denne gruppe omfatter interventioner, som ikke tidligere er blevet evalueret, eller hvis seneste evalueringer er så gamle, at det er nødvendigt at ajourføre dem. I ESF+-programmerne er der mange af disse interventioner, fordi kontrafaktiske konsekvensanalyser først for nylig er blevet udbredte, og "nettoeffekten" er ukendt.

Gennemførlighed vedrører både interventionernes karakteristika og de omstændigheder, hvorunder de gennemføres. Planlæggerne bør kunne besvare følgende spørgsmål bekræftende:

- Er interventionsbehandlingen diskret, karakteristisk og tilstrækkeligt homogen?
- Er sammenligningen mellem behandlings- og kontrolgrupper egnet til at måle virkningen?
- Er målgruppen stor og veldefineret?
- Er den teori, der forbinder interventionen med de tilsigtede outcomes, logisk sammenhængende?
- Kan behandlingsgruppen klart identificeres inden for målgruppen?
- Er behandlingsgruppens størrelse tilstrækkelig?
- Kan der findes en troværdig kontrolgruppe?
- Kan forskellen mellem behandlings- og kontrolgruppernes erfaringer opretholdes over en tilstrækkelig lang periode til at måle virkningen?

Data er af afgørende betydning. Kernen i kontrafaktiske konsekvensanalyser er måling, og måling kræver kvantitative oplysninger om både behandlings- og kontrolgrupper og den sammenhæng, hvori evalueringen foretages. Nøjagtigt hvilke data der kræves, afgøres normalt af teorien om interventionen og den strategi, der anvendes til at fastslå det kontrafaktiske scenarie. Ved udvælgelsen af interventioner til kontrafaktisk konsekvensanalyse skal en forvaltningsmyndighed, der planlægger en kontrafaktisk konsekvensanalyse, spørge:

- Hvad er vigtigt at vide om mål- og kontrolgruppernes medlemmer?
- Hvad er det afgørende at vide om arten af den intervention, der faktisk er leveret til

behandlingsgruppen?

- Får kontrolgruppen ingen behandling eller en anden behandling? Findes der data herom?
- Hvilke data er tilgængelige fra administrative og andre kilder?
- Foreligger der data, der beskriver individuelle karriereforløb?
- Kan individualiserede data fra forskellige kilder sammenkædes?

Yderligere oplysninger om disse spørgsmål findes senere i dette kapitel.

2.1.1. Prioritering af interventioner til konsekvensanalyse

Inden man prioriterer specifikke interventioner med henblik på kontrafaktisk konsekvensanalyse, er der nogle bredere strategiske spørgsmål, der skal overvejes. Der bør fokuseres på at udvælge de interventioner, hvor konsekvensanalyser giver det største udbytte, når det drejer sig om at finde ud af, hvad der virker. Fordelene ved veldesignede og grundige evalueringer kommer ikke bare forvaltningsmyndigheden og de medlemsstater, der iværksætter dem, til gode, men også andre medlemsstater og deres forvaltningsmyndigheder, andre interessenter og Kommissionen.

Bidrag til begrundelse af udgifter

I betragtning af at de kontrafaktiske konsekvensanalyser fokuserer på spørgsmål, der er af afgørende betydning for de politiske beslutningstagere og navnlig dem, der har ansvaret for at fordele ressourcerne, giver det mening at koncentrere konsekvensanalyseindsatsen om programmer og interventioner, der er særligt ressourcekrævende. Jo mere tid og andre ressourcer et bestemt program eller en bestemt intervention kræver, jo vigtigere er det at forstå, om de virker, og derfor om fordelene overstiger de påløbne omkostninger. Dyre interventioner, der ikke skaber social eller økonomisk værdi, skal muligvis tages op til fornyet overvejelse, mens andre med dokumenteret merværdi kan kræve øget finansiering og opmærksomhed.

*Fokus på
ressourcekrævende
interventioner*

Resultaterne af nylige evalueringer af ESF-interventioner, der blev finansieret i programmeringsperioden 2014-2020, har vist, at der er behov for strategiske justeringer og øget fokus på centrale politiske mål. Interventionerne vedrørende beskæftigelse og arbejdskraftens mobilitet viste sig at være mindre effektive for ældre og dem, der er længere væk fra arbejdsmarkedet. Disse grupper kræver nye og mere effektive instrumenter²⁰. Selv om socialpolitiske interventioner udviser store forskelle i omkostninger pr. deltager og type operation, foretages der generelt ikke nogen systematisk cost-benefit-analyse²¹. Uddannelsesinterventionerne viste positive resultater, men det begrænsede antal konsekvensanalyser betyder, at det er vanskeligt at foretage en fuldstændig vurdering af deres langsigtede virkninger²². Kontrafaktiske konsekvensanalyser kan give mulighed for at vurdere interventioner med henblik på at identificere de mest effektive tilgange for bestemte målgrupper og derved maksimere "valuta for pengene" i de nye ESF+-programmer.

²⁰ [Fondazione G. Brodolini, Metis GmbH, Applica, Ockham IPS \(2020\).](#)

²¹ [ICF, Cambridge Econometrics and Eurocentre \(2020\).](#)

²² [Ecorys, Ismeri Europa \(2020\).](#)

Bidrag fra en intervention til en reformproces

Interventioner, der udgør et centralt element i et bredere reformprogram, vil ofte tiltrække betydelige midler. Det forhold, at en ESF-intervention er af central betydning for en social integrationsstrategi eller en vigtig del af et aktivt arbejdsmarkedsprogram, vil naturligvis betyde, at den kræver større opmærksomhed.

Interventioner, der bidrager til politisk innovation

Innovativ og sonderende

Nye og innovative pilotinterventioner er åbenlyse kandidater til en kontrafaktisk konsekvensanalyse. For at teste virkningerne af interventioner gennem et pilotprojekt eller et forsøg er det klart nødvendigt at foretage en grundig evaluering. Evaluering ved hjælp af en veldesignet kontrafaktisk konsekvensanalyse er så meget desto vigtigere, når der er en klar vilje til at opskalere eller rulle indsatsen mere bredt ud, hvis den opfattes som vellykket.

Bidrag til læring

Behovet for at koncentrere opmærksomheden og ressourcerne om specifikke programområder — og specifikke interventioner på disse områder — øges, når der kun er begrænset eller slet ingen dokumentation for, hvad der virker inden for det pågældende politikområde. Det gælder f.eks., når der er reel usikkerhed om politikken fremadrettet, og der er risiko for overdreven tillid til dokumentation, som måske ikke er direkte relevant (f.eks. dokumentation fra andre lande).

Udarbejdelse af dokumentation

Evalueringer af høj kvalitet kan betragtes som et offentligt gode. De fordele, de skaber i form af læring, omfatter også interessenter ud over dem, der arbejder for en specifik forvaltningsmyndighed. Det er derfor vigtigt at overveje, hvilke interessenter der kan drage fordel af den foreslåede konsekvensanalyse. Det kan være bemyndigede organer eller agenturer, der beskæftiger sig med interventioner inden for samme program, andre forvaltningsmyndigheder eller bemyndigede organer i den pågældende medlemsstat eller agenturer og institutioner, der beskæftiger sig med nationale eller regionale fonde. En anden indlysende ekstern interessent, der bør overvejes, er Kommissionen, og der er også interessenter i andre medlemsstater, som måske kan lære af en evaluering. Hensyntagen til behovene hos andre end de umiddelbare interessenter er et vigtigt bidrag, som de politiske beslutningstagere og programforvalterne kan yde til gensidig læring.

En sidste strategisk overvejelse ved udvælgelsen af fokusområder i forbindelse med udarbejdelsen af kontrafaktiske konsekvensanalyser, er de interventioner, der kan vise fordelene ved kontrafaktiske konsekvensanalyser og fungere som model.

Vis fordelene ved kontrafaktiske konsekvensanalyser frem

2.1.2. Udvalgelse af interventioner, der egner sig til en kontrafaktisk tilgang

Efter en gennemgang af de bredere strategiske hensyn, der kan ligge til grund for udvælgelsen af bestemte interventioner til kontrafaktisk konsekvensanalyse, ser dette afsnit nærmere på de særlige karakteristika ved interventioner, der kan gøre dem egnede til en kontrafaktisk tilgang. Disse karakteristika er mange og varierede. Nogle af elementerne i en intervention kan være egnet til en kontrafaktisk konsekvensanalyse i visse situationer, mens de i andre kan stå i vejen for dens gennemførelse. Som følge heraf er det ikke muligt at lave en udtømmende liste over de

overvejelser, der skal gøres. Det er imidlertid værd at nævne de karakteristika ved interventioner, der med størst sandsynlighed vil sikre en vellykket kontrafaktisk konsekvensanalyse.

Tekstboks 4 De mest almindelige interventionstyper og målgrupper til kontrafaktiske konsekvensanalyser for ESF

Størstedelen af de kontrafaktiske konsekvensanalyser, der foretages for ESF-interventioner, fokuserer på aktive arbejdsmarkedspolitikker rettet mod arbejdsløse og undergrupper blandt dem, der er berørt af en særlig ulempe. Det store antal konsekvensanalyser af interventioner for unge afspejler, at reglerne om ungdomsbeskæftigelsesinitiativet kræver, at sådanne evalueringer foretages med bestemte mellemrum.

Omkring halvdelen af de kontrafaktiske konsekvensanalyser, der er udpeget i Evaluation Helpdesk-projektet²³ siden 2015, vedrører tematisk mål 8 "fremme af bæredygtig beskæftigelse af høj kvalitet og støtte til arbejdskraftens mobilitet". I de fleste kontrafaktiske konsekvensanalyser, der undersøgte interventioner finansieret under tematisk mål 9, er der ligeledes fokus på interventionernes effektivitet med hensyn til at integrere sårbare arbejdsløse på arbejdsmarkedet (f.eks. den tyske intervention rettet mod langtidsledige, der finansieres under OP Bund ESF 2014-2020).

De hyppigst analyserede former for støtte til arbejdsløse er uddannelse, praktikophold eller andre former for erhvervs erfaring og jobtilskud. Dette fremgår klart af de eksempler, der analyseres i denne vejledning (se eksemplerne fra den italienske evaluering i Marche, evalueringen af kurser for migranter i Tyskland og erhvervsuddannelse for NEET'er i Letland). Analyse af rådgivnings- og jobmatchningstjenester er mindre almindelig (som det er tilfældet med det svenske eksempel i denne vejledning, hvor en pilotaktion med henblik på at øge støtten til arbejdsløse til offentlige arbejdsformidlinger evalueres ved hjælp af en randomiseret tilgang). Ifølge Helpdesk-data evalueres ESF-interventioner til støtte for selvstændig erhvervs virksomhed eller virksomhedsetablering ofte ikke²⁴.

I nogle tilfælde analyserer de kontrafaktiske konsekvensanalyser forskellige interventionstyper samlet, hvilket indebærer en risiko for at blande forskellige interventionslogikker og mindske pålideligheden af sammenligningen mellem behandlings- og kontrolgrupperne.

Arbejdet med at evaluere virkningerne af interventionerne for sårbare personer er særligt interessant, ikke kun med hensyn til beskæftigelsesoutcomes, men også med hensyn til "bløde outcomes": En kontrafaktisk konsekvensanalyse, der blev gennemført i Tyskland, fokuserede på dette aspekt ved at vurdere virkningen af jobskabelsesordninger for langtidsledige på selvvurderet sundhedstilstand, tilfredshed med livet, følelsen af at høre til og social status.

ESF-interventioner på uddannelsesområdet vurderes mindre hyppigt ved hjælp af en kontrafaktisk tilgang. Der er færre kontrafaktiske konsekvensanalyser vedrørende tematisk mål 10 end tematisk mål 8 eller 9, og der fokuseres ofte på interventioner finansieret under IP 10.IV vedrørende erhvervsuddannelse og dens virkninger med hensyn til integration på arbejdsmarkedet. To årsager hertil er mere begrænset adgang til datasættene og strengere regler for beskyttelse af privatlivets fred²⁵. Der er dog gjort forsøg i denne retning: I Spanien (Asturien) blev der foretaget en vurdering af en indsats på sekundærtrinnet for at modvirke skolefrafald. I Polen (Podalskie) undersøgte de kontrafaktiske konsekvensanalyser virkningerne af et projekt, der havde til formål at fremme erhvervsuddannelse blandt unge studerende. I Portugal blev stipendier til studerende på videregående uddannelser analyseret.

Kontrafaktiske konsekvensanalyser kan gennemføres på forskellige politiske niveauer (dvs. en eller flere prioriterede akser, specifikke mål eller operationer²⁶ i et program) og kan omfatte nationale eller regionale ESF+-programmer, kan fokusere på homogene målgrupper (mænd eller kvinder, unge eller sårbare befolkningsgrupper, langtidsledige osv.) eller

²³ Se https://ec.europa.eu/regional_policy/en/policy/evaluations/member-states/.

²⁴ Eksempler herpå er: [Borik V. et al., 2015](#) [Ires Piemonte, 2019](#), [Openfield, 2019](#).

²⁵ Se eksemplet [Ismeri Europa — Ecorys — Institute for Employment Studies, 2019](#).

²⁶ Artikel 2, nr. 4), i Rådets forordning (EU) 2021/1060 indeholder følgende definition: "operation": a) et projekt, en kontrakt, en aktion eller en gruppe projekter, som er udvalgt under de pågældende programmer, b) i forbindelse med finansielle instrumenter, et programbidrag til et finansielt instrument og den efterfølgende finansielle støtte, som det finansielle instrument yder til slutmodtagere.

interventionstyper (f.eks. uddannelse, tjenester til social inklusion eller modvirkning af skolefrafald (se foregående tekstboks)).

Eksempler fra medlemsstaterne viser, at en række instrumenter, der anvendes inden for ESF+, er egnede til en kontrafaktisk konsekvensanalyse, herunder uddannelse, beskæftigelsesincitamenter og arbejdsmarkedstjenester (f.eks. jobrådgivning, coaching). På den anden side er jobrotation og jobdelingsforanstaltninger, opstartsincitamenter eller støtte til systemer og strukturer samt interventioner inden for uddannelse og social inklusion en større udfordring.

Det er hensigtsmæssigt at overveje, hvilke interventioner der er mere egnede til en kontrafaktisk konsekvensanalyse, ved at tage stilling til følgende spørgsmål:

Er interventionen diskret, karakteristisk og forholdsvis homogen?

Den eller de behandlinger, der er omfattet af en intervention, skal kunne skelnes fra dem i andre interventioner. Desuden skal der være en meningsfuld kontrast mellem det, som deltagerne i en intervention modtager, og hvad andre lignende persongrupper drager fordel af. Hvis behandlinger er uklare, således at det ikke er muligt at identificere en særskilt gruppe af modtagere, er kontrafaktiske tilgange ikke mulige eller ønskelige.

Behandlingen skiller sig tydeligt ud

Det bliver meget komplekst og vanskeligt at udføre en kontrafaktisk konsekvensanalyse, hvis en given enheds (en virksomheds eller persons) behandlingsstatus påvirker andre enheders potentielle resultat (gennem såkaldte bredere "generelle ligevægtsvirkninger"). I praktikforløb kan dette ske, når kandidater fra programmet gør det vanskeligt for andre ikkepraktikanter at finde arbejde på kort sigt. Hvis dette anses for at være et væsentligt problem (f.eks. i forbindelse med omfattende interventioner), kan det være nødvendigt med makroøkonomiske analyser for at vurdere omfanget af substitutions- og forskydningseffekter. Forvaltningsmyndighederne bør indhente ekspertrådgivning, når det er sandsynligt, at sådanne effekter vil være til stede.

Selve interventionen bør være forholdsvis homogen. Det betyder, at alle deltagere i en intervention bør modtage eller blive eksponeret for stort set den samme pakke af foranstaltninger. Der er en række konsekvenser for den kontrafaktiske konsekvensanalyse, hvis rækken af foranstaltninger, der leveres til deltagerne i en enkelt intervention, er for forskelligartet. For det første giver det måske ikke mening at tale om en sammenhængende intervention, men snarere om interventioner med separate årsagsprocesser. For det andet vil det være vanskeligt at fortolke virkninger, der indberettes som gennemsnitlige nettovirkninger for en gruppe af forskellige interventioner. For det tredje kan undergruppeanalyse være relevant, men hvis der er for mange undergrupper inden for en behandlingsgruppe, kan begrænsninger i stikprøvestørrelsen begrænse muligheden for at indberette brugbare resultater.

Homogene interventioner

Sammenlignes behandlingen med ingen behandling, eller findes der andre relevante former for behandling?

ESF medfinansierer nationale og regionale arbejdsmarkedspolitikker og politikker for social inklusion. Enhver evalueringsordning med kontrafaktisk konsekvensanalyse skal derfor nøje tage hensyn til, om interventionen er klart identificerbar, og enkeltpersoner har mulighed for at modtage tjenester

fra andre (nationale eller regionale) programmer og finansieringskilder. Det vigtige er, at de behandlinger, der evalueres, rent faktisk ændrer deltagernes muligheder eller ressourcer i forhold til, hvad der er til rådighed for kontrolgruppen, og at forskelle kan måles og overvåges.

Sådanne problemer med "kompleks behandling" har tendens til at være kontekstspecifikke. De komplicerer designet og gennemførelsen af kontrafaktiske konsekvensanalyser. Hvis disse problemer eksisterer, understreger det betydningen af omhyggelig evalueringsplanlægning — udarbejdelse af evalueringsordningen — forud for gennemførelsen.

*Komplekse
behandlinger*

Er der en stor og veldefineret målgruppe?

Kontrafaktiske konsekvensanalyser kræver store stikprøvestørrelser i forhold til andre former for evaluering. Målgrupper, der består af et tilstrækkeligt antal personer, er af afgørende betydning, og det skal være muligt at finde kontrolgrupper af en tilstrækkelig størrelse. Dette spørgsmål behandles nærmere nedenfor.

*Stor
stikprøvestørrelse*

Det er vigtigt, at den intervention, der overvejes til kontrafaktisk konsekvensanalyse, er rettet mod en veldefineret gruppe. Uden en klar forståelse af, hvem målgrupperne er, er det vanskeligt at identificere en meningsfuld kontrolgruppe. Nogle interventioner har bevidst til formål at rekruttere personer til behandling gennem uformelle mekanismer, der omfatter processer, der ikke er foruddefinerede eller for præskriptive (f.eks. kan vanskelige sociale målgrupper, såsom unge NEET'er eller dårligt stillede grupper, inddrages gennem lejlighedsvis procedurer og procedurer, der skifter fra person til person). Dette kan gøre det vanskeligt præcist at identificere de behandlede personer og den tilknyttede kontrolgruppe.

*Fastlæggelse af
målgruppens
identitet*

Findes der en klar årsagsmekanisme?

Som nævnt ovenfor, da de vigtigste evalueringsmetoder og "forandringsteorien" blev præsenteret, er det ofte hensigtsmæssigt, at en teoribaseret evaluering er blevet gennemført forud for eller i kombination med en kontrafaktisk konsekvensanalyse. Udvikling af en forandringsteori eller en detaljeret interventionslogik kan hjælpe ved design af en kontrafaktisk konsekvensanalyse på en række forskellige måder. Det gælder især, når det skal afgøres, om en intervention har en sammenhængende årsagsmekanisme, der understøtter den. Det er usandsynligt, at interventioner uden en klar og overbevisende årsagsmekanisme vil få virkninger af et tilstrækkeligt omfang til statistisk at kunne identificere dem ved hjælp af en kontrafaktisk konsekvensanalyse.

*Særskilt politisk
mekanisme*

Kan outcomes defineres kvantitativt?

Der er behov for kvantificerbare mål for outcomes (eller resultater). Sådanne data og indikatorer kan indhentes fra administrative kilder eller specifikt målrettede undersøgelser.

*Nødvendigt at måle
resultater*

Under visse omstændigheder kan hensigten med interventioner være at opnå resultater, der kræver særlige forholdsregler for, at de kan måles kvantitativt. En intervention kan f.eks. vedrøre ændring af holdninger, overbevisninger eller meninger. I sådanne tilfælde skal der foretages undersøgelser for at måle disse ændringer. Nogle interventioner har ret vage eller dårligt definerede resultater. Også her kan udviklingen af en interventionslogik bidrage til at styrke forståelsen af, hvad en intervention

søger at opnå, og hvordan det er hensigten, at den skal opnå de ønskede resultater.

Er interventionen organiseret på en måde, der gør det muligt at finde en meningsfuld kontrolgruppe?

For at identificere en meningsfuld kontrolgruppe er det vigtigt at overveje, hvordan behandlede enheder (personer eller virksomheder) udvælges til en intervention, eller hvorfor de beslutter at deltage, om samme datakilde — f.eks. det samme undersøgelsesinstrument — kan anvendes på kontrolstikprøven som på behandlingsgruppen, og endelig, om det er nødvendigt at udtage stikprøver af kontrolgruppen, som er underlagt de samme arbejdsmarkedsbetingelser som behandlingsgruppen. Nedenstående tekstboks giver nogle eksempler.

Mekanisme for udvælgelse til behandling

Hvis en intervention er obligatorisk og leveres til hele målpopulationen mere eller mindre samtidigt, kan det vise sig vanskeligt at finde en ubehandlet del af målpopulationen, der kan fungere som kontrolgruppe.

Tekstboks 5 Definition af kontrolgrupper

I de eksempler på kontrafaktiske konsekvensanalyser, der indgår i denne vejledning, var udvælgelsen af kontrolgrupper baseret på interventionernes karakteristika (f.eks. udvælgelseskriterier) og tilgængeligheden af relevante data.

Sammenlignet med erfaringerne fra kontrafaktiske konsekvensanalyser i den foregående programmeringsperiode er identifikationen og udvælgelsen af kontrolgruppen oftere baseret på administrative data, navnlig arbejdsløshedsregistre. Den mest almindelige strategi er at identificere potentielle kontrolpersoner med lignende karakteristika, der er registreret som arbejdsløse hos offentlige arbejdsformidlinger i en vis periode, som krævet i henhold til udvælgelseskriterierne. Dette gælder for nogle evalueringer i Italien (Marche, provinsen Trento og Piemonte), Polen (Lubelskie og Podlaskie) samt de tyske evalueringer af uddannelseskurser for migranter, jobskabelsesordninger og integrationsforanstaltninger for arbejdsløse i Baden-Württemberg.

I andre evalueringer, som ikke indgår blandt de eksempler, der er beskrevet i denne vejledning, blev personer, der havde tilmeldt sig, men som ikke blev udvalgt til behandling, efterfølgende udvalgt til kontrolgruppen. Dette er tilfældet i den italienske evaluering af ungdomsgarantien, den tyske evaluering af pilotprojektet om borgerarbejde, som blev finansieret i 2007-2013, og evalueringen af PIPOL-programmet, der blev gennemført i Friuli-Venezia Giulia²⁷. Denne strategi anvendes imidlertid ikke hyppigt, da overvågningsinformationssystemerne i de fleste lande og regioner ikke indeholder oplysninger om personer, der har ansøgt, men ikke deltaget.

Den tyske tilgang til evaluering af interventioner for langtidsledige²⁸ var anderledes, da der i evalueringen blev foretaget en vurdering af "behandlingsintentionen" og ikke af "gennemsnitlige behandlingseffekter på behandlede personer". I dette tilfælde bestod den behandlede gruppe af personer, der potentielt var berettiget til behandling i hele programmets gennemførelsesperiode, uanset om de rent faktisk blev behandlet (fra 2015), og kontrolgruppen bestod af personer, som opfyldte samme udvælgelseskriterier, men som var langtidsledige forud for gennemførelsen af programmet (2010-2012).

Kun i det svenske eksempel, som anvendte en randomiseret tilgang, blev kontrolgruppen identificeret tilfældigt: I designet blev behandlingen (intensiveret støtte fra offentlige arbejdsformidlinger) ydet tilfældigt til en gruppe unge, mens kontrolgruppen modtog den almindelige støtte, som de offentlige arbejdsformidlinger tilbød.

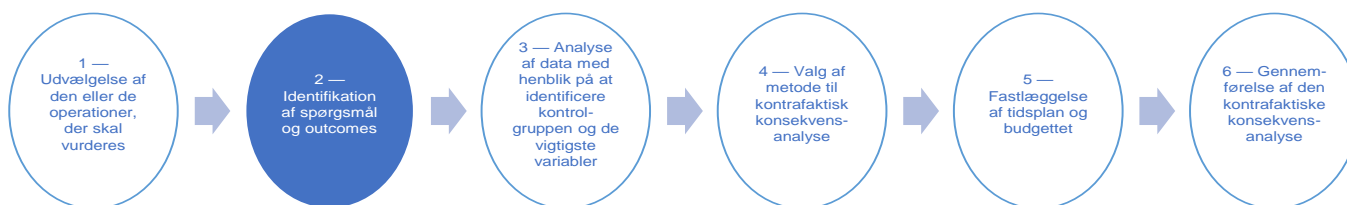
I de få evalueringer, der fokuserer på ESF-foranstaltninger for virksomheder, blev afgrænsningen mellem behandlings- og kontrolgrupper fastlagt mellem finansierede og ikkefinansierede ansøgere som i nogle

²⁷ Se henholdsvis [Isfol, 2016](#) — [IAW Institut für Angewandte Wirtschaftsforschung](#) — [ISG Institut für Sozialforschung und Gesellschaftspolitik, GmbH, 2015](#) — [Ismeri Europa, 2018](#).

²⁸ Se [Boockmann B. et al., 2019](#).

danske evalueringer, hvor resultaterne for de virksomheder, der finansieres gennem ESF, sammenlignes med resultaterne for en stikprøve af virksomheder med lignende egenskaber, som ikke modtog støtte²⁹.

2.2. Evalueringsspørgsmål og outcomevariabler



2.2.1. Hvad er interventionens mål og formål?

I forbindelse med udarbejdelsen af en evalueringsordning er det først og fremmest tilrådeligt at beskrive formålene med og de vigtigste elementer af interventionen.

Kombination af design med viden fra interventionslogikken

I mange tilfælde vil der allerede foreligge dokumenter, der beskriver formålene med interventionen. I forbindelse med en kontrafaktisk konsekvensanalyse er det imidlertid vigtigt at være specifik om de resultater og ændringer, som interventionen skal opnå, og dermed de forventede virkninger.

Det er ofte gavnligt at formulere en foranstaltningsteori, som beskriver, hvordan dens forskellige input og aktiviteter skal knyttes til output, outcomes (eller resultater) og dermed virkninger.

2.2.2. Hvad er formålet med evalueringen?

Ved udarbejdelsen af en evalueringsordning for en kontrafaktisk konsekvensanalyse er det vigtigt nøje at overveje formålet med evalueringen. Uden en klar forståelse af, hvorfor evalueringen er nødvendig, er det usandsynligt, at den vil give den nødvendige dokumentation. I forbindelse med evalueringer af ESF-finansierede interventioner er det nødvendigt at overveje en række spørgsmål:

- Hvad er formålet med og arten af evalueringen inden for rammerne af Kommissionens forskriftsmæssige krav og retningslinjer?
- Hvem er evalueringens vigtigste interessenter?
- Hvordan vil evalueringens resultater blive anvendt?
- Hvilke specifikke spørgsmål skal evalueringen besvare?

Hvad er formålet med og arten af evalueringen?

For det første skal **begrundelsen** for at gennemføre evalueringen defineres. I forordning (EU) 2021/1060 hedder det følgende: "Medlemsstaten eller forvaltningsmyndigheden foretager evalueringer af programmerne vedrørende et eller flere af følgende kriterier: effektivitet, produktivitet, relevans, sammenhæng og merværdi på EU-plan med det formål at forbedre kvaliteten af udformningen og gennemførelsen af

Evalueringens formål og art

²⁹ F.eks.: [Danmarks Statistik et al., 2017](#) og [Danmarks Statistik et al., 2018](#).

programmerne. Evalueringerne kan også omfatte andre relevante kriterier, såsom inklusion, ikkeforskelsbehandling og synlighed, samt mere end ét program³⁰. Som beskrevet ovenfor, vedrører resultaterne af en kontrafaktisk konsekvensanalyse effektivitet (i hvilket omfang de forventede outcomes er opnået) og produktivitet (interventionens omkostningseffektivitet eller cost-benefit-forhold).

Mere generelt opfordrer Kommissionen medlemsstaterne til at følge resultatorienterede tilgange i deres politikudformning og til at foretage evalueringer, der opfylder medlemsstaternes interne behov med hensyn til deres anvendelsesområde, udformning og tidsramme. I den forbindelse kan kontrafaktiske konsekvensanalyser i ESF+-programmerne også gennemføres for at besvare specifikke evalueringsspørgsmål eller i overensstemmelse med den nationale evalueringspolitik.

For det andet skal **evalueringens art** fastlægges³¹:

- Ved evalueringer af **konsekvenserne** undersøges virkningerne af et program eller en gruppe af programmer i forhold til EU's og medlemsstaternes prioriteter (dette kan være ESF's makroøkonomiske virkninger, fokus på specifikke politikker og temaer eller horisontale prioriteter som f.eks. barndom og lige muligheder).
- **Procesrelaterede (eller gennemførelsesrelaterede)** evalueringer understøtter gennemførelsen af et program, analyserer fremskridt og gennemførelsesmetoder og fremsætter anbefalinger om forbedringer af programmet.

I princippet kan den kontrafaktiske tilgang anvendes på konsekvensevurderinger, mens procesevaluering kræver andre metoder (se også Figur 1 ovenfor).

Forordningen om fælles bestemmelser kræver ikke et specifikt antal konsekvensanalyser i programmeringsperioden 2021-2027 i modsætning til 2014-2020. I stedet kræver den en evalueringstrategi, der kan vurdere, hvordan støtte fra EU-fonde har bidraget til at nå programmets mål med hensyn til alle de vigtigste strategiske profiler³². Desuden bestemmer forordningen om fælles bestemmelser ikke, hvilke prioriteter eller interventioner der skal fokuseres på i evalueringen, men overlader det til forvaltningsmyndigheden og evalueringsplanen. Det betyder, at evalueringstrategien for hvert enkelt program skal fastlægge sammensætningen af konsekvens- og procesevalueringer, og hvilke prioriteter og interventioner der skal fokuseres mest på. Der skal udarbejdes en endelig og generel konsekvensanalyse senest i juni 2029³³, men der er ingen andre tidsmæssige begrænsninger for de andre evalueringer, som skal fastsættes i evalueringsplanen.

Hvem er hovedmålgruppen?

Det bør defineres, hvem der er målgruppe for evalueringen. Afhængigt af evalueringens art kan den omfatte politiske beslutningstagere, forvaltningsmyndigheder og programforvaltere, andre forvaltningsmyndigheder eller gennemførelsesorganer i medlemsstaten

Hvem er interessenterne?

³⁰ Artikel 44, stk. 1, i forordning (EU) 2021/1060.

³¹ [Se Europa-Kommissionen, 2007.](#)

³² Artikel 44, stk. 1, i forordningen om fælles bestemmelser (EU) 2021/1060.

³³ Artikel 44, stk. 2, i forordningen om fælles bestemmelser (EU) 2021/1060.

eller nationale eller regionale myndigheder, der forvalter lignende programmer. Hvis data leveres af institutioner uden for programforvaltningen, bør disse organer ("dataejere") også betragtes som interessenter. Det er vigtigt at inddrage alle vigtige interessenter i en evalueringstyringsgruppe for at sikre fælles ejerskab af processen med udformning og gennemførelse af evalueringen samt visse evalueringsekspertter fra akademiske eller offentlige institutioner, som kan yde teknisk rådgivning til forvaltningsmyndigheden.

Hvordan vil evalueringresultaterne blive anvendt?

Når målgruppen for evalueringen er blevet identificeret, kan anvendelsen af resultaterne bestemmes. Dette kan i praksis opnås ved at inddrage styringsgruppen i udviklingen af evalueringsspørgsmålene og drøftelserne om kommissoriet.

To centrale beslutninger, som resultaterne af den kontrafaktiske konsekvensanalyse ofte bidrager til, er:

- Hvorvidt en eksisterende intervention bør fortsætte.
- Hvorvidt en ny interventionstype bør gennemføres mere bredt (dvs. opskaleret).

I det første tilfælde kan en kontrafaktisk konsekvensanalyse forsøge at vurdere effektiviteten af et eksisterende eller igangværende program, hvor budgetterne er under pres, og der er potentielle alternative anvendelser af de pågældende ressourcer. I denne situation er det usandsynligt, at interventionen tidligere er blevet evalueret ved hjælp af en kontrafaktisk metode.

I det andet tilfælde kan der være begrænsninger for gennemførelsen af interventionen. Den kan f.eks. gennemføres i en bestemt region eller et bestemt område i en medlemsstat eller kun i en begrænset periode. I disse sammenhænge kan resultaterne fra en kontrafaktisk konsekvensanalyse anvendes til at afgøre, om den pågældende intervention er effektiv og derfor kan gennemføres andre steder. I sådanne situationer gennemføres der pilotprojekter eller afprøvninger af en intervention inden en bredere udrolning.

Hvilke spørgsmål skal besvares?

Når målene for interventionen og evalueringens formål og endelige anvendelsesformål er fastlagt, og målgruppen er klart identificeret, bør det være muligt at præcisere de spørgsmål, som den kontrafaktiske konsekvensanalyse skal behandle. Under mange omstændigheder vil en række målgrupper og interessenter have spørgsmål vedrørende årsagssammenhæng, som de ønsker evalueret i den kontrafaktiske konsekvensanalyse. Forvaltningsmyndigheden eller evaluatoren bør indsamle disse spørgsmål gennem omfattende høringer, hvor der tages hensyn til forskellige synspunkter og forslag fra personer, der er involveret på forskellige stadier af interventionen. Efterfølgende skal forvaltningsmyndigheden og/eller evaluatoren prioritere spørgsmålene og fokusere den kontrafaktiske konsekvensanalyse på dem, der er vigtigst mest og hensigtsmæssige.

Nogle af de spørgsmål, der kan tages i betragtning i forbindelse med færdiggørelsen af en liste over de centrale researchspørgsmål til en kontrafaktisk konsekvensanalyse, omfatter:

*Centrale
researchspørgsmål*

- Har interventionen skabt eller bidraget til de tilsigtede outcomes på kort, mellemlang og lang sigt? Og afveg kortsigtede virkninger væsentligt fra virkningerne på lang sigt? Disse spørgsmål bør prioriteres.
- Er det muligt at få et kvantitativt mål for outcomes? Den kontrafaktiske konsekvensanalyse skal baseres på et passende sæt data (administrative data eller direkte undersøgelser sendt til deltagerne) uafhængigt af en foretrukken kilde.
- I hvilket omfang kan ændringer i deltagernes forhold eller i den socioøkonomiske kontekst tilskrives interventionerne? Denne type spørgsmål kræver en måling af foranstaltningens nettoeffekt med henblik på sammenligning med kontroldeltagere og kontekstindikatorer.
- Var virkningerne af interventionen de samme for alle medlemmer af målgruppen? Var f.eks. virkningen af en intervention rettet mod langtidsledige den samme for mænd og kvinder? Og for personer under 25 år eller over 50 år? Den kontrafaktiske konsekvensanalyse er god til undersøgelse af virkninger i forskellige undergrupper, hvis antallet af personer i behandlings- og kontrolgrupperne er tilstrækkeligt højt.
- Har interventionen været omkostningseffektiv (sammenlignet med alternativer)? Og hvad er forholdet mellem omkostninger og fordele? Denne effektivitetsanalyse kræver, at den kontrafaktiske konsekvensanalyse giver et nøjagtigt billede af virkningerne samt præcise oplysninger om direkte og indirekte omkostninger og fordele.
- Findes der mange oplysninger og stor viden om virkningerne af lignende interventioner? I hvilket omfang er denne viden relevant for den undersøgte intervention? Når en politik er velkendt, og dens virkninger er blevet grundigt undersøgt, kan det være nyttigt at fokusere evalueringsspørgsmålene på specifikke aspekter af politikken og således undgå at gentage andre analyser. En detaljeret gennemgang af eksisterende litteratur kan danne grundlag for beslutninger i denne henseende, og det er under alle omstændigheder en nyttig hjælp ved udformningen af kontrafaktiske konsekvensanalyser.

Det er vigtigt at have en klar idé om den vifte af researchspørgsmål, som en kontrafaktisk konsekvensanalyse skal besvare, inden den iværksættes. Drøftelse af de spørgsmål, som evalueringen vil behandle, er et centralt element i enhver evalueringsordning.

Det er vigtigt at prioritere spørgsmål og ikke komme til at følge den populære tendens med at overfylde en evaluering med for mange spørgsmål. Det er en vanskelig balance at sikre, at evalueringen er relevant for en række interessenter med forskellige interesser, og at den samtidig er til at gennemføre. Hvis en evaluering har behandlet en for bred vifte af researchspørgsmål, kan den miste fokus og ende med at spænde over for mange emner, hvilket ikke bliver optimalt. Det handler ofte om "less is more" — prioriteringen er en afgørende fase i planlægningen af evalueringen.

*Prioritering af
spørgsmål*

For at prioritere evalueringsspørgsmål er det nødvendigt at undgå dubletter og tildele point til hvert spørgsmål i henhold til relevante principper. Det kan være: Betydningen af og det reelle engagement hos den interessent, der

formulerede spørgsmålet, spørgsmålets egnethed og overensstemmelse med programmets forandringsteori, spørgsmålets relevans for det generelle formål med evalueringen, muligheden for at stille spørgsmålet i betragtning af de data, den tid og de ressourcer, der er til rådighed³⁴. Den deraf følgende rangorden vil sortere spørgsmålene efter vigtighed og gøre det muligt at udvælge de mest relevante.

I nogle tilfælde er det muligt at stille en gruppe evalueringsspørgsmål under et mere generelt spørgsmål. F.eks. under et spørgsmål som "hvad har interventionens nettovirkning været?". Andre spørgsmål, såsom virkningerne for forskellige grupper af deltagere, måske endda på forskellige tidspunkter, kunne være underspørgsmål. Denne proces skal imidlertid altid sikre, at man ender med et sæt spørgsmål, der er til at besvare og bruge.

Hvilke evalueringskriterier kan der knyttes til evalueringen?

Forholdet mellem evalueringskriterier og -spørgsmål er blevet nævnt ovenfor, men det bør uddybes yderligere. Evalueringskriterier (produktivitet, effektivitet, EU-merværdi, sammenhæng osv.) er nødvendige for at tillægge alle indsamlede beviser en værdi og nå frem til en vurdering af en politik (produktivitet osv.). Evalueringsspørgsmålene er derimod nødvendige for at sikre, at opdragshaverens krav fremgår tydeligt, og at der fokuseres på de vigtigste politiske spørgsmål, der står på spil. Evalueringskriterier og evalueringsspørgsmål er imidlertid forbundne. Alle spørgsmål kan generelt knyttes til et specifikt kriterie, og denne tilknytning er vigtig for både forvaltningsmyndigheder og evaluatore, fordi det forbinder evalueringens udformning, der nødvendigvis er baseret på spørgsmålene, med kravene i forordningen om fælles bestemmelser for 2021-2027, som vedrører evalueringskriterierne på en anden måde.

Nedenfor gives nogle eksempler på typiske evalueringsspørgsmål grupperet med deres tilhørende kriterier. De er udvalgt og tilpasset fra Kommissionens værktøjskasse for bedre regulering³⁵ (værktøj nr. 47 om evalueringskriterier og -spørgsmål) for at vise eksempler på spørgsmål, der kræver en konsekvensanalyse og, hvor det er muligt, en kontrafaktisk konsekvensanalyse:

- *Typiske eksempler på spørgsmål vedrørende effektivitet*
 - Hvilke kvantitative virkninger har interventionen haft?
 - I hvilken grad kan disse ændringer/virkninger tilskrives interventionen?
 - I hvilken grad kan faktorer, der påvirker de observerede resultater, knyttes til EU-interventionen?
- *Typiske eksempler på spørgsmål vedrørende produktivitet*
 - I hvilken grad har interventionen været omkostningseffektiv?
 - I hvilken grad er omkostningerne til foranstaltningen berettiget og står i rimeligt forhold til de ændringer/den effekt, der er opnået?
 - Hvis der er betydelige forskelle i omkostninger (eller fordele) mellem områder, hvad skyldes de så? Hvordan hænger disse forskelle sammen med interventionen?
- *Typiske spørgsmål om EU-merværdi*

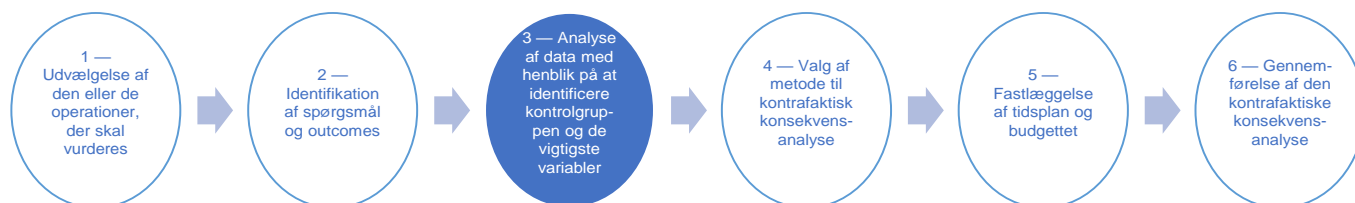
*Evalueringskriterier
og
evalueringsspørgsmål*

³⁴ Se f.eks. [Centre for Disease Control and Prevention CDC, 2013](#).

³⁵ Se [Europa-Kommissionen, 2017](#).

- Hvad er merværdien af ESF+-interventionerne sammenlignet med, hvad der skabes af lignende nationale og/eller regionale interventioner?
- Hvad ville de mest sandsynlige konsekvenser være af at stoppe eller annullere den eksisterende ESF+-intervention?

2.3. Data til identifikation af kontrolgruppen og måling af outcomevariablerne



2.3.1. Er de relevante data til rådighed, eller kan de stilles til rådighed?

Ifølge forvaltningsmyndigheder og evalueringsekspertter fra hele EU tyder det på, at adgang til relevante data er en af de største udfordringer i forbindelse med gennemførelsen af kontrafaktiske konsekvensanalyser. En vigtig praktisk overvejelse er, om de nødvendige datatyper er tilgængelige. I dette afsnit præsenteres en forenklet kategorisering af de typer data, der kræves, sammen med en redegørelse for, hvilke kilder sådanne data kan hentes fra, eller hvilke typer af primære metoder til dataindsamling der kan være behov for. Det meget vigtige spørgsmål om databeskyttelse behandles også.

Vigtigheden af grundig planlægning kan ikke fremhæves nok. Nogle kontrafaktiske konsekvensanalyser er til en vis grad blevet hæmmet af en mangel på data, fordi man ikke lagde en plan i god tid. I forbindelse med eksisterende interventioner er det vigtigt at identificere medlemmerne af de behandlede og ikkebehandlede grupper og etablere mekanismer til indsamling af data fra dem, da de skal danne grundlag for evalueringen. I forbindelse med nye interventioner bør der på et tidligt tidspunkt tages skridt til at sikre, at de rette typer data indsamles på passende tidspunkter.

Planlægning af dataindsamling

Hvilke typer data skal bruges?

Primære datakilder

Generelt kræves der tre typer data for at kunne gennemføre en kontrafaktisk konsekvensanalyse. I nogle tilfælde kan en enkelt datakilde indeholde en eller flere af disse typer. Det drejer sig om: fortegnelser over behandlings- og kontrolgrupper, outcomes og kontekstdata.

- **Fortegnelser over behandlings- og kontrolgrupper:** Der kræves datakilder, som gør det muligt for evaluatorene at identificere individuelle behandlings- og kontrolgrupperenheder (virksomheder, personer eller potentielt geografiske områder).
- **Fortegnelser over outcomes:** Som vist i figur 6 og 7 i kapitel 3 i denne vejledning kræver en kontrafaktisk konsekvensanalyse, at outcomes måles for både behandlings- og kontrolgrupper. Ideelt set bør data om outcomes for begge grupper indsamles ved hjælp af de samme indsamlingsmetoder og resultatmålinger foretaget på de samme

tidspunkter.

- **Fortegnelser over kontekstdata:** Der kræves data til at udvælge velmatchede kontrol- og behandlingsgrupper og gøre det muligt at kontrollere eventuelle resterende forskelle mellem de to grupper i analysen. Det er vigtigt at indsamle så mange data som muligt om enhedernes egenskaber og faktorer, der kan være forbundet med både valget om at deltage i en intervention og de potentielle resultater, navnlig resultatindikatorer, der måles forud for behandlingen. Kontekstdata kan også omfatte data, der beskriver lokale arbejdsmarkeder (f.eks. lokale arbejdsløshedstal eller mål for ophedning på arbejdsmarkedet), og data, der gør det muligt at analysere de enkelte undergrupper.

Tabel 2 nedenfor viser disse tre datatyper og angiver de mulige kilder til dem. Eksempler på data, der anvendes til kontrafaktiske analyser for ESF, vises i den første boks nedenfor, mens den anden boks illustrerer eksempler på integrerede datasæt, der anvendes til analyse af arbejdsmarkedsspørgsmål og i evalueringer af arbejdsmarkedspolitikker.

Tabel 2 Datatyper og kilder

Datatyper	Kilder
Fortegnelser over behandlingsgrupper	<ul style="list-style-type: none"> - Fortegnelser over deltagere i interventioner (normalt hos støttemodtagerne) - ESF+-overvågningsdata (beskrivelse af interventionen, start og slut, henvisninger, ansøgninger)
Fortegnelser over kontrolgrupper	<ul style="list-style-type: none"> - Administrative data såsom fortegnelser over socialsikring, uddannelse og arbejdsløshedsunderstøttelse - Fortegnelser over deltagere (personer, der var berettigede til at deltage, men som ikke gjorde det af andre årsager³⁶) - Eksisterende nationale undersøgelser, såsom arbejdsstyrkeundersøgelsen
Fortegnelser over outcomes (påkrævet for både behandlings- og kontrolgrupper)	<ul style="list-style-type: none"> - Administrative oplysninger: F.eks. kan oplysninger om social sikring og arbejdsløshed også anvendes til at sammensætte resultatmål (indtægter fra ydelser/socialsikring), nationale forsikrings- og skattefortegnelser (indkomst- og beskæftigelsesoutcomes) - Administrative fortegnelser fra uddannelsesinstitutioner (standardiserede test af opnåede kompetencer, færdiggørelse, optagelse og fremmøde) - Officiel virksomhedstælling eller skattefortegnelser (produktivitet eller omsætning før og efter intern uddannelse eller nye ansættelser) - Beskæftigelses- eller outputdata fra officielle statistikker (i kontrafaktiske territoriale analyser til måling af beskæftigelses- og BNP-niveauer) - Skræddersyede undersøgelser af behandlings- og kontrolgrupper
Kontekstdata/kontrolvariabler (påkrævet for både behandlings- og kontrolgrupper)	<ul style="list-style-type: none"> - Administrative systemer — f.eks. fortegnelser over ydelser, der viser, hvilke ydelser den pågældende har modtaget før behandlingen, nationale forsikrings- og skattefortegnelser, historiske fortegnelser over indkomst og beskæftigelse - Officielle statistikker om arbejdsmarkedet eller uddannelse (f.eks. arbejdsstyrkeundersøgelse, grundlæggende data på regionalt eller nationalt plan, som også giver mikrodata på individuelt niveau med henblik på konkrete vurderinger) - Undersøgelser af kontrol- og behandlingsgrupper. Hvis behandlingsreglerne er klare, kan kontrolgrupperne identificeres på forhånd, og der kan indsamles basislinjedata - Overvågningssystemer — i nogle tilfælde kan overvågningssystemerne bruges til at indsamle data fra både behandlings- og kontrolgrupper, f.eks. ansøgningssystemer, hvor ansøgere, der ikke er blevet valgt, kan anvendes som kontrol

Tekstboks 6 Eksempler på data til kontrafaktiske konsekvensanalyser

Blandt de eksempler, der gennemgås nærmere i denne vejledning, er de tyske evalueringer sandsynligvis dem, der er baseret på de mest omfattende administrative datasæt. Integrierten Erwerbsbiografie (arbejdsmarkedsregister (IEB)) integrerer en række forskellige administrative informationskilder, der forvaltes af det føderale arbejdsformidlingskontor, og indeholder oplysninger om en persons beskæftigelse (med undtagelse af selvstændig virksomhed) og arbejdsløshedsepisoder, om sociodemografiske

³⁶ Disse er deltagere, som er "berettigede, men ikke optaget". Hvis deltagere ikke optages, er det ofte, fordi finansieringen er ophørt, eller af andre eksterne årsager (f.eks. sygdom hos underviseren, transportproblemer osv.). Disse kan være en relevant kontrolgruppe, fordi de er berettiget til at deltage af de samme grunde og er lige så villige til at deltage som de faktiske deltagere. Det er imidlertid nogle gange ikke nok til, at de kan udgøre en kontrolgruppe, eller det kan ske, at deres oplysninger ikke er registreret.

karakteristika, om overførselsindkomst (arbejdsløshedsunderstøttelse) og om deltagelse i aktive arbejdsmarkedspolitikker. IEB har eksisteret i en lang årrække, og det har krævet en løbende indsats og investering at udvikle og forvalte datasættet. Tekstboks 7 indeholder flere oplysninger om IEB³⁷. I den tyske evaluering af jobskabelsesordninger blev disse data suppleret af en undersøgelse opdelt i tre bølger med henblik på at indsamle oplysninger om "bløde outcomes" (som opfattet af de enkelte personer) på forskellige stadier under interventionerne.

I de italienske evalueringer, der gennemgås nærmere i vejledningen, blev den samme type data anvendt. Arkiverne over arbejdsløse registreret hos de offentlige arbejdsformidlinger blev brugt til at identificere potentielle kontrolgrupper, mens administrative data om beskæftigelsesmæssige ændringer (Comunicazioni obbligatorie — COB), der registrerede alle virksomheders og offentlige arbejdsgiveres ansættelseskontrakter (undtagen selvstændige erhvervsdrivende), blev anvendt til at måle outcomevariablerne. Da COB forvaltes på regionalt plan, varierer informationskvaliteten. Der er imidlertid blevet indført landsdækkende standarder i de seneste år, og kvaliteten er forbedret. I en evaluering (for provinsen Trento) var evaluatoren også i stand til at flette disse data sammen med oplysningerne om selvangivelser fra INPS (det nationale institut for social sikring), således at man kunne måle virkningen i forhold til den enkeltes indkomst.

I Polen blev der i evalueringerne anvendt administrative data fra de offentlige arbejdsformidlingers arbejdsløshedsregistre til at identificere kontrolgrupper. Disse data indeholder imidlertid ikke oplysninger om den enkeltes beskæftigelsesstatus eller historie, og evalueringerne skulle baseres på en proxy (annullering af status som arbejdsløs hos de offentlige arbejdsformidlinger) eller på indsamling af oplysninger gennem undersøgelser af en stikprøve af personer, der både behandles og ikke behandles.

I Letland var evaluatoren i stand til at sammenlægge to vigtige administrative datasæt. Data fra Letlands statslige arbejdsformidlingskontor gav oplysninger om både deltagere og ikkedeltagere, der var registreret som arbejdsløse på bestemte datoer, og data fra skattemyndighederne gav oplysninger om ansættelsesvilkår på forskellige datoer samt om enkeltpersoners indkomst. Dette gjorde det muligt for evaluatoren at vurdere virkningerne både med hensyn til sandsynligheden for at være i beskæftigelse på forskellige tidspunkter og med hensyn til indkomst.

Det eneste eksempel på randomisering, det svenske eksempel, anvendte både administrative data til at måle outcomevariabler og en undersøgelse af arbejdsformidlingskontorer og formidlere til at måle intensiteten og arten af den ydede støtte.

Evalueringerne på uddannelsesområdet anvendte administrative data til at identificere kontrolgruppen og vurdere outcomevariablerne, med undtagelse af Polen, som var nødt til at bruge en undersøgelse til at måle de variabler, der blev anvendt som outcomes.

Generelt viser eksemplerne betydningen af passende administrative data for kontrafaktiske konsekvensanalyser, ideelt set både med henblik på identifikation af kontrolgrupper og måling af outcomevariabler. Ud fra dette synspunkt bør forvaltningsmyndigheder, der planlægger kontrafaktiske konsekvensanalyser, på forhånd sikre, at de administrative data til evalueringerne er tilgængelige, og træffe passende foranstaltninger til at tackle potentielle problemer med tilgængelighed, integration af data eller andre problematiske aspekter.

Tekstboks 7 Eksempler på integrerede databaser til kontrafaktiske konsekvensanalyser

Jobseekers Longitudinal Dataset (JLD) i Irland

Jobseekers Longitudinal Dataset (JLD) er en administrativ database, der forvaltes af departementet for social beskyttelse (Department of Social Protection (DSP)).

JLD er et ambitiøst forsøg på at tilpasse administrative data til forskningsformål. Man begyndte at udvikle den for ca. 10 år siden, efter at DSP havde bedt University College, Dublin, om at gennemføre en indledende undersøgelse af forvaltningen af Live Register³⁸ og mere generelt data vedrørende

³⁷ Et lignende eksempel ses i Irlands "Jobseekers Longitudinal Dataset" (JLD), hvor oplysninger om betalinger til sociale bistandsordninger, arbejdsmarkedsprogrammer og beskæftigelsehistorik samles i én database. JLD blev bl.a. brugt i evalueringen af aktiveringsprogrammet JobBridge, som blev finansieret i programmeringsperioden 2007-2013. Se [Indecon, 2016](#).

³⁸ Live Register indeholder oplysninger om de personer, der har ansøgt om social sikringsydelse (Jobseekers' Benefit

arbejdsmarkedet. Undersøgelsen indeholdt flere forslag til forbedring af dataindsamlingen og identificerede en række udfordringer (f.eks. duplikering af data i forskellige datasystemer, manglende oplysninger osv.).

JLD integrerer flere informationskilder: betalingsdata og administrative data fra DSP — f.eks. betalinger af social bistand og socialforsikring til befolkningen i den erhvervsaktive alder, der er opført i Live Register, og data vedrørende aktive arbejdsmarkedsprogrammer, der forvaltes af DSP, data om arbejdsmarkedsprogrammer, der forvaltes af SOLAS, som er det nationale uddannelsesorgan, og data indsamlet af skattemyndighederne (Revenue Commissioners). Med hensyn til ESF-interventioner dækker JLD delvis de interventioner, der finansieres af fonden, men mere specifikt dem, der finansieres af DSP og SOLAS³⁹.

Data fra ovennævnte kilder omgrupperes som en række episoder, hvor én episode begynder, når personen begynder en periode med arbejdsløshed, og slutter, når personen skifter til beskæftigelse eller en anden aktivitet eller en uddannelse. Når en persons status ændres igen, begynder en ny episode så. Starten på en arbejdsløshedsperiode falder sammen med, at en person begynder at modtage Jobseekers' Benefit og Jobseekers' Allowance⁴⁰. I JLD er en fordel ved denne struktur med opdeling i episoder, at sammenhængende perioder på Jobseekers' Benefit og Jobseekers' Allowance kan kædes sammen og vises som én episode med arbejdsløshed. Det er værd at bemærke, at episoder som nævnt ovenfor kan overlape hinanden, og at det derfor kan være vanskeligt at prioritere én episode frem for en anden.

JLD har siden 2004 registreret udbetalinger af sociale ydelser, aktivering, uddannelse og beskæftigelsehistorik og indeholder ca. 13 mio. individuelle episoder omfattende perioder med ydelser og arbejde for omkring 2 millioner mennesker. Hver episode har en start- og slutdato og en operationel kode⁴¹, der gør det muligt at identificere en enkeltpersons situation i hver periode. JLD omfatter en lang række "variabler": køn, alder, ægteskabelig status, statsborgerskab, uddannelsesniveau, tidligere beskæftigelse, beskæftigelses- og arbejdsløshedshistorik (varighed og antal episoder) og jobkarakteristika (f.eks. sektor), kurser for arbejdsløse (type, varighed og antal episoder), ydelsestype (JA, JB), antal børn, forsørgelsesberettigede pårørende, familiejydelsestype (dvs. tilskud til voksne og børn, kun voksne osv.), indkomst og skat og geografisk placering⁴².

Desuden forbinder individuelle identifikatorer JLD med andre administrative data. I forbindelse med evalueringen af JobsPlus blev JLD f.eks. matchet med en separat overvågningsdatabase med detaljerede oplysninger om start- og slutdatoer for JobsPlus, antal dage i Live Register i begyndelsen af JobsPlus og behandlingstype.

Der er to primære kanaler til at få adgang til JLD-data: forskning på kontrakt, som finder sted, når DSP lancerer en proces for anmodning om tilbud (Request for Tender (RFT)), eller efter anmodning fra forskere med passende kvalifikationer. I sidstnævnte tilfælde behandles anmodninger om adgang fra sag til sag, og hvis de imødekommes, skal der underskrives en juridisk bindende aftale om datadeling mellem forskerne/institutionerne og DSP. Alle data er pseudonymiserede, og kun de datafelter, der er relevante for forskningsemnet, overføres. Forskerens eller forskningsinstitutionens adgang til data er begrænset, og data skal slettes ved afslutningen af forsknings- eller evalueringsprojektet. Der stilles en kodebog til rådighed for forskere sammen med JLD-dataene, selv om der er behov for en betydelig forbedring af dokumentationen.

JLD har i årenes løb været anvendt til en række forskningsaktiviteter og evalueringer: Back to Education Allowance (2015), JobBridge Activation Programme (2016), Back to Work Enterprise Allowance (2017), JobPath (2019) og JobPlus (2020) er nogle af de initiativer, der er blevet udvalgt til evaluering.

Strukturen og indholdet i JLD er blevet forbedret siden starten. Det kræver dog en konstant indsats for at

(JB)) eller arbejdsløshedsunderstøttelse (Jobseekers' Allowance (JA)) eller forskellige andre lovbestemte ydelser hos DSP's lokalafdelinger.

³⁹ Programmer på områder som retlige anliggender (Youth Diversion Projects og Young Persons Probation Projects samt Integration and Employment of Migrants), uddannelse (Third Level Access og Adult Literacy), lokalsamfund (Community Training Centres) og forsvar (Defence Forces Employment Support Scheme) er ikke omfattet af JLD.

⁴⁰ Jobseekers' Benefit er en ugentlig betaling fra DSP til personer, der er arbejdsløse, enten fuldt ud eller i deltidsbeskæftigelse, fordi deres arbejdstid er blevet nedsat af deres arbejdsgiver. Den udbetales til arbejdsløse, som har betalt det krævede socialsikringsbidrag (PRSI), og som har indbetalt tilstrækkeligt i det relevante skatteår, og har indbetalt mindst 104 bidrag i alt. Jobseekers' Allowance er en indtægtsbestemt ydelse til arbejdsløse, som ikke er berettiget til at modtage Jobseekers' Benefit, eller hvis ret til Jobseekers' Benefit er udløbet.

⁴¹ Koden repræsenterer en kombination af aktiverings-/uddannelsesaktiviteter, offentlige ydelser og arbejdsløshedsperiodens varighed.

⁴² Det skal bemærkes, at nogle data henviser til forskellige perioder, hvor f.eks. indkomst registreres på årsbasis, mens udbetalinger af sociale sikringsydelser opdateres hver uge.

holde den ajour og udvikle forbedringer og udfylde hullerne, f.eks. ved at medtage mere detaljerede oplysninger om uddannelse og integrere mere programspecifikke data (indhold, gennemførelse osv.) og forbedre indkomstdataene. Et kommende projekt bør sikre forbedringer af JLD med hensyn til: regelmæssig opdatering (f.eks. månedlig) med automatisering og afprøvning af kilderne til dataene og tilføjelse eller udskiftning af datakilder for at sikre fuld dækning af beskæftigelses- og uddannelsesstøtte, andre sociale beskyttelsesprogrammer og realtidsindkomstdata⁴³.

Instituttet for beskæftigelsesforskning (IAB) i Tyskland og arbejdsmarkedsregisteret (IEB)

I Tyskland blev adgangen til data til videnskabelige formål forbedret efter arbejdsmarkedsreformerne i 2003-2005, idet der blev lagt vægt på evalueringen af politikker og Kommissionens henstillinger om "forbedring af informationsinfrastrukturen mellem videnskab og statistik" og oprettet et forskningsdatacenter hos hver offentlig producent af mikrodata. Som følge heraf oprettede det føderale arbejdsformidlingskontor i 2004 et forskningsdatacenter under Institut für Arbeitsmarkt- und Berufsforschung (institut for arbejdsmarkedsforskning (IAB)). IAB er ansvarlig for at udtrække data fra administrative processer med henblik på at udarbejde datasæt, der kan anvendes til empirisk forskning.

IAB forvalter mere end 15 datasæt, og de er tilgængelige for forskere. Dataene stammer fra administrative data fra socialsikringsystemet, den føderale arbejdsformidlings interne processer og undersøgelser foretaget af IAB⁴⁴.

I forbindelse med de sociale sikringsordninger skal alle arbejdsgivere indberette en række oplysninger og karakteristika for deres ansatte, og disse data giver et omfattende sæt oplysninger om folks beskæftigelsehistorik. De administrative data vedrørende de interne procedurer i den føderale arbejdsformidling omfatter obligatorisk arbejdsløshedsforsikring, arbejdsløshedsunderstøttelse og de tilsvarende optjeningsperioder, rådgivningsmøder med arbejdsløse, jobtilbud og aktive arbejdsmarkedsforanstaltninger. IAB kombinerer dataene i et enkelt omfattende datasæt, "Integrierten Erwerbsbiografie (IEB)". Indsamlingen af disse administrative data begyndte i 1975, men ikke alle variabler er tilgængelige for hele observationsperioden. IEB kan integreres med flere data fra undersøgelser gennemført af IAB, såsom: IAB's virksomhedspanel, IAB's undersøgelse af ledige stillinger, undersøgelsen German Management and Organizational Practices (GMOP), panelundersøgelsen om arbejdsmarkedet og social sikring (PASS), undersøgelsen om arbejde og læring i en verden i forandring, medarbejderundersøgelsen om bonusbetalinger, lønstigninger og retfærdighed (BLoG), sammenkædede data om arbejdsgivere og medarbejdere IAB (LIAB), Panelet "WeLL" — arbejdstagerundersøgelse i forbindelse med projektet om videreuddannelse som led i livslang læring.

IAB ajourfører regelmæssigt sine dataprodukter og tilbyder forskellige udsnit af disse omfattende administrative datakilder til forskningsformål. For hvert dataprodukt leverer IAB detaljeret dokumentation på tysk og engelsk. Retsgrundlaget for dataadgang findes i den tyske sociallovbog (flere versioner), og mere specifikt er der mulighed for fire former for adgang til data for forskersamfundet: Campusfiler (fuldt anonymiserede og kan kun bruges til undervisning), og filer om videnskabelig anvendelse, som er anonymiserede mikrodata, der indsendes til videnskabelige institutioner i Tyskland og EU's medlemsstater med henblik på forskningsprojekter inden for arbejdsmarkedsforskning, men ikke med henblik på undervisning eller kommercielle forskningsinteresser. Datasikkerheden skal garanteres af den videnskabelige institution, der anmoder om oplysningerne. Svagt anonymiserede data med mere detaljerede oplysninger er kun tilgængelige på stedet. IAB stiller separate arbejdspladser til rådighed i et sikkert databehandlingsmiljø i Nürnberg og forskellige steder i Tyskland, USA og Det Forenede Kongerige. Forskere har direkte adgang til data, men de kan først få resultaterne af deres programmer, når de er blevet kontrolleret med hensyn til databeskyttelse af IAB's personale. Fjernbehandling betyder, at forskere forbereder deres programmer med kunstige data og uploader dem i joboverførselsapplikationen (JoSuA). I denne proces ser forskere aldrig de oprindelige data, og de modtager kun deres resultater. Der findes standardiserede anmodningsformularer for al dataadgang. Når en anmodning er godkendt, indgås der mellem forskerens institution og IAB en kontrakt om anvendelse af data til et specifikt projekt inden for en bestemt periode. I kontrakten specificeres databeskyttelsesreglerne, og der pålægges strenge sanktioner i

⁴³ Vi vil gerne takke Frank Humphreys, Ciaran Judge, Saidhbhín Hardiman og Krzysztof Gigon fra DSP for oplysningerne om JLD.

⁴⁴ I 2011 blev der oprettet et center for sammenkædning af data i IAB i et fælles projekt mellem Duisburg-Essen Universitet finansieret af den tyske forskningsfond. Centeret har til formål at forenkle sammenkædningen af datasæt uden en specifik identifikator.

tilfælde af overtrædelse. Nogle af de datasæt, der forvaltes af IAB, er kun tilgængelige til brug på stedet (f.eks. de sammenkædede datasæt).

På grundlag af data fra IAB evalueres aktive arbejdsmarkedspolitikker regelmæssigt ved hjælp af de seneste empiriske metoder, og i nogle tilfælde har resultaterne ført til en ændring af politikkerne⁴⁵. IAB giver ikke kun forskere adgang til sine datasæt i Tyskland, men også i andre lande, og antallet af brugere er støt stigende (f.eks. var næsten en tredjedel af alle aftaler om dataanvendelse i 2016 fra en institution uden for Tyskland).

Med hensyn til anvendelsen af administrative data til kontrafaktiske konsekvensanalyser og dermed forbundne praktiske spørgsmål henvises også til den europæiske vejledning udarbejdet af forskere fra JRC, [Europa-Kommissionen, 2020](#).

Kan der være problemer i forhold til databeskyttelse?

Det kan være vanskeligt at indhente data, der identificerer personer eller virksomheder, der har deltaget i ESF+-finansierede interventioner⁴⁶. Kontrafaktiske konsekvensanalyser kræver mikrodata — data, der indeholder observationer om de enkelte enheder i både behandlings- og kontrolgrupperne.

ESF+-forordningen 2021/1057 (bilag I) kræver oplysninger om deltagere opdelt efter køn, beskæftigelsesstatus, aldersgruppe, uddannelsesniveau og sårbare grupper (indvandrere, minoriteter, handicappede, andre dårligt stillede). Forordningen om fælles bestemmelser og ESF+-forordningen for 2021-2027 fastsætter en retlig forpligtelse for forvaltningsmyndighederne til at indsamle og behandle personoplysninger i form af individuelle deltagerregistreringer. Desuden præciseres følgende i artikel 17, stk. 6, i ESF+-forordningen (EU) 2021/1057 om indikatorer og rapportering: "Hvor data er tilgængelige i registre eller fra tilsvarende kilder, kan medlemsstaterne sætte forvaltningsmyndigheder og andre organer med ansvar for indsamling af data, der er nødvendige for overvågningen og evalueringen af generel støtte fra ESF+-indsatsområdet under delt forvaltning, i stand til at indhente data fra disse registre eller tilsvarende kilder i overensstemmelse med artikel 6, stk. 1, litra c) og e), i forordning (EU) 2016/679."

*ESF+-forordningen
om datastyring*

Disse regler, der er fastsat i forordningen om fælles bestemmelser og ESF+-forordningen, letter adgangen til og anvendelsen af de personoplysninger, der er nødvendige for ESF+-overvågningen og, for så vidt angår en kontrafaktisk konsekvensanalyse, for definitionen af den behandlede gruppe. Adgang til personoplysninger, der er nødvendige for at danne kontrolgrupperne, og behandlingen af behandlings- og kontrolgruppedata skal imidlertid være i overensstemmelse med forordning (EU) 2016/679 (generel forordning om databeskyttelse), som omhandler generel overførsel og brug af personoplysninger, herunder særlige kategorier af oplysninger⁴⁷ inden for EU. Nedenstående tekstboks giver et overblik over det vigtigste indhold og de

⁴⁵ Som eksempel kan nævnes en evaluering af den obligatoriske integrationsaftale mellem den arbejdsløse og sagsbehandleren. Ved hjælp af et randomiseret feltforsøg og ved hjælp af arbejdsmarkedsfortegnelserne for de personer, der var med i forsøget, kunne IAB dokumentere, at den obligatoriske aftale for nogle grupper af arbejdsløse kan virke mod hensigten og bør erstattes af en mere fleksibel anvendelse af instrumentet (van den Berg et al. 2016).

⁴⁶ Se [Ismeri Europa — Ecorys — Institute for Employment Studies, 2019](#).

⁴⁷ I forordningen bruges udtrykket følsomme data ikke længere, men i stedet henvises der til "særlige kategorier af personoplysninger". Ifølge artikel 9 omfatter disse: "personoplysninger om race eller etnisk oprindelse, politisk, religiøs eller filosofisk overbevisning eller fagforeningsmæssigt tilhørsforhold samt behandling af genetiske data, biometriske data med det formål entydigt at identificere en fysisk person, helbredsoplysninger eller oplysninger om en fysisk persons seksuelle forhold eller seksuelle orientering".

vigtigste forpligtelser, der er reguleret i den generelle forordning om databeskyttelse.

Tekstboks 8 EU's regelsæt for behandling af personoplysninger

Den relevante lovgivning om behandling af personoplysninger i Europa består hovedsagelig af forordning (EU) 2016/679 og de retningslinjer og foranstaltninger, der er vedtaget af kompetente myndigheder såsom Det Europæiske Databeskyttelsesråd (EDPB). Denne lovgivning fastsætter mange betingelser og begrænsninger for behandlingen af personoplysninger for at beskytte registreredes rettigheder og frihedsrettigheder. Under alle omstændigheder skal det behov for beskyttelse, der følger af bestemmelserne og forordningerne, finde en balance i forhold til behovet for ikke at begrænse den videnskabelige forskning og faktisk fungere som et aktiv i udviklingen heraf. Derfor og inden for disse retlige rammer har Den Europæiske Union fastsat, at der på særlige betingelser kan gøres undtagelser for at tillade forskningsaktiviteter og formidling af resultaterne, så længe de berørte personers vigtigste rettighed, dvs. retten til privatlivets fred, beskyttes⁴⁸.

DEN GENERELLE FORORDNING OM DATABESKYTTELSE (FORORDNING (EU) 2016/679)

Den generelle forordning om databeskyttelse trådte i kraft den 24. maj 2016 og er blevet anvendt fuldt ud i alle medlemsstater siden den 25. maj 2018. Forordningen finder anvendelse på "behandling af personoplysninger, der helt eller delvis foretages ved hjælp af automatisk databehandling, og på anden ikkeautomatisk behandling af personoplysninger, der er eller vil blive indeholdt i et register"⁴⁹.

For så vidt angår det territoriale anvendelsesområde finder forordningen anvendelse både på behandling af personoplysninger, der foretages af personer, der befinder sig på Den Europæiske Unions område, og i tilfælde, hvor behandlingen involverer registrerede, der befinder sig i Den Europæiske Union, selv når den dataansvarlige eller databehandleren befinder sig uden for EU.

Ud fra et materielt synspunkt og sammenlignet med den tidligere lovgivning styrker forordningen de registreredes rettigheder og pålægger de dataansvarlige en række forpligtelser i henhold til en logik baseret på risikoanalyse og princippet om ansvarlighed. Desuden fastsætter forordningen en række krav, som den dataansvarlige skal opfylde, og som ikke var omfattet af den tidligere lovgivning: Det gælder f.eks. konsekvensanalysen vedrørende databeskyttelse (artikel 35-36), minimeringen af behandlingsaktiviteterne i henhold til kriterierne om databeskyttelse gennem design og databeskyttelse gennem standardindstillinger (artikel 25), vedtagelsen af fortegnelsen over behandlingsaktiviteter (artikel 30) og udnævnelsen af den databeskyttelsesansvarlige (artikel 37-39).

Det er vigtigt at fremhæve hovedprincipperne i forordningens artikel 5, som de, der behandler personoplysninger, skal overholde. Det gælder navnlig følgende:

- a) Lovlighed, rimelighed og gennemsigtighed: Personoplysninger skal behandles lovligt, rimeligt og på en gennemsigtig måde.
- b) Formålsbegrænsning: Der indsamles oplysninger til udtrykkeligt angivne og legitime formål, og senere behandling heraf må ikke være uforenelig med disse formål. I tilfælde af behandling til statistiske formål eller videnskabelig forskning vedtager den registeransvarlige de nødvendige garantier og beskyttelsesforanstaltninger.
- c) Dataminimering: Kun data, der er strengt nødvendige for at opfylde specifikke formål, bør indsamles.
- d) Rigtighed: De indsamlede oplysninger bør være korrekte og om nødvendigt ajourførte.
- e) Opbevaringsbegrænsning: Personoplysninger skal opbevares på en sådan måde, at det ikke er muligt at identificere de registrerede i et længere tidsrum end det, der er nødvendigt til de formål, hvortil de pågældende personoplysninger behandles. Med henblik herpå kan personoplysninger opbevares i længere tidsrum, hvis personoplysningerne alene behandles til arkivformål i samfundets interesse, til videnskabelige eller historiske forskningsformål eller til statistiske formål i overensstemmelse med artikel 89. I sådanne tilfælde skal der implementeres passende tekniske og organisatoriske foranstaltninger for at sikre de registreredes rettigheder og frihedsrettigheder.

⁴⁸ Anden relevant europæisk lovgivning, der indgår i den retlige ramme for behandling af personoplysninger, omfatter: forordning (EU) 2013/557 om europæiske statistikker for så vidt angår adgang til fortrolige data til videnskabelige formål, forordning (EU) 2018/1725 om beskyttelse af fysiske personer i forbindelse med behandling af personoplysninger i Unionens institutioner, organer, kontorer og agenturer og om fri udveksling af sådanne oplysninger.

⁴⁹ Artikel 2 i Europa-Parlamentets og Rådets forordning (EU) 2016/679.

- f) *Integritet og fortrolighed*: personoplysninger behandles på en måde, der sikrer en tilstrækkelig sikkerhed for dem, herunder beskyttelse mod uautoriseret eller ulovlig behandling og mod hændeligt tab, tilintetgørelse eller beskadigelse, under anvendelse af passende tekniske eller organisatoriske foranstaltninger.
- g) *Ansvarlighed*: Et af de vigtigste principper i forordningen. Det fastsætter, at den registeransvarlige er ansvarlig for databehandlingen og skal kunne påvise, at eventuelle nødvendige foranstaltninger er gennemført.

Dette sæt principper udgør den vigtigste struktur, som den generelle forordning om databeskyttelse er baseret på, og fastlægger en række forpligtelser, som den dataansvarlige og databehandleren skal opfylde.

BEHANDLING AF PERSONOPLYSNINGER I FORBINDELSE MED DEN KONTRAFAKTISKE KONSEKVENSANALYSE: BETINGELSER, BEGRÆNSNINGER OG VIGTIGSTE SPØRGSMÅL

Gennemførelsen af en kontrafaktisk konsekvensanalyse indebærer behandling af store mængder data, herunder personoplysninger. Med henvisning til de retlige rammer, der er beskrevet ovenfor, er det vigtigt at kende de betingelser, begrænsninger og vigtigste spørgsmål, som medlemsstaterne og forvaltningsmyndighederne står over for, når de gennemfører kontrafaktiske konsekvensanalyser.

Behandlingens retsgrundlag og formål

For at kunne behandle personoplysninger skal mindst én af de betingelser for lovlighed, der er anført i artikel 6 i den generelle forordning om databeskyttelse, være opfyldt⁵⁰. Hvis de oplysninger, der skal behandles, derimod er "særlige kategorier af oplysninger", skal der også henvises til forordningens artikel 9.

I forbindelse med kontrafaktiske konsekvensanalyser er offentlighedens interesse i henhold til artikel 6, stk. 1, litra e), i den generelle forordning om databeskyttelse ("udførelse af en opgave i samfundets interesse eller som henhører under offentlig myndighedsudøvelse, som den dataansvarlige har fået pålagt") det mest hensigtsmæssige retsgrundlag for databehandling. Dette retsgrundlag skal fremgå af EU-retten eller medlemsstaternes nationale ret som fastsat i artikel 6, stk. 3, i den generelle forordning om databeskyttelse. "Offentlighedens interesse" repræsenterer klart de forpligtelser, der påhviler forvaltningsmyndighederne som defineret i ESF+-forordningen 2021/1057, artikel 17, stk. 6, om anvendelse af data, der er "tilgængelige i registre eller fra tilsvarende kilder". Desuden hedder det i artikel 4 i forordningen om fælles bestemmelser (EU) 2021/1060, at "[m]edlemsstaterne og Kommissionen må kun behandle personoplysninger, hvis det er nødvendigt for at opfylde deres respektive forpligtelser i henhold til denne forordning, navnlig med henblik på overvågning, rapportering, kommunikation, offentliggørelse, evaluering [...]". Den nationale lovgivning kan give dataansvarlige tilsvarende beføjelser til at forvalte og behandle data i offentlighedens interesse.

Der kan henvises til andre retsgrundlag for en kontrafaktisk konsekvensanalyse under andre betingelser, der er anført i artikel 6, stk. 1, i den generelle forordning om databeskyttelse, navnlig:

- Den registreredes **samtykke** (artikel 6, stk. 1, litra a)). Samtykket kan f.eks. være et passende retsgrundlag, når oplysninger til den kontrafaktiske konsekvensanalyse indsamles ved en undersøgelse, og de registrerede nemt kan give deres samtykke til behandlingen (se artikel 7 i den generelle forordning om databeskyttelse om samtykke). Generelt er samtykke mere komplekst, hvis det ikke planlægges i god tid. Det kan betragtes som et "supplerende" retsgrundlag for en kontrafaktisk konsekvensanalyse, når andre retsgrundlag ikke finder anvendelse.
- Behandlingen er nødvendig for at **overholde en retlig forpligtelse**, som påhviler den dataansvarlige (artikel 6, stk. 1, litra c), jf. artikel 17 i ESF+-forordningen). Dette retsgrundlag skal fastlægges ved lov og kan omfatte private eller offentlige enheder. Det kan f.eks. være, at en privat eller offentlig enhed med ansvar for et datasæt, på grund af specifikke regler skal samarbejde med forvaltningsmyndigheden om den kontrafaktiske konsekvensanalyse.

⁵⁰ Betingelserne er som følger: a) den registrerede har givet samtykke til behandling af sine personoplysninger til et eller flere specifikke formål, b) det er nødvendigt at behandle dem af hensyn til opfyldelse af en kontrakt, som den registrerede er part i, eller af hensyn til gennemførelse af foranstaltninger, der træffes på den registreredes anmodning forud for indgåelse af en kontrakt, c) det er nødvendigt at behandle dem for at overholde en retlig forpligtelse, som påhviler den dataansvarlige, d) behandling er nødvendig for at beskytte den registreredes eller en anden fysisk persons vitale interesser, e) behandling er nødvendig af hensyn til udførelse af en opgave i samfundets interesse eller som henhører under offentlig myndighedsudøvelse, som den dataansvarlige har fået pålagt, f) behandling er nødvendig for, at den dataansvarlige eller en tredjemand kan forfølge en legitim interesse, medmindre den registreredes interesser eller grundlæggende rettigheder og frihedsrettigheder, der kræver beskyttelse af personoplysninger, går forud herfor, navnlig hvis den registrerede er et barn.

Desuden kan den dataansvarlige **anvende de indsamlede oplysninger til andre formål, hvis disse er forenelige med de oprindelige formål**. I denne forbindelse skal den dataansvarlige vurdere betingelserne i artikel 6, stk. 4, og artikel 5, stk. 1, litra b), i den generelle forordning om databeskyttelse. Disse bestemmelser er særlig relevante i forbindelse med behandling af data til videnskabelige forskningsformål eller statistiske formål. De gør det muligt at anvende administrative data til formål, der afviger fra de oprindelige, og som ikke kræver specifikt samtykke til den nye anvendelse, men skal overholde de beskyttelsesregler, der er fastsat i artikel 89, stk. 1, hovedsagelig pseudoanonymisering (se nedenfor). Disse bestemmelser kan f.eks. være relevante i forbindelse med data fra arbejdsløshedsregistre, der skal anvendes i en kontrafaktisk konsekvensanalyse.

På grund af de særlige forhold, der gør sig gældende for kontrafaktiske konsekvensanalyser, kan den dataansvarlige, der indsamler oplysningerne, og den enhed, der udfører forskningen, være forskellige. I dette tilfælde skal der være en betingelse, der legitimerer dataoverførslen og gør det muligt for modtageren at gå videre med den kontrafaktiske konsekvensanalyse.

Anonymisering og pseudonymisering: behandling til statistiske formål

Artikel 89 i den generelle forordning om databeskyttelse fastsætter, at databehandling, der foretages i samfundets interesse eller i forbindelse med videnskabelig forskning eller til statistiske formål, skal være underlagt fornødne garantier for registreredes rettigheder og frihedsrettigheder og navnlig overholde princippet om "minimering". Det betyder, at der skal anvendes pseudonymiseringsteknikker⁵¹.

Hvis formålene kan opfyldes ved efterfølgende behandling, som ikke tillader eller ikke længere tillader identifikation af de registrerede, skal disse formål opfyldes ved at anonymisere oplysningerne og derefter behandle dem i sammenfattet form. I kontrafaktiske konsekvensanalyser aggregeres resultaterne pr. definition, og denne risiko er ikke til stede, medmindre de oprindelige datasæt offentliggøres af videnskabelige årsager. I så fald skal datasæt anonymiseres (se eksemplet i tekstboksen nedenfor).

Datalagring og sikker behandling

Et af de vigtigste aspekter ved databehandling er lagring af data. Lovgivningen specificerer ikke, hvordan data skal lagres, men ovennævnte principper kræver, at lagring og behandling altid er knyttet til formålet med forskningen. Når formålet med behandlingen er opfyldt, og det ikke længere er nødvendigt at opbevare oplysningerne, ophører den. Denne generelle regel skal angives i den databeskyttelseserklæring, der gives til registrerede, når de registreres i administrative datasæt, eller, da det er usandsynligt, at det er muligt at informere alle registrerede personligt, ved alternative metoder til at give oplysningerne (f.eks. ved at offentliggøre en informationsside, der indeholder politikken for beskyttelse af privatlivets fred, om forskningsaktiviteterne på forvaltningsmyndighedens websted).

Det gælder især, at når den kontrafaktiske konsekvensanalyse er gennemført, anses de resultater, der er gjort tilgængelige og aggregeret til statistisk brug, for at være opfyldt. På dette tidspunkt bør data enten anonymiseres (hvis de skal bruges igen på et senere tidspunkt og til andre formål) eller slettes. Love eller forskrifter, der tillader opbevaring, kan være ubegrænset i tid, men kun i henhold til udtrykkelige bestemmelser.

Den anden betingelse, der kræves i den generelle forordning om databeskyttelse, er, at behandlingen skal være omgivet af passende sikkerhedsforanstaltninger. I henhold til forordningens artikel 32 skal den, der foretager behandlingen (den dataansvarlige eller databehandleren), gennemføre passende tekniske og organisatoriske foranstaltninger for at sikre et sikkerhedsniveau, der passer til risikoen. Det betyder, at der ikke foreskrives standardforanstaltninger, men at de skal defineres fra sag til sag med særlig henvisning til de risici, der er forbundet med den enkelte behandling.

⁵¹ Pseudonymisering betyder, at personoplysningerne behandles på en sådan måde, at personoplysningerne ikke længere kan henføres til en bestemt registreret uden brug af supplerende oplysninger. Anonymisering henviser til behandling af personoplysninger på en måde, der gør det umuligt at identificere enkeltpersoner ud fra dem. Der findes en oversigt over pseudonymiseringsteknikker på <https://www.enisa.europa.eu/publications/pseudonymisation-techniques-and-best-practices>.

RETLIGE FORPLIGTELSER

Underretning af registrerede

Den dataansvarliges første opgave er at informere de registrerede. Denne forpligtelse er fastsat i artikel 13 i den generelle forordning om databeskyttelse, når oplysningerne indsamles fra den registrerede, men når dette ikke er tilfældet, er det artikel 14, der gælder. Den registrerede skal informeres om formålene med og metoderne til behandlingen, retsgrundlaget, opbevaringsperioderne og de rettigheder, som den registrerede kan udøve. For kontrafaktiske konsekvensanalyser gives disse oplysninger på tidspunktet for indsamlingen af de data, der senere vil blive anvendt til statistisk og evalueringsmæssig forskning, f.eks. i ESF-overvågning eller datasæt fra arbejdsløshedsregistre. Når denne meddelelse er praktisk umulig, kan der som nævnt ovenfor og navnlig i forbindelse med enkeltpersoner i registre, der anvendes til en kontrafaktisk konsekvensanalyse, anvendes et websted eller andre generelle informationsværktøjer.

Styring af forholdet mellem de forskellige enheder, der deltager i behandlingen

Generelt involverer forskningsaktiviteterne mere end ét organ, og i sådanne tilfælde skal forholdet mellem dem reguleres af særlige ordninger, der kaldes "databehandlingsaftaler"⁵². Disse skal defineres fra sag til sag og afspejle bidragene fra de forskellige interessenter. Eksempler på mulige forhold er:

- *Dataansvarlig — dataansvarlig*: Dette forekommer, når enhederne samarbejder om gennemførelsen af et projekt, selv om de arbejder under forskellige betingelser og med forskellige opgaver og hver især har deres eget særskilte behandlingsformål.
- *Dataansvarlig — databehandler*: Dette forekommer, når en enhed (dataansvarlig) vælger metoderne til og formålene med behandlingen og anvender en anden enhed (databehandler) til at udføre visse behandlingsaktiviteter. I dette tilfælde findes referencen og forpligtelserne i artikel 28 i den generelle forordning om databeskyttelse.
- *Fælles dataansvarlige*: Denne type forhold er fastsat i artikel 26 i den generelle forordning om databeskyttelse, hvori det hedder, at: "*Hvis to eller flere dataansvarlige i fællesskab fastlægger formålene med og hjælpemidlerne til behandling, er de fælles dataansvarlige*". I sådanne tilfælde skal parterne fastlægge deres respektive ansvar for overholdelse af reglerne på en gennemsigtig måde, navnlig med hensyn til de registreredes rettigheder.

Konsekvensanalyse vedrørende databeskyttelse

Artikel 35 i den generelle forordning om databeskyttelse fastsætter, at hvis behandlingen af personoplysninger sandsynligvis vil indebære en høj risiko for de registreredes rettigheder og frihedsrettigheder, foretager den dataansvarlige forud for behandlingen en analyse af de påtænkte behandlingsaktiviteters konsekvenser for beskyttelse af personoplysninger⁵³. En konsekvensanalyse i kontrafaktiske konsekvensanalyser vil kun være vigtig i forbindelse med behandling af særlige kategorier af registrerede i stor målestok. I sådanne tilfælde kan den dataansvarlige i forvaltningen foretage konsekvensanalysen i overensstemmelse med de regler og værktøjer, som den nationale databeskyttelsesmyndighed stiller til rådighed.

FREMTIDSUDSIGTER: FORORDNINGEN OM DATASTYRING

De rammer, der er skitseret ovenfor, afspejler det aktuelle tekniske niveau og de gældende regler i EU. Den generelle forordning om databeskyttelse blev indført specifikt for at muliggøre et gennembrud inden for beskyttelse af personoplysninger og for at skabe et fælles system i hele EU i overensstemmelse med den aktuelle teknologiske udvikling og dagens datasamfund. Andre reformer er under forberedelse og kan få en betydelig indvirkning på indsamlingen af personoplysninger i de kommende år.

Loven om dataforvaltning er stadig på forslagsstadiet i Kommissionen⁵⁴. Det udtrykkelige formål med denne retsakt er at fremme tilgængeligheden af brugbare data ved at styrke tilliden til dataformidlere og forbedre dataudvekslingsmekanismerne i hele EU. Personoplysninger vil sandsynligvis være omfattet af den nye standard, og deres anvendelse til statistiske og forskningsmæssige formål kan blive mere vidtrækkende takket være indførelsen af en ny funktion: formidleren til deling af personoplysninger. En sådan

⁵² Dette er f.eks. tilfældet, når den dataansvarlige og de, der faktisk udfører forskningen, ikke er de samme. I dette tilfælde er det nødvendigt at finde en aftale, der regulerer forholdet mellem parterne med hensyn til at få adgang til dataene og udføre forskningen.

⁵³ Der findes flere oplysninger i retningslinjerne fra Databeskyttelsesrådet <https://ec.europa.eu/newsroom/article29/items/611236>.

⁵⁴ Kommissionen vedtog forslaget til forordningen om datastyring i slutningen af 2020. Den aktuelle status for forordningen kan ses her: <https://eur-lex.europa.eu/legal-content/DA/HIS/?uri=CELEX:52020PC0767>.

foranstaltning kan bidrage til at skabe gunstigere rammer for kontrafaktiske konsekvensanalyser og dermed lette adgangen til en stor mængde oplysninger i stor målestok.

Det er klart, at reglerne i den generelle forordning om databeskyttelse ikke udelukker gennemførelsen af en kontrafaktisk konsekvensanalyse. Som følge heraf skal nogle grundlæggende procedurmæssige og operationelle skridt respekteres i alle medlemsstater for at gøre den kontrafaktiske konsekvensanalyse sikker. I et normalt tilfælde, hvor overvågningsdata (behandlede personer) kombineres med data fra et offentligt register (kontrolgruppe), og en ekstern evaluator udfører den kontrafaktiske konsekvensanalyse, er de vigtigste skridt f.eks. følgende:

De vigtigste trin til overholdelse af den generelle forordning om databeskyttelse

1. Forvaltningsmyndigheden indgår en aftale med de forvaltninger, der er ansvarlige for de data, der er nødvendige for at identificere kontrolgruppen og analysere behandlings- og kontrolgrupper (f.eks. arbejdsløshedsregistre, skatteregister osv.). Forvaltningsmyndigheden kontrollerer, at anvendelsen af disse data er i overensstemmelse med artikel 6, stk. 1, litra e), eller artikel 6, stk. 4, og artikel 5, stk. 1, litra b), i den generelle forordning om databeskyttelse, hvis der ikke er indhentet et udtrykkeligt samtykke.
2. Forvaltningsmyndigheden indgår aftaler med de andre enheder (dataejere og evaluator) om at regulere informationsstrømmen og det gensidige ansvar i overensstemmelse med reglerne i den generelle forordning om databeskyttelse. Tjenesteydelseskontrakten mellem evaluatoren og forvaltningsmyndigheden skal indeholde en særlig bestemmelse om databeskyttelse. For så vidt angår andre forvaltninger, kan et aftalememorandum eller specifikke nationale procedurer regulere databeskyttelse i forbindelse med gennemførelsen af kontrafaktiske konsekvensanalyser.
3. På grundlag af ovennævnte aftale modtager forvaltningsmyndigheden data i pseudoanonymiseret form fra dataejeren og overfører data til behandling af behandlings- og kontrolgrupper til evaluatoren i overensstemmelse med forordningen⁵⁵.
4. Opbevaringen af personoplysninger om behandlings- og kontrolgrupper opfylder kun reglerne for opbevaring af data, så længe forskningsprojektet varer, og i overensstemmelse med de grundlæggende sikkerhedsregler som fastsat i den generelle forordning om databeskyttelse. Disse regler overholdes af alle de enheder, der er involveret i den kontrafaktiske konsekvensanalyse.

National praksis er generelt i overensstemmelse med den generelle forordning om databeskyttelse, men forskellig fra medlemsstat til medlemsstat, og evaluatorene rapporterer, at nationale databeskyttelsesregler stadig udgør alvorlige hindringer for anvendelsen af mikrodata. Dette kan skyldes den tid, der er nødvendig for at tilpasse nationale regler og vaner til den seneste generelle forordning om databeskyttelse, eller forskellige fortolkninger af forordningen i forskellige nationale forvaltninger, overregulering, der gør forordningen strengere i visse lande, eller andre fejlfortolkninger. I nogle lande vil det være nødvendigt med et specifikt initiativ fra forvaltningsmyndighederne eller de nationale

⁵⁵ Hvis behandlings- og kontrolgrupper analyseres ved hjælp af en undersøgelse, skal denne omfatte samtykke til at anvende svarpersonernes data til forskningsformål.

myndigheder for at overvinde disse hindringer i ånden i den generelle forordning om databeskyttelse og ESF+-forordningen. Den italienske Veneto-regionens håndtering af nogle af de mest almindelige problemer er beskrevet i tekstboksen nedenfor. En god praksis er at kontakte den nationale databeskyttelsesmyndighed for at drøfte de foreslåede ordninger, inden de færdiggøres.

Tekstboks 9 Beskyttelse og udveksling af data

Et godt eksempel på at få adgang til anonymiserede personoplysninger på relativt kort tid kan findes i en kontrafaktisk evaluering i "Pilot and feasibility study on the sustainability and effectiveness of results for European Social Fund participants using counterfactual impact evaluations"⁵⁶, der blev gennemført for GD EMPL i 2019.

Evalueringen, en ud af fire cases i undersøgelsen, fokuserede på ESF-interventioner (IP 9.i), der blev gennemført i Veneto-regionen i 2015/2016, og som var rettet mod langtidsledige. Overvågningsdataene om deltagere og typer af foranstaltninger blev kombineret med data om beskæftigelsesperioder før og efter ESF-interventionerne (kaldet "Comunicazioni obbligatorie"). Processen varede ca. to måneder (fra udgangen af juli til udgangen af september) fra det første møde om datakrav med forvaltningsmyndigheden og den institution, der lå inde med dataene, det regionale institut i Veneto Lavoro. Veneto Lavoro var ansvarlig for anonymisering af dataene, og alle de leverede datasæt kunne let kædes sammen med hinanden⁵⁷ ved hjælp af en unik fælles identifikator, som Veneto Lavoro leverede.

Selv om de involverede organisationer ikke direkte bestilte evalueringen, skyldes de positive erfaringer to hovedfaktorer: Et stærkt (om end uformelt) samarbejde mellem forvaltningsmyndighedens regionale kontorer og det regionale institut i Veneto Lavoro, som er nødvendigt for at løse databeskyttelsesproblemer, og en database (Mercurio), der forvaltes af Veneto Lavoro og er et offentligt tilgængeligt register til forskningsformål⁵⁸.

Mercurio er en statistisk database med alle oplysninger, der er registreret af arbejdsformidlingskontorer i Veneto-regionen vedrørende beskæftigelses- og arbejdsløshedsperioder i befolkningen. Den indeholder desuden detaljerede oplysninger om alle registrerede arbejdstagere og virksomheder. På tidspunktet for analysen indeholdt Mercurio oplysninger om mere end 4 millioner arbejdstagere, 17,6 millioner arbejdsløshedsændelser og 4 millioner arbejdsløshedsændelser⁵⁹.

Datasættet "opdateres og renses regelmæssigt", og dette gør det relativt ligetil at matche det med overvågningsdataene fra forvaltningsmyndigheden. En mulig begrænsning er, at det offentligt tilgængelige register ofte ikke opdateres nok (f.eks. blev den seneste version, der er tilgængelig, opdateret i december 2020), da det kræver en enorm indsats for at sikre en god datakvalitet⁶⁰. Erfaringerne viser, at forvaltningsmyndigheden bør træffe foranstaltninger på forhånd, samtidig med at der udarbejdes evalueringsplaner for: mulige former for samarbejde med eksterne organer, der forvalter de nødvendige data til en kontrafaktisk øvelse, løsninger på potentielle juridiske hindringer, samt afklaring af, hvordan man kan få adgang til, forvalte og levere administrative data i anonymiseret form.

2.3.2. Hvordan identificeres den "behandlede" gruppe?

For at kunne gennemføre en kontrafaktisk konsekvensanalyse er det nødvendigt med en klar definition af, hvad det betyder at blive behandlet eller

⁵⁶ [Ismeri Europa — Ecorys — Institute for Employment Studies, 2019.](#)

⁵⁷ Der blev leveret endnu et datasæt med start- og slutdatoerne for hver arbejdsdygtighedserklæring underskrevet af alle behandlede og ikkebehandlede personer.

⁵⁸ <https://www.venetolavoro.it/public-use-file> indeholder en beskrivelse af Mercurio og reglerne for adgang til den.

⁵⁹ Følgende relevante variabler blev udtrukket af de administrative data: start- og slutdatoer for alle ansættelses- og arbejdsløshedsperioder, kontrakttype for hver ansættelsesperiode (fast, tidsbegrænset, praktik), tidsbegrænsede kontrakters forventede varighed, sociodemografiske oplysninger (køn, fødselsdato, uddannelse, statsborgerskab).

⁶⁰ Det betyder, at det er nødvendigt at arbejde på at sikre mere opdaterede oplysninger end dem, der er i den senest tilgængelige version af Mercurio, og at der specifikt er behov for yderligere arbejde for at integrere og rense de mest ajourførte oplysninger om ansættelsesperioder.

at have deltaget i interventionen. Når der er opnået en klar forståelse af, hvornår en person eller virksomhed anses for at være blevet behandlet, er det desuden vigtigt, at de kan identificeres. Her introduceres de vigtigste spørgsmål vedrørende definitionen af behandlings- og kontrolgrupper, og kapitel 3 indeholder en mere detaljeret redegørelse for deres metodologiske karakteristika.

Ved første øjekast kan det være ligetil at definere deltagelsen. Der er dog en række spørgsmål, som måske ikke er umiddelbart åbenlyse, men som er afgørende og kræver grundige overvejelser. Anses praktikanter i en uddannelsesordning, der dropper ud af interventionen, f.eks. for at være blevet behandlet? Hvor mange gange skal deltagerne på et kursus f.eks. møde op til et kursus, før de betragtes som deltagere? Forventningsrelaterede virkninger skal også overvejes. I forventning om, at de skal deltage i en intervention, vil nogle modtagere af sociale sikringsydelser måske opgive deres ydelser for at undgå aktiveringsforanstaltninger. Behandles disse personer, selv om de f.eks. aldrig fysisk møder op til deres aftaler hos den offentlige arbejdsformidling?

Definition af behandlingsgruppen

Ved definitionen af "behandlingsgruppe" sondres der mellem "behandlingsintention" og "behandling af den behandlede". Ud fra et politisk synspunkt er det centrale spørgsmål normalt, om der fokuseres på virkningerne af at få mulighed for at deltage i en intervention, eller virkningerne af faktisk at deltage. I førstnævnte tilfælde kan de personer, der tilbydes en intervention, deltage eller lade være. I sidstnævnte tilfælde omfatter den behandlede gruppe, når interessen ligger i behandlingens virkning på den behandlede gruppe, kun de personer, der deltager⁶¹.

Behandlingsintention eller behandling af den behandlede

Som udgangspunkt antager de politiske beslutningstagere ofte, at de er interesserede i at bestemme nettoeffekten af behandlingen for dem, der deltager. Ved et nærmere kig er spørgsmålene imidlertid måske mindre klare. Hvis dem, der tilbydes behandling, kan identificeres, kan det ud fra et politisk synspunkt være mere nyttigt at definere dem som den "behandlede" gruppe. Dette er navnlig tilfældet, hvor deltagelse i en intervention ikke er obligatorisk. De politiske beslutningstagere kan ikke tvinge dem, der har fået tilbud om en intervention, til at deltage, så det relevante spørgsmål er: Hvad er indvirkningen efterfølgende på beskæftigelse og indkomst for dem, der fik mulighed for at deltage i et praktikforløb?

Tilbud eller faktisk behandling

For at kunne vurdere virkningerne af tilbuddet om behandling på en række resultater skal dem, der modtager tilbuddet, kunne identificeres. I mange tilfælde kan dette være vanskeligt.

Hvor finder man passende data?

Når der er truffet afgørelse om definitionen af, hvem der behandles, og hvad der udgør behandling, er det vigtigt at overveje, hvordan de personer, der behandles, vil blive identificeret i forbindelse med evalueringen. Dette betyder altid, at der skal findes en datakilde, hvorfra behandlede enheder, uanset om de er personer eller virksomheder, enten kan opregnes fuldt ud eller udtages til stikprøve. Disse optegnelser stammer normalt fra ESF-

Datakilder for behandlede personer

⁶¹ Hvis deltagelsen i en intervention er obligatorisk, er der i det væsentlige ingen forskel mellem disse to — alle, der tilbydes behandling, skal deltage. I de fleste tilfælde er interventionerne dog ikke obligatoriske (som det antages gennem hele denne vejledning).

overvågningssystemerne og — hvis de foreligger — fra yderligere fortegnelser oprettet for den pågældende intervention.

På grund af kravene til overvågning og rapportering fra ESF (og ESF+) skal modtagerorganisationerne registrere antallet af og visse personlige karakteristika for dem, der modtager tjenester gennem en intervention. Med henblik på kontrafaktisk konsekvensanalyse skal interventionerne gå videre og give mikrodata om dem, der har deltaget. Evaluatorene vil kræve registrering for hver behandlet enhed (virksomhed eller person) med data om deres vigtigste karakteristika (køn, alder, uddannelsesniveau osv.). Disse oplysninger kan anonymiseres/pseudonymiseres af hensyn til databeskyttelse, men når der er behov for en undersøgelse for at gennemføre den kontrafaktiske konsekvensanalyse, vil det være nødvendigt at fastslå, om betingelserne i den generelle forordning om databeskyttelse gør det muligt at overføre disse enheders identitet (navne, adresser, telefonnumre osv.), så de kan indgå i stikprøven. Der kræves også entydige identifikatorer for hver enkelt enhed for at lette sammenkædningen af data på tværs af datakilder.

2.3.3. Faktorer, der skal tages i betragtning ved identificering af en kontrolgruppe

For at få et skøn over den kontrafaktiske situation vil det normalt være nødvendigt at identificere en kontrolgruppe. Valget af en kontrolgruppe vil normalt være begrænset af, om interventionen er obligatorisk for deltagerne eller ej, samt af, om interventionen gennemføres universelt inden for en jurisdiktion eller er begrænset til et bestemt område eller over et begrænset tidsrum. Ved valget af en passende kontrolgruppe skal der tages hensyn til tre aspekter: 1) det analytiske, 2) det politikrelaterede og 3) det praktiske.

Tre aspekter, der skal tages i betragtning ved definitionen af en kontrolgruppe

Definition af en kontrolgruppe ud fra et analytisk perspektiv

Formålet med en kontrafaktisk konsekvensanalyse er at opnå objektive skøn over en interventions indvirkning på en række resultater. Til det formål er der behov for skøn over kontrafaktiske resultater. Skøn over kontrafaktiske resultater indhentes fra en kontrolgruppe (se afsnit 1.1). Som vist i figur 6 og 7 anslås virkningen ved at trække et skøn over det kontrafaktiske resultat fra et observeret resultat for den "behandlede" gruppe. Hvorvidt der er bias i virkningen, afhænger af, i hvilket omfang det kontrafaktiske resultat, der er beregnet for kontrolgruppen, repræsenterer det resultat, der alt andet lige ville være fremkommet for den behandlede gruppe, hvis de ikke var blevet behandlet.

For at finde en tilsvarende kontrolgruppe uden randomisering skal den svare til behandlingsgruppen i alle vigtige henseender, både med hensyn til observerbare og ikkeobserverbare dimensioner. Dette gælder for alle kvasiexperimentelle metoder og er dermed en nødvendig betingelse for, at alle kontrafaktiske konsekvensanalyser ikke anvender en randomiseret tilgang.

Da næsten alle ESF-interventioner enten er a) frivillige (målgruppen er ikke tvunget til at deltage i en intervention) og/eller b) begrænset på anden måde — hvis der er tale om pilotinterventioner eller instrumenter, der er begrænset til en bestemt region eller jurisdiktion, vil evaluatorene have en pulje af enheder, der kan udvælges til brug som kontrol. En vis sortering i denne

Muligheder for udvælgelse af kontrolgrupper

potentielle pulje vil være nødvendig for at finjustere det endelige valg af kontroller, så de passer godt til deltagerne (den behandlede gruppe). I mange tilfælde er der potentielt fire muligheder:⁶²

- **Placering** — Kontroller, der svarer til dem, der deltager i en intervention, men som befinder sig i områder i medlemsstaten, hvor interventionen ikke er tilgængelig (hvis sådanne områder findes). Difference-in-differences-metoden er ofte den foretrukne tilgang i tilfælde, hvor sådanne kontrolgrupper og de rette data er tilgængelige. Populationer andre steder kan være meget ens, og sådanne grupper vil ikke have haft mulighed for at deltage i interventionen eller afvise tilbuddet om at deltage, og derfor vil denne vigtige kilde til potentiel bias ikke eksistere. Populationer andre steder vil imidlertid være underlagt forskellige arbejdsmarkedsforhold. Difference-in-differences-teknikkerne for sådanne variationer fungerer ret godt, og forskellene i de lokale arbejdsmarkedsforhold har tendens til at være rimeligt uændrede over tid. Det er imidlertid mindre tilrådeligt at udtage kontrolprøver fra forskellige lokale arbejdsmarkeder i tilfælde, hvor matchning anvendes til at vurdere virkningerne. Det er blevet påvist, at den bias, der er forbundet med at udvælge kontrolprøver fra forskellige arbejdsmarkeder, kan være større end selektionsbias⁶³.
- **Tid** — Kontroller, der svarer til deltagerne, men som observeres på forskellige tidspunkter, enten før eller efter interventionen. Kontrolgrupper, der udvælges på denne måde, er ofte nødvendige, når en intervention er universel og obligatorisk — dvs. hvor alle medlemmer af målgruppen er forpligtet til at deltage, og programmet gennemføres i en hel jurisdiktion. Kontrolgrupper, der er dannet på en sådan måde, har en betydelig ulempe, nemlig at deres resultater vil blive målt på forskellige tidspunkter i forhold til behandlingsgruppens, således at de er sårbare over for konjunkturudsving, ændringer i sammensætningen og skiftende makroøkonomiske tendenser, hvilket kan stå i vejen for, at der kan findes et kontrafaktisk resultat uden bias. En sådan kontrolgruppe bør kun overvejes, hvis der er begrænset variation i resultaterne over tid, og hvis der ikke er en samtidig kontrolgruppe til rådighed.
- **Støtteberettigelse** — Her udvælges kontrolgruppen fra grupper på samme sted og tidspunkt, men blandt kandidater, som af en eller anden grund ikke var berettiget til at deltage. Sådanne kontroller søges ofte, når en intervention er universel, deltagelsen er høj, eller deltagelsen er obligatorisk, og hvor der er klare regler for udvælgelse, således at f.eks. dem, der "ikke er berettiget til at deltage", udgør en potentiel kilde til kontrol. Målet er at finde grupper, der ligner dem, der behandles, men som af velkendte og faste årsager (som kan kvantificeres i dataene) ikke var berettiget til behandling. Adgangen til interventioner under ESF+ er ofte baseret på særskilte udvælgelseskriterier, som let kan måles og ikke kan manipuleres (deltagernes specifikke alder, arbejdsløshedsstatus osv.). Ved udvælgelsen af kontroller kan disse tærskler derfor anvendes som diskriminerende faktorer, som behandlede og kontroller fordeles på.
- **Valg/bevidsthed** — I det væsentlige er både behandlings- og kontrolgruppen (snarere end blot behandlingsgruppen) underlagt

⁶² Dette afsnit er baseret på [Card, D., Ibarrran, P. og Villa, J.M., 2011.](#)

⁶³ Se [Heckman, J.J., Ichimura, H., Smith, J. og Todd, P., 1998.](#)

udvælgelsesprocesser baseret på valg, der er motiveret af potentielt uobserverede faktorer⁶⁴. Kontroller kan udvælgelse blandt dem, der var berettiget til at deltage, men som ikke deltog. Fordelen ved dette er, at de normalt kommer fra det samme arbejdsmarked som dem, der blev behandlet. Sådanne kontroller bør overvejes nøje, navnlig når der anvendes en tilsvarende udformning af den kontrafaktiske konsekvensanalyse, og hvor der findes omfattende data til udvælgelsesbeslutningen. Under andre omstændigheder, f.eks. hvor difference-in-differences-metoden anvendes, vil kontrolgrupper baseret på valg/bevidsthed være mindre attraktive.

Endnu et punkt er værd at bemærke. Hvis der findes målinger af resultaterne før behandling for både behandlings- og kontrolgrupperne, er det vigtigt at undersøge tendenser inden interventionen i resultatmålinger for begge grupper. Ved at kontrollere den såkaldte antagelse om "en fælles tendens" tages der fat på problemet med midlertidige fald i beskæftigelsesfrekvensen og indkomsten forud for intervention for nogle af dem, der er berettiget til aktive arbejdsmarkedspolitikker (ellers ville de ikke være støtteberettigede — det såkaldte "Ashenfelter's Dip"-problem). Evaluatoren søger at finde lignende tidstendenser med hensyn til resultatmålinger for både behandlings- og kontrolgrupper, således at genopretningen efter kortvarige fald i beskæftigelsen eller indkomsten ikke forveksles med de langsigtede relative gevinster, som den kontrafaktiske konsekvensanalyse forsøger at afdække.

Udvælgelsen af passende kontrolgrupper er en teknisk og metodologisk kompleks øvelse. I forbindelse med udviklingen af evalueringsordninger anbefales det, at embedsmændene gør sig bekendt med de vigtigste begreber og tager skridt til at identificere potentielle kontroller på et tidligt tidspunkt. Det er vigtigt, at de, der iværksætter en evaluering, inddrager eksperter tidligt i designfasen for at modtage støtte og rådgivning.

Hvad er de relevante politikrelaterede overvejelser?

Udvælgelsen af en passende kontrolgruppe er ikke blot en teknisk eller analytisk proces. Selv om analytiske aspekter i forbindelse med identifikation af passende kontroller er af grundlæggende betydning, er det også vigtigt, at en kontrolgruppe udgør et relevant alternativ til interventionen, når den betragtes ud fra et politisk synspunkt.

Kontrafaktiske konsekvensanalyser kan antage forskellige former: De kan f.eks. sammenligne resultaterne af en behandlingsgruppe eller en række behandlingsgrupper med en kontrolgruppe, der ikke modtager behandling. De kan også sammenligne en behandling med en anden uden en kontrolgruppe uden behandling. Valget af kontrolgruppe vil være baseret på, hvilken type sammenligning der er mest politisk relevant, og om det endda er muligt at finde en kontrolgruppe uden behandling. Tekstboks 10 nedenfor viser et eksempel på en sammenligning af en behandling med en anden uden en kontrolgruppe uden behandling, hvor formålet er at vurdere, om man skal fortsætte én intervention i stedet for en anden. Det skal også bemærkes, at sammenligninger af et program med et andet uden at have en kontrolgruppe uden behandling kan give anledning til tvetydighed (dette drøftes yderligere i tekstboks 11).

Analyse af tendenser forud for intervention

Definition af et alternativ til en intervention

Sammenligning af behandling med ingen behandling eller med et alternativ

⁶⁴ Det er det, som [Card, D., Ibarrran, P. og Villa, J.M., 2011](#) kalder "tosidet selektionsbias".

Tekstboks 10 Politiske spørgsmål vedrørende et praktikforløb

Hvis vi tager et eksempel, hvor den politiske beslutningstager indfører en ny praktikintervention, som skal finansieres over ESF — lad os kalde den intervention A. Hvis vi så antager, at medlemsstaten allerede har en praktikordning (B), som er rettet mod de samme personer, men som finansieres med nationale midler. Her kan et politisk spørgsmål være: Er beskæftigelses- og indkomstniveau for deltagerne i intervention A større end for deltagerne i intervention B efter deltagelse? Og giver intervention A dermed mere valuta for pengene? Hvis indkomsten er højere for deltagerne i intervention A, er det indlysende politiske svar at indstille intervention B til fordel for intervention A, hvis leveringen heraf også viser sig at være omkostningseffektiv.

Tekstboks 11 Fortolkning af nettoeffekterne

Måske konkluderes det i en undersøgelse, at der ikke er nogen indkomstforskel mellem deltagerne i intervention A og dem i intervention B. Det er muligvis ikke klart, hvordan de politiske beslutningstagere skal forholde sig til det, hvis intervention B f.eks. var meget effektiv i forhold til dem, der ikke modtog behandling. Dette ville betyde, at begge interventioner er yderst effektive. I nogle tilfælde kan det imidlertid være, at der ikke er noget bevis for, at intervention B er effektiv i forhold til ingen behandling. Alternativt kan intervention A og B begge være ineffektive, selv om den ene intervention kan forekomme relativt mere effektiv end den anden. I tilfælde, hvor visse befolkningsgrupper kan være omfattet af mere end én intervention, kan det stadig være mere informativt at finde en passende gruppe af ubehandlede enheder, der kan fungere som sammenligningsgrundlag.

Bemærk, at difference-in-differences-metoden ikke kan anvendes til at sammenligne flere behandlinger, hvis der ikke findes en kontrolgruppe uden behandling.

Hvilke praktiske hensyn skal der tages ved udvælgelsen af en kontrolgruppe?

Ud over analytiske og politiske overvejelser skal der tages hensyn til de praktiske aspekter af udvælgelsen af kontrolgrupper. Udvalgelse eller stikprøveudtagning af enheder (personer eller virksomheder), der skal fungere som kontroller, kræver, at der findes en passende stikprøveramme. Stikprøverammerne bør desuden indeholde individuelle enheder, der opfylder de analytiske og politikrelaterede krav. Præcis hvordan dette bedst gøres, vil variere fra den ene evaluering til den anden, afhængigt af den specifikke kontekst for den intervention, der testes.

I mange tilfælde anvendes der ofte to datakilder til at identificere egnede kontrolgrupper. De kræver begge, at behandlingsgruppens identitet er kendt.

Folkeregistre af forskellig art kan bruges til at finde kontroller. Hvis f.eks. en aktiv arbejdsmarkedsintervention er rettet mod 18-24-årige, der modtager arbejdsløshedsunderstøttelse, kan fortegnelserne over deres understøttelse bruges til at identificere målgruppen. Hvis den behandlede gruppe er kendt og kan matches med data om arbejdsløshedsunderstøttelse, kan de 18-24-årige, der ikke er blevet behandlet, og dermed er potentielle kontroller, desuden findes. Alternativt kan det antages, at en intervention er rettet mod små og mellemstore virksomheder. Nationale virksomhedsregistre (hvis sådanne findes) kan anvendes til at definere målpopulationen og med tilgængelige oplysninger om, hvilke virksomheder der behandles, kan potentielle kontrolgrupper findes.

Folkeregistre og virksomheders skatteregnskaber

Fortegnelser over ansøgere kan anvendes, hvis interventionen ikke er universel. Det gælder f.eks. hvis ikke alle, der ansøger om et praktikforløb, accepteres (en valg-/bevidsthedsgruppe). Tilsvarende vil ikke alle de virksomheder, der ansøger om finansiering, blive udvalgt, og virksomheder, der ikke er godkendt til uddannelse eller finansiering, kan i nogle tilfælde anvendes som kontrol (se tidligere diskussion i dette afsnit om den forsigtighed, der bør udvises ved udvælgelsen af kontrolgrupper under disse omstændigheder).

Fortegnelser over ansøgere

2.3.4. Hvilke typer spørgsmål vedrørende data skal overvejes i evalueringsordningen?

Hvilke typer data er nødvendige, og hvordan vil de blive indsamlet?

Som det er blevet bemærket, kræver kontrafaktiske konsekvensanalyser normalt adgang til betydelige mængder mikrodata (i nogle tilfælde kan der anvendes grupperede data — f.eks. regionale data). Disse data skal indsamles, sammenholdes og dokumenteres. Data fra forskellige kilder skal sammenkædes på grundlag af fælles identifikationsfelter. De skal opbevares og overføres sikkert mellem dem, der iværksætter og foretager den kontrafaktiske konsekvensanalyse. Samtidig skal der udarbejdes analytiske datasæt ud fra disse datakilder for at gøre det lettere at vurdere virkningerne.

Styring af datakilder

I forbindelse med udarbejdelsen af en evalueringsordning er det vigtigt at overveje følgende datarelaterede spørgsmål:

En grundlæggende tjekliste for datastyring

Tilgængelighed

- Hvilke kilder kan anvendes til at indhente forskellige typer grupperede data/mikrodata?
- Er de nødvendige individuelle data tilgængelige? Gælder dette også for "særlige kategorier" af data (om nødvendigt)?
- Når en undersøgelse er nødvendig, identificeres mål- og kontrolgrupper på en måde, der gør det muligt at følge op på dem gennem spørgeskemaundersøgelser — er kontaktoplysninger tilgængelige og ajourførte?

Konsekvens

- Findes der én enkelt datakilde, eller er det nødvendigt at sammenkæde datakilder (f.eks. statistikker over arbejdsløshed, sociale ydelser, social sikring, virksomheds-/etableringsdata osv.)?
- Er kilderne i overensstemmelse med hinanden? Kan enkeltpersoner identificeres i dem på et ensartet grundlag på tværs af kilderne?

Tilgængelighed

- Er det muligt at få adgang til nationale datakilder om individuelle karriereforløb, indkomst eller sociale ydelser for at sammenligne ESF-deltagere med en potentiel kontrolgruppe?

Aftaler

- Er der indgået specifikke aftaler om datatilgængelighed mellem forvaltningsmyndigheden og andre dataejere? Hvilke juridiske eller organisatoriske hindringer skal drøftes?
- Hvem får ansvaret for at forhandle om adgang til og tilladelse til brug af data?

Opbevaring

- Hvor vil dataene blive opbevaret? Hvilke it-systemer og -infrastrukturer vil der være brug for?
- Hvilke skridt vil der blive taget for at sikre, at data opbevares sikkert, og at adgangen er forbeholdt dem, der har brug for dataene med henblik på evaluering?
- Hvordan anonymiseres oplysningerne? Er det muligt at følge enkeltpersoner over tid og gennem sammenkædede datakilder?

Overholdelse af databeskyttelse

- Er tidligere løsninger i overensstemmelse med de grundlæggende regler i den generelle forordning om databeskyttelse og ESF+-forordningen? Er eventuelle hindringer for datatilgængelighed begrundet i anvendelsen af forskellige retsgrundlag i henhold til artikel 6, litra a), c) og e), i den generelle forordning om databeskyttelse og artikel 17 i ESF+-forordningen og den mulige anvendelse af anonymiserede data?

Hvordan behandles data?

I mange tilfælde vil kontrafaktiske konsekvensanalyser kræve mikrodata — dvs. data, der indeholder observationer om individuelle enheder (sædvanligvis personer eller virksomheder) i både behandlings- og kontrolgrupperne (undertiden kan der anvendes grupperede data, f.eks. regionale data eller data for de enkelte offentlige arbejdsformidlinger). Vi har sondret mellem tre hovedtyper af data, der kræves: a) behandlings- og kontrolgruppefortegnelser, b) resultatfortegnelser og c) kontekstdata (data, der anvendes til at kontrollere, om der er væsentlige potentielle forskelle mellem behandlings- og kontrolgrupper). Disse data kan stamme fra samme eller forskellige kilder. Kilderne skal struktureres, så de danner analytiske datasæt (eller analytiske stikprøver), der anvendes til at estimere virkningerne. Denne strukturering vil i mange tilfælde indebære sammenkædning af fortegnelser over personer eller virksomheder på tværs af kilderne. En sådan sammenkædning kræver enten individuelle identifikatorer (f.eks. individuelle socialsikringsnumre), som gør det muligt at afstemme en persons data (f.eks. i skatteoplysninger) med deltagerfortegnelser, eller at der er tilstrækkelige data til at sammenkæde fortegnelser på tværs af kilder (f.eks. skal navn og fødselsdato være tilgængelige på tværs af kilder). Det er vigtigt at overveje, hvilke datakilder der vil blive anvendt til den kontrafaktiske konsekvensanalyse, der planlægges, men også om det vil være muligt at sammenkæde fortegnelser på tværs af kilderne.

Sammenkædning af mikrodata på tværs af kilder

2.3.5. Hvad er de vigtigste begrænsninger i forbindelse med analyse af data og resultater?

Som nævnt ovenfor bestemmes virkningerne i kontrafaktiske konsekvensanalyser normalt ved at sammenligne resultaterne i behandlingsgruppen med resultaterne i kontrolgruppen, og denne grundlæggende sammenligning er i det væsentlige en del af metoden til kontrafaktiske konsekvensanalyser. Forskellen mellem behandlings- og kontrolgrupper kaldes virkningen eller nettoeffekten af interventionen. Den nøjagtige måde, hvorpå virkningerne vurderes, vil dog afhænge af det valgte forskningsdesign.

Ved planlægningen af en kontrafaktisk konsekvensanalyse er det vigtigt at overveje, om interventionen er stor nok og har sandsynlighed for at skabe virkninger, der kan påvises statistisk. Metoderne til kontrafaktiske konsekvensanalyser er faktisk baseret på princippet om, at mål- og kontrolgrupper er en stikprøve af hele de potentielt behandlede populationer og kontrolpopulationer, og at antallet af individer i de to grupper skal være statistisk signifikant for at opnå signifikante resultater.

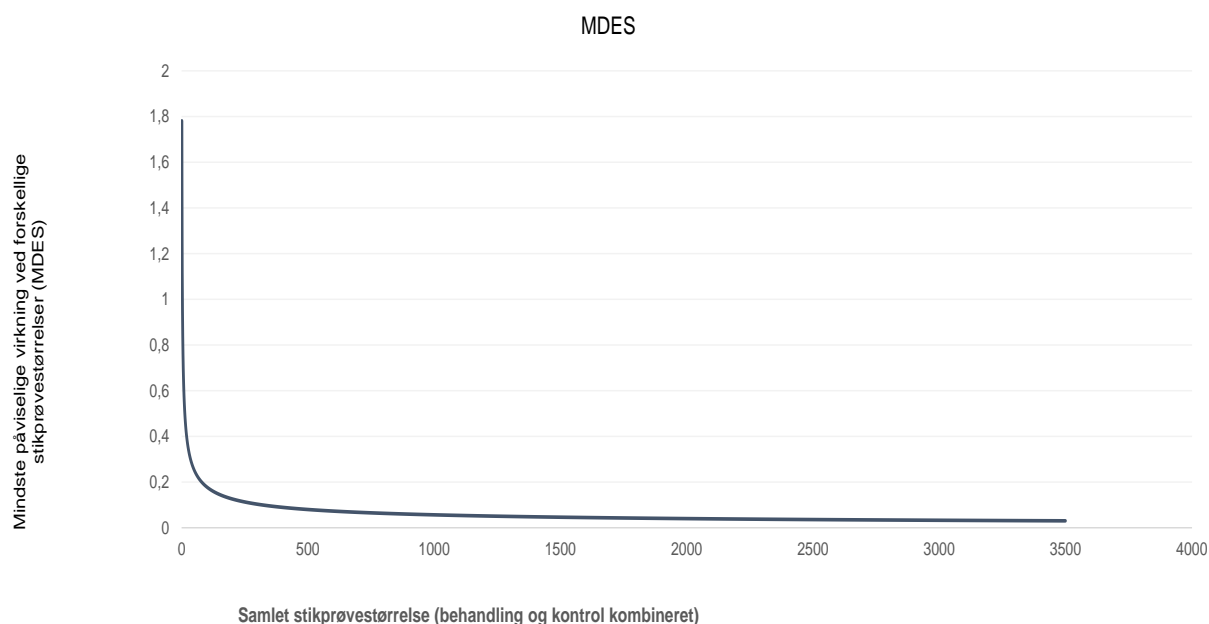
Ved vurderingen af, om en stikprøve er tilstrækkelig stor til at kunne analyseres, er et nyttigt begreb "den mindste påviselige virkning"⁶⁵. Kort sagt er den mindste påviselige virkning den mindste reelle virkning, som en stikprøvestørrelse kan påvise ved standardniveauer for statistisk konfidens og effekt⁶⁶.

*Vurdering af
stikprøvestørrelse
og effekt*

Figur 4 nedenfor viser, hvordan den mindste påviselige virkning varierer med den samlede stikprøvestørrelse (det samlede antal stikprøver i behandlings- og kontrolgrupper) med et statistisk konfidensniveau på 95 % og ved et statistisk effektniveau på 80 %. Fra venstre mod højre falder den mindste påviselige virkning hurtigt, når stikprøvestørrelsen nærmer sig 500 (250 behandlingsenheder og 250 kontroller). Udformningen af den kontrafaktiske konsekvensanalyse er med andre ord mere præcis og i stand til at påvise mindre virkninger, efterhånden som den samlede stikprøvestørrelse øges.

Figur 4 Mindste påviselige virkning ved forskellige stikprøvestørrelser

Bemærk: 95 % statistisk signifikans og 80 % statistisk effekt. Figuren er baseret på en randomiseret metode



⁶⁵ Bloom H. S., 1995, indeholder praktisk vejledning i, hvordan man beregner de mindste påviselige virkninger til eksperimentelle forsøg. I forbindelse med kvasiexperimentelle tilgange vil sådanne beregninger kræve en justering. Generelt kræver kvasiexperimentelle tilgange større stikprøvestørrelser i forhold til et eksperimentelt forsøg.

⁶⁶ Den statistiske konfidens er et statistisk mål for estimeringsprocedurens pålidelighed, mens den statistiske effekt er sandsynligheden for, at en test afdækker en reel effekt. Disse to parametre afhænger også af stikprøvestørrelsen, og de kan omvendt bruges til at beregne stikprøvestørrelsen i forhold til passende signifikansniveauer. Generelt er den statistiske konfidens 95 %, og den statistiske effekt er 80 %.

Ved planlægningen af en kontrafaktisk konsekvensanalyse er det nyttigt at anslå stikprøvernes sandsynlige størrelse med henblik på at definere behandlings- og kontrolgrupper. Dette skøn er baseret på prognoser over det antal enheder, der vil blive behandlet, udformningen af den sandsynlige metode, der vil blive anvendt i den kontrafaktiske konsekvensanalyse, og størrelsen af de tilsvarende kontrolgrupper udledt af tilgængelige administrative data eller en passende undersøgelse. Disse oplysninger kan så under visse forudsætninger bruges til at kontrollere, om de deraf følgende mindste påviselige virkninger er tilstrækkeligt signifikante, og om en kontrafaktisk konsekvensanalyse er et fornuftigt valg. Et centralt aspekt i denne øvelse er, om det er sandsynligt, at den pågældende intervention vil skabe virkninger af en størrelse, der svarer til de anslåede mindste påviselige virkninger. På dette punkt kan en undersøgelse af den eksisterende litteratur og andre lignende evalueringer være en hjælp.

*Stikprøvens
sandsynlige
størrelse*

Det er også værd at bemærke, at kontrafaktiske konsekvensanalyser ofte har til formål at undersøge virkningerne på forskellige delpopulationer, der deltager i interventionen (mænd/kvinder, unge/ældre, mindre/mere uddannede osv.). Disse analyser reducerer nødvendigvis størrelsen af de behandlede grupper og kontrolgrupperne for hver delpopulation (mænd, kvinder, unge osv.), og den statistiske signifikans falder derfor. Når undersøgelsen af virkningerne i delpopulationerne er et vigtigt element i evalueringen, er det nødvendigt at have et tilstrækkeligt antal personer i behandlings- og kontrolgruppen i hver delpopulation.

Definitionen af stikprøvestørrelsen kræver tekniske færdigheder, som ikke altid findes i forvaltningsmyndigheden. I så fald skal evaluatoren beregne stikprøvestørrelsen og indgå aftale med forvaltningsmyndigheden om gennemførelsen og omfanget af den kontrafaktiske konsekvensanalyse. Det er ikke muligt at definere den tærskelstikprøvestørrelse, over hvilken det altid er muligt at udføre en kontrafaktisk konsekvensanalyse, fordi tærsklen kan variere afhængigt af de anvendte metoder, de forventede analyser af delpopulationer, de accepterede niveauer for statistisk signifikans og andre elementer. I mange evalueringer har et antal på 2 000 personer for behandlingsgruppen og kontrolgruppen i alt vist sig at være tilstrækkelige til en rimeligt nøjagtig analyse (herunder de vigtigste undergrupper såsom mænd/kvinder osv.). Denne indikation kan ikke bruges som en videnskabelig "tærskel", men blot som en hjælp ved undersøgelsen af, om en kontrafaktisk konsekvensanalyse skal gennemføres.

Stikprøvestørrelsen påvirker pålideligheden af resultaterne af den kontrafaktiske konsekvensanalyse. Nogle eksempler på usikkerheder i fortolkningen af resultater og deres relation til stikprøvestørrelsen forklares i tekstboksen nedenfor.

Tekstboks 12 Usikkerheder i fortolkningen af resultaterne

Blandt eksempler på evalueringer af ESF-finansierede interventioner gav evalueringen af flere foranstaltninger for langtidsledige i Marche-regionen i Italien ikke statistisk signifikante resultater i en af de fire analyserede foranstaltninger med vouchers til uddannelse. Dette skyldes sandsynligvis, at den stikprøve af behandlede personer, der analyseres, var lille i forhold til det store antal faktiske modtagere. Et lignende tilfælde gjorde sig gældende i forbindelse med evalueringen af foranstaltninger til fremme af erhvervsuddannelse i Podlaskie-regionen i Polen. Ud over spørgsmålet om statistisk signifikans viser dette eksempel, at anvendelsen af for små stikprøver fra behandlingsgrupper ved analyse af virkningerne kan føre til problemer med generalisering i resultaterne.

Evalueringen af sprogundervisningen for indvandrere i Tyskland viser positive resultater i modsætning til nogle tidligere undersøgelser, der vurderer virkningerne af sprogundervisningsprogrammer for indvandrere. En mulig forklaring på denne forskel er ifølge forfatteren, at det undersøgte tyske program også bestod af et praktikmodul. Forfatteren kunne ikke teste denne hypotese med de tilgængelige data, da det var umuligt at skelne mellem elementerne sprogundervisning og praktik i interventionen. Dette kan være af potentiel interesse for fremtidig forskning. Eksemplet viser, at det undertiden er nødvendigt at gentage kontrafaktiske øvelser med forbedrede og mere omfattende data eller suppleret med andre evalueringsmetoder for at forklare de mekanismer, der bestemmer resultaterne af en intervention.

2.3.6. En tjekliste til forberedelse af en kontrafaktisk konsekvensanalyse og vurdering af, om den bør gennemføres

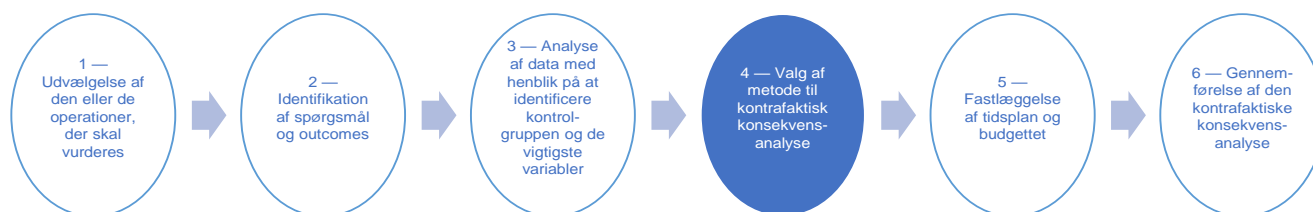
På nuværende tidspunkt er det nyttigt at opstille en tjekliste med de vigtigste faktorer, som en forvaltningsmyndighed bør overveje og kontrollere i forbindelse med udarbejdelsen af en kontrafaktisk konsekvensanalyse:

En liste over de vigtigste faktorer, der skal tages i betragtning

- Er den valgte intervention egnet til en kontrafaktisk konsekvensanalyse?
 - o Er interventionen diskret, karakteristisk og forholdsvis homogen?
 - o Viser "forandringsteorien" for interventionen, at der er en overbevisende årsagsmekanisme for de outcomes, som en kontrafaktisk konsekvensanalyse skal undersøge?
 - o Er det let at identificere en kontrolgruppe i overensstemmelse med reglerne for interventionen?
- Er virkningerne for deltagerne kvantitativt målbare?
 - o Måles outcomes af interventionen klart og i overensstemmelse med forandringsteorien for interventionen?
 - o Er der gået tilstrækkelig tid siden behandlingens afslutning til at påvise outcomes for deltagerne?
- Kræver evalueringsspørgsmålene en måling af virkningen (nettoeffekten) af interventionen?
- Er passende data (behandlings- og kontrolgruppefortegnelser, resultatfortegnelser, kontekstdata) tilgængelige, eller kan de gøres tilgængelige?
 - o Er de tilgængelige data i overensstemmelse med den generelle forordning om databeskyttelse? Hvis ikke, kan de nødvendige administrative aftaler og tekniske løsninger da aktiveres for at bringe dataene i overensstemmelse?
 - o Er det muligt at definere en passende kontrolgruppe med de tilgængelige data?
 - o Er stikprøvens størrelse stor nok til at opnå den nødvendige statistiske signifikans i resultaterne af den kontrafaktiske konsekvensanalyse?

Et bekræftende svar på alle disse spørgsmål betyder, at det er muligt at udforme og gennemføre en kontrafaktisk konsekvensanalyse. Et afkræftende svar på nogle af disse spørgsmål udelukker ikke gennemførelsen af en kontrafaktisk konsekvensanalyse, men kræver yderligere kontrol af betingelserne for gennemførelse og forbedringer i datatilgængeligheden eller andre vigtige faktorer.

2.4. Hvilken metode skal anvendes til den kontrafaktiske konsekvensanalyse?



Valget af den metode, der skal anvendes til den kontrafaktiske konsekvensanalyse, er et afgørende skridt i forbindelse med udarbejdelsen af evalueringen, da den valgte metode påvirker både resultaternes kvalitet og kvaliteten generelt. Generelt er ingen metode bedre end andre, men én metode kan være mere velegnet til at udnytte de tilgængelige data, mens en anden kan være mere hensigtsmæssig til at undersøge en bestemt type intervention. Identificering af den bedst egnede metode til en specifik evaluering er et afgørende trin i udformningen.

Grundlæggende betingelser for at finde den bedst egnede metode

De tekniske aspekter og de særlige karakteristika ved hver metode gennemgås i detaljer i **kapitel 3**, men her vil vi kort gøre rede for de praktiske konsekvenser af den metode, der vælges.

For det første skal der fastlægges en randomiseret tilgang — som tilfældigt allokerer enkeltpersoner til behandlings- eller kontrolgruppen — inden interventionen påbegyndes. Dette forudsætter både, at designet er på plads, og at man har en organisation, der er i stand til at håndtere opgaven. I ESF-interventionerne er denne tilgang sjældent blevet anvendt, men som vi vil se, er den på visse betingelser mulig og meget effektiv ud fra et metodologisk synspunkt.

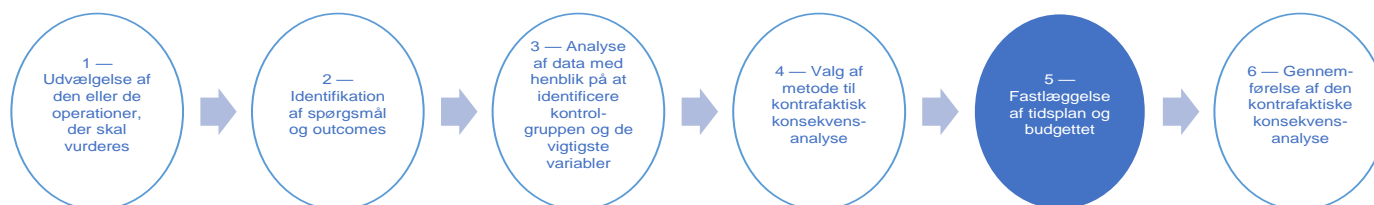
For det andet kan valget af metode kræve mere kapacitet, end forvaltningsmyndigheden har, fordi den kræver avancerede tekniske færdigheder og specifik erfaring. I så fald vælger evaluatoren den bedst egnede metode, men naturligvis først efter at en evaluator er blevet udpeget. Dette kan have den ekstra bonus, at det fungerer som et kriterie for vurdering af evaluatorernes kapacitet, når der skal vælges en evaluator. Forvaltningsmyndigheden skal under alle omstændigheder føle sig sikker på, at en kontrafaktisk konsekvensanalyse skal udføres, efter at have givet bekræftende svar på alle eller de fleste af spørgsmålene i den foregående tjekliste og — internt eller eksternt — skaffet den information, der er nødvendig for at godkende valget af evaluator.

For det tredje kan randomiserede metoder afvige fra kvasiexperimentelle metoder med hensyn til varighed og omkostninger, og i randomiserede kontrolforsøg bør evaluatoren inddrages fra begyndelsen af interventionen. Kvasiexperimentelle metoder adskiller sig ikke væsentligt fra hinanden. Evaluatoren kan påbegynde arbejdet, når interventionen er påbegyndt (men i tide til at organisere og forberede de nødvendige data), og kvaliteten af de tilgængelige data kan gøre forskellen med hensyn til tid og omkostninger.

Endelig giver interventionstypen og datatilgængeligheden nogle gange mulighed for at anvende mere end én kvasiexperimentel metode. I dette

tilfælde kan det være nyttigt at anvende flere metoder, fordi konstatering af de samme virkninger ved hjælp af forskellige metoder gør resultaterne mere robuste.

2.5. Tidsplan og budget



2.5.1. Hvilke ressourcer er til rådighed?

Et centralt spørgsmål, der skal overvejes i forbindelse med udarbejdelsen af en kontrafaktisk konsekvensanalyse, er ressourcer. Der kan være tale om en lang række overvejelser, som her er inddelt i tre overskrifter: a) eksperter, b) tid og c) finansielle ressourcer.

Hvilke eksterne eksperter og interne medarbejdere kræves der til en kontrafaktisk konsekvensanalyse?

I de fleste tilfælde vil udarbejdelsen af en konsekvensanalyse blive lagt ud til en ekstern leverandør. Kontrakten skal dog forvaltes inden for forvaltningsmyndigheden af personale med kendskab til metoder til kontrafaktisk konsekvensanalyse. En sådan viden er nødvendig for at sikre kvalitet og effektivt samarbejde med eksterne eksperter. Der kan også være behov for andre former for ekspertise inden for forvaltningsmyndigheden, f.eks. statistiske færdigheder og ekspertise inden for dataindsamling og -forvaltning. Det er vigtigt på forhånd at overveje, om forvaltningsmyndigheden har adgang til tilstrækkeligt kvalificeret og uddannet personale, og at disse medarbejdere har kapacitet til at deltage i evalueringen.

Internt personale

iværksættelse af en effektiv kontrafaktisk konsekvensanalyse kræver, at de nødvendige færdigheder og den nødvendige erfaring til at udføre sådanne evalueringer er til stede. Desuden skal egnede leverandører forstå den politiske og administrative kontekst i medlemsstaten, have kendskab til potentielle datakilder og beherske de relevante sprog. Det er værd at overveje, hvad der kræves for at oprette en leverandørbase i en medlemsstat (se yderligere om dette emne i kapitel 4).

Eksternt personale

Effektive kontrafaktiske konsekvensanalyser kræver samarbejde fra dem, der forvalter det program eller den intervention, der evalueres. For eksempel vil der være krav om adgang til registre, der føres af de ansvarlige for interventionerne. Disse registre indeholder oplysninger om personer eller virksomheder, der har deltaget i en intervention.

Personale, der forvalter programmer/interventioner

Program- og interventionsforvaltere kan yde rådgivning og vejledning om disse typer data. De kan også være forpligtet til at føre fortegnelser ud over, hvad de ville skulle gøre uden en konsekvensanalyse.

For at løse problemet med dataindsamling fra forskellige kilder vil de, der planlægger en kontrafaktisk konsekvensanalyse, skulle samarbejde med personale, der beskæftiger sig med officielle datakilder (f.eks. arbejdsløshedsregistre, socialsikringsdata, statistiske kontorer osv.). Dette vil gøre det muligt at planlægge leveringen af data i god tid i forvejen.

Statistisk ekspertise

Hvilke faktorer er relevante for tidsplanen for en kontrafaktisk konsekvensanalyse?

Gennemførelsen af en kontrafaktisk konsekvensanalyse kræver bidrag fra en række forskellige menneskelige ressourcer. Samtidig gennemføres sådanne evalueringer over en længere periode. En evalueringsordning bør indeholde en kort tidsplan med vigtige projektdelmål, som vil gælde for dem, der er involveret i selve interventionen, samt for dem, der beskæftiger sig med evalueringen. Tidsplanen skal tilrettelægges på tværs af både evaluerings- og interventionsaktiviteter samt de vigtigste politiske delmål.

Kontrafaktiske konsekvensanalyser tager tid

Det er vanskeligt at udarbejde en meningsfuld og realistisk tidsplan for en kontrafaktisk konsekvensanalyse. På den ene side skal forvaltningsmyndigheden (eller det bemyndigede organ), der planlægger evalueringen, overveje de vigtigste datoer for, hvornår de beslutninger, der afhænger af evalueringens resultater, skal træffes. På den anden side vil der være uundgåelige begrænsninger, som vil påvirke timingen af rapporterne. Nogle resultater vil tage flere år, og dataindsamlings-, analyse- og rapporteringstidsplanerne skal så vidt muligt afspejle dette (se afsnit 2.5.2 nedenfor).

Afsæt ekstra tid for at afvente virkningerne

Evalueringen skal gennemføres så tidligt i programmeringsperioden, at der kan foretages ændringer, og at de indhøstede erfaringer kan udnyttes i den resterende tid. Under visse omstændigheder kan de samme eller lignende interventioner støttes i flere på hinanden følgende programmeringsperioder. Resultater fra kontrafaktiske konsekvensanalyser, der fokuserer på interventioner fra tidligere programmeringsperioder, kan være yderst nyttige som grundlag for gennemførelse og udformning i efterfølgende programmeringsperioder.

Fokus på bestemte tidspunkter

Det er også vigtigt at overveje, hvordan timingen af en kontrafaktisk evaluering passer med timingen af andre evalueringskomponenter. Det er sandsynligt, at en teoribaseret evaluering vil skulle gennemføres inden en kontrafaktisk konsekvensanalyse. For innovative interventioner (f.eks. ESF-interventioner for 2014-2020, der støttede børn i Tjekkiet, eller tjenester til fremme af social inklusion for personer, der modtager den nye indkomststøtte i Italien, eller de forskellige foranstaltninger til støtte for arbejdstagere og studerende under fjernarbejde under covid-19-nedlukningen i mange medlemsstater) kan det være nødvendigt at indberette centrale elementer i en procesevaluering enten tidligere for at kunne indarbejde forbedringer i designet af en kontrafaktisk konsekvensanalyse eller på et senere tidspunkt for at kunne foretage en detaljeret undersøgelse af årsager og virkninger. Når der gennemføres en kontrafaktisk konsekvensanalyse for en moden løbende intervention, vil det være mere relevant, hvis

Planlæg rækkefølgen af forskellige evalueringstyper

procesevalueringen gennemføres sideløbende med konsekvensanalysen.

En tidsplan vil også blive påvirket af tilgængeligheden af data. Det kan tage lang tid at opdatere datakilderne, som det ofte er tilfældet med skatteoplysninger. Det kan også være tidskrævende og dyrt at overvinde juridiske og institutionelle hindringer for indhentning af de nødvendige data. Desuden kræver det meget tid og en betydelig indsats at trække på data fra en række kilder for at sikre, at de er kompatible, kontrollere deres kvalitet og konvertere dem til en form, der kan bruges til at vurdere virkningerne.

Dataindsamling kan være tidskrævende

Hvordan kan omkostningerne vurderes?

Det er vigtigt at udarbejde et vejledende budget for, hvor meget en forvaltningsmyndighed er i stand til og villig til at bruge på en kontrafaktisk konsekvensanalyse. Budgettet består af to komponenter: interne og eksterne omkostninger. Førstnævnte henviser til den indsats, der kræves af de interne ressourcer for at følge evalueringen. I forvaltningerne føres der generelt ikke en detaljeret oversigt over disse omkostninger, men de skal tages i betragtning for at sikre tilstrækkelig kontrol af og opfølgning på evalueringen. Omkostningerne til at bestille eksterne eksperter til en kontrafaktisk konsekvensanalyse skal vurderes omhyggeligt for at sikre en evaluering af høj kvalitet til rimelige udgifter. Der fokuseres her på eksterne omkostninger.

Der skal skelnes mellem evalueringer af rutinemæssige interventioner, hvor udgifterne generelt er lavere, og innovative aktioner eller pilotprojekter, hvor indsamling af relativt store mængder data, anvendelse af nye datakilder og inddragelse af nye interessenter kan berettige højere udgifter. Dette er imidlertid ikke en regel og skal ses i sammenhæng med evalueringsspørgsmålene, hvis kompleksitet og antal kan kræve færre eller flere finansielle ressourcer.

Forslag til metode til vurdering af omkostninger

Valget af evalueringsmetode gør også en forskel. Som nævnt ovenfor kræver en randomiseret tilgang, at evaluatoren er til stede for at indsamle og verificere oplysninger, så længe interventionen varer. På den anden side afhænger indsatsen ved kvasiexperimentelle metoder i vid udstrækning af antallet af datakilder, deres kvalitet og tilgængelighed.

En vejledning fra Kommissionen⁶⁷ giver en indikation af det beløb, der er nødvendigt for evalueringen af et program, og anfører, at storstilede og rutinemæssige programmer ikke bør afsætte mere end 1 % af deres programbudget til evaluering, mens innovative initiativer eller pilotinitiativer kan afsætte op til 10 %. Denne vejledning omhandler imidlertid ikke eksplicit ressourcebehovene til kontrafaktiske konsekvensanalyser, og vi kan gå ud fra, at disse procentsatser er loftet for evalueringssomkostninger, herunder en eller flere kontrafaktiske konsekvensanalyser.

Som man kan forestille sig, er det umuligt at angive vejledende omkostninger til en kontrafaktisk konsekvensanalyse for så mange forskellige interventioner og operationelle vilkår i ESF+-programmer i 27 medlemsstater. Her er det mere nyttigt at foreslå en metode til at anslå

⁶⁷ Se [Europa-Kommissionen, 2013](#).

et rimeligt budget for kontrafaktiske konsekvensanalyser i forskellige sammenhænge. Følgende tabel illustrerer en simpel tilgang: I rækkerne angives de vigtigste opgaver, der skal udføres, og i kolonnerne angives de vigtigste omkostninger.

Varigheden og kompleksiteten af hver aktivitet er afgørende for den indsats, der kræves i forhold til evalueringsteamets arbejdsdage og dermed de vigtigste omkostninger. Omkostningerne til dataindsamling i form af kvantitative undersøgelser af deltagere og kontrolgruppemedlemmer er betydelige. Hvis en kontrafaktisk konsekvensanalyse er afhængig af eksisterende administrative datakilder, og disse er forholdsvis let tilgængelige, vil de samlede omkostninger være lavere. Administrative data kræver imidlertid ofte en væsentlig "forberedelsesaktivitet" (rettelse af manglende eller forkerte data, tilpasning af databasens format, kontrol af de administrative regler for data osv.), hvilket kan øge omkostningerne.

Tabel 3 Overblik over de primære omkostninger til en kontrafaktisk konsekvensanalyse

Primære opgaver	Evalueringsteam			Andre omkostninger (Udstyr, materialer, forsyninger, rejser osv.)	Bemærkninger
	Programforvalter	Senior-eksperter	Junior-eksperter		
Planlægning og koordinering					Planlægningsprocessen omfatter gennemførligheden af en kontrafaktisk konsekvensanalyse, tilrettelæggelse af den og færdiggørelsen af metodologien efter aftale med forvaltningsmyndigheden. I denne fase skal der afsættes tilstrækkelig tid til at analysere potentielle datamangler. Koordineringen dækker hele den kontrafaktiske konsekvensanalyses varighed og omfatter tilrettelæggelse af arbejdet og kontakt med forvaltningsmyndigheden og andre relevante interessenter med henblik på at færdiggøre evalueringsspørgsmålene.
Litteraturgennemgang					Litteraturgennemgangen støtter udarbejdelsen af forandringsteorien, identifikationen af de variabler, der skal anvendes, og forståelsen af konteksten. Den gør det også muligt at finpudse evalueringsspørgsmålene og udnytte tidligere resultater og formulere hypoteser.
Dataindsamling og -forberedelse					Den indsats, der kræves til denne aktivitet, varierer meget afhængigt af dataindsamlingsmetode(r). Indsamling af originaldata via undersøgelser kan være tidskrævende og dyr. Andre metoder kan kræve investeringer i teknologi (software eller hardware) eller indgåelse af aftaler med dataejere. Omkostningerne kan reduceres ved at inddrage dataejerne i udarbejdelsen af data.
Dataanalyse					Dataanalyse kræver avancerede færdigheder og stor erfaring. Den tid, det tager at udføre denne aktivitet, afhænger af metode, analysernes kvantitet og datakvalitet.
Udarbejdelse af rapport(er)					Den indsats, der kræves til denne aktivitet, varierer afhængigt af antallet og typen af rapporter og andre kommunikationsværktøjer. Omkostningerne kan omfatte trykning og grafisk design ud over den der tid, der bruges på udarbejdelse af rapporten.
Opfølgingsmøder					Opfølgingsmøder og andre formidlingsaktiviteter er et vigtigt skridt i formidlingen af resultaterne. Udgifterne hertil kan omfatte leje af lokaler og mad.
I alt					

I mangel af andre referencer kan personaleomkostningerne beregnes på grundlag af ESF+-omkostninger til uddannelses- eller ansættelsestjenester, der er vedtaget i hvert land for senior- og juniorundervisere eller -eksperter. Disse beløb kan generelt sammenlignes med en juniorforskers og seniorforskers vederlag, og selv om de ikke nødvendigvis skal være de samme, kan de fungere som en nyttig rettesnor. En simpel markedsanalyse og nogle interviews med

nogle få forskere kan bidrage til at belyse de forventede omkostninger til den kontrafaktiske konsekvensanalyse.

2.5.2. Hvornår skal interventionen evalueres?

Det er afgørende at fastslå, hvornår i en intervention det er mest hensigtsmæssigt at foretage en konsekvensanalyse, samt de vigtige spørgsmål om, hvornår resultaterne skal måles, og hvilke virkninger der skal vurderes.

Hvornår skal nye og igangværende interventioner evalueres?

Spørgsmålet om, hvornår det i en interventions løbetid er hensigtsmæssigt at foretage en kontrafaktisk konsekvensanalyse, vil afhænge af, om interventionen er ny eller en modnet løbende ordning. For en ny intervention er der brug for mere tid til at blive moden og stabil. Det vil være for tidligt at gennemføre en kontrafaktisk konsekvensanalyse før dette tidspunkt, og det kan potentielt give vildledende dokumentation. I tilfælde af nye interventioner er det ofte nyttigt med en indledende procesvurdering, der gennemføres forud for en kontrafaktisk konsekvensanalyse, for at identificere startproblemer og pege på løsninger.

Tidsplanen er forskellig for nye og igangværende interventioner

Ved fastlæggelsen af den optimale tidsplan for en kontrafaktisk konsekvensanalyse for en ny intervention bør der tages hensyn til en række andre faktorer, herunder: Skridt til at sikre, at der er passende datakilder til rådighed, oprettelse af et internt projektteam bestående af tilstrækkeligt uddannet personale og udvælgelse af en ekstern leverandør. Desuden vil behovet i den beslutningsproces, som evalueringen i sidste ende er rettet mod, være en kritisk hindring.

Ressourcespørgsmål

Tidsplanen for en evaluering af virkningerne af en igangværende intervention vil hovedsagelig afhænge af praktiske og politikrelaterede krav. Interventionen bør have fundet sit leje og nået et modenhedsniveau, hvor det er passende at gennemføre en kontrafaktisk konsekvensanalyse. Et andet spørgsmål, der bør overvejes, er, om der findes andre reformer, der gennemføres parallelt med den intervention, der evalueres. Virkningerne af disse reformer kan påvirke virkningen af den pågældende intervention. De politiske beslutningstagere vil skulle overveje, om andre eksisterende reformer i det politiske landskab er relevant for de politiske beslutninger, som skal træffes på grundlag af den pågældende kontrafaktiske konsekvensanalyse.

ESF-evalueringerne fokuserer normalt på én programmeringsperiode. Især i tilfælde af stabile interventioner, der allerede var en del af ESF-programmet i den foregående periode, kan det dog være hensigtsmæssigt at kombinere en retrospektiv evaluering af den foregående periode med en løbende evaluering af den nuværende periode for at dække en større del af den periode, interventionen løber over.

Fleere programmeringsperioder

Hvornår skal resultaterne måles og virkningerne beregnes?

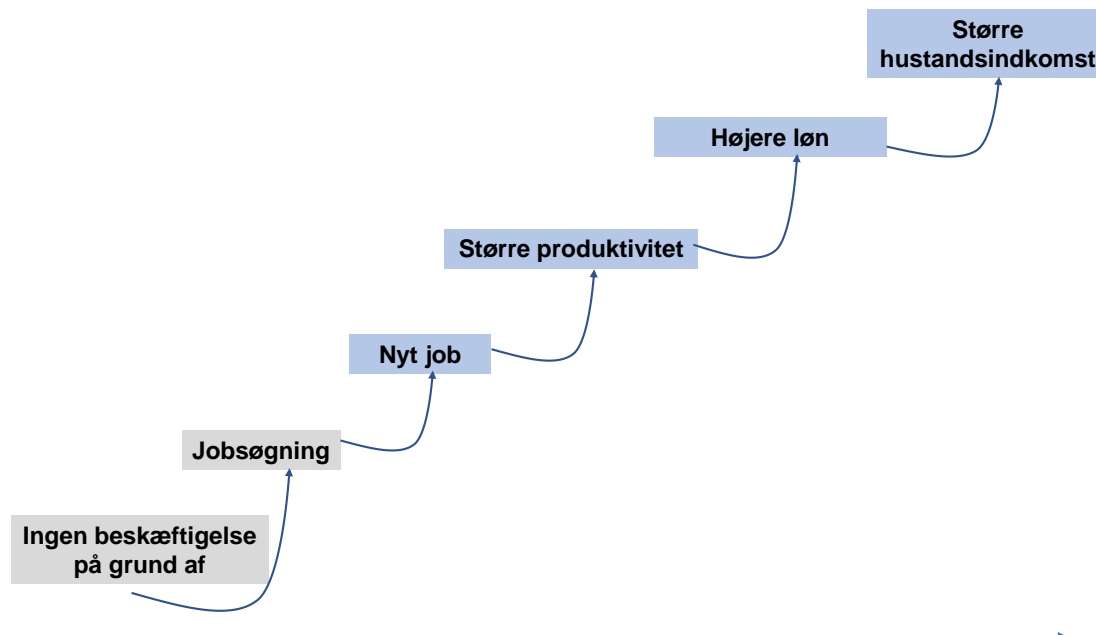
Det andet centrale spørgsmål i forbindelse med tidsplanen for en evaluering er, hvornår virkningerne skal måles og estimeres, eller mere specifikt hvornår det forventes, at virkningerne vil opstå efter en intervention.

Når man undersøger en uddannelsesintervention rettet mod arbejdsløse, er spørgsmålet, i hvilken periode der kan forventes højere beskæftigelsesfrekvenser. Det er velkendt, at uddannelser på kort sigt har tendens til at nedbringe beskæftigelsen blandt deltagerne. Dette skyldes den såkaldte "lock-in"-effekt. Der er en tendens til, at uddannelsesinterventioner afholder arbejdsløse fra at søge job, mens de deltager i kurser. Kontrolpersonerne arbejder derimod for at finde et job. Hvis virkningerne beregnes for tidligt, kan de meget vel være negative eller blive undervurderet. Ved planlægningen af en kontrafaktisk konsekvensanalyse er det vigtigt at være realistisk med hensyn til tidspunktet for virkningerne, og hvornår de sandsynligvis kan måles. En forenklet model for efterfølgende virkninger findes i figur 5 nedenfor.

Overvejelser om, hvornår det er bedst at måle resultater og vurdere virkningerne, skal tage hensyn til de politiske beslutningstageres behov for at modtage oplysninger inden for visse frister. I tilfælde af interventioner, der har til formål at forbedre beskæftigelsesegnetheden på lang sigt, kan det ud fra et analytisk perspektiv være hensigtsmæssigt at følge op på deltagerne to år eller mere, efter at de har modtaget behandling, for at se, om deres indkomst og beskæftigelsesfrekvens er højere end en tilsvarende gruppe af ubehandlede personer. I modsætning hertil har programforvalterne ofte brug for resultater hurtigt, i hvilket tilfælde der skal opnås et kompromis mellem, hvad der er et rimeligt opfølgingsinterval set i lyset af interventionen, og beslutningstageres behov for rettidig dokumentation.

Mellem det rimelige og det mulige

Figur 5 Forenklet tidsplan for resultaterne af et praktikforløb



Hvis resultatmålinger opnås fra administrative kilder (f.eks. socialsikringsregistre med oplysninger om beskæftigelse og indkomst), vil det være praktisk at spore resultaterne gentagne gange over en længere periode og vurdere virkningerne (selv på månedsbasis). Her er der risiko for, at resultaternes karakter kan ændre sig med tiden. Hvis der kræves indsamling af primære data til måling af resultaterne i form af stikprøveundersøgelser, vil det blive meget dyrt at vurdere virkningerne med regelmæssige mellemrum, medmindre det er muligt at indsamle retrospektive data om resultaterne.

Fokus på bestemte tidspunkter

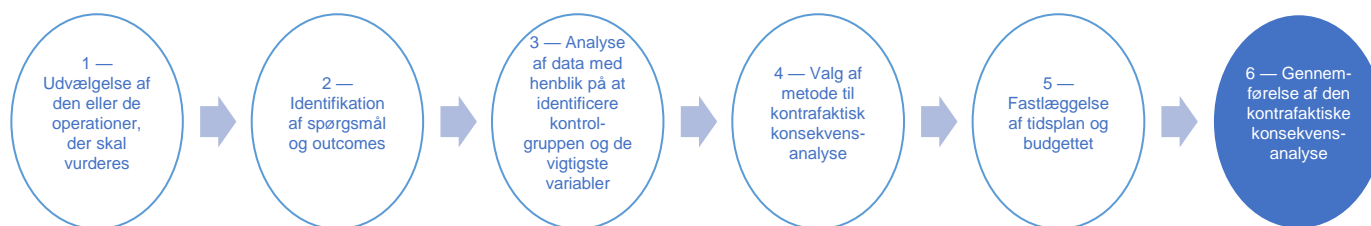
Omkostningerne til at udtrække data fra flere administrative systemer og skabe et enkelt analytisk datasæt bør dog ikke undervurderes.

Som nævnt i afsnit 1.5 kan formuleringen af en forandringsteori (eller interventionslogik) bidrage til at bestemme, hvornår virkningerne skal vurderes.

Et alternativ til dem, der planlægger en kontrafaktisk konsekvensanalyse i mangel af en forandringsteori (men også for dem, der har en klar forandringsteori til rådighed), er at foretage en kort gennemgang af tidligere undersøgelser, der evaluerer interventioner, som ligner den, der overvejes. En nøje gennemgang af resultaterne fra tidligere undersøgelser kan give et godt fingerpeg om den mest hensigtsmæssige måling af resultater og beregning af virkninger.

... eller gennemgang af andre nylige undersøgelser

2.6. Gennemførelse af den kontrafaktiske konsekvensanalyse



Med fastlæggelsen af tidsplanen og budgettet er planlægningen af den kontrafaktiske konsekvensanalyse afsluttet, så nu kan man gå videre til gennemførelsen. Det sidste trin i denne vejledning har til formål at illustrere de vigtigste aktiviteter i forbindelse med gennemførelsen af evalueringen, og hvordan forvaltningsmyndigheden kan overvinde potentielle hindringer, som de møder. I dette afsnit anvendes navnlig erfaringerne fra Evaluation Helpdesk, hvorigennem Kommissionen og GD for Beskæftigelse yder bistand til forvaltningsmyndighederne i alle medlemsstater. Gennemførelsen af evalueringen drøftes generelt i Evalsed, som er et interaktivt onlineredskab til evaluering af den samfundsøkonomiske udvikling⁶⁸. Her gennemgås fire hovedaktiviteter, der er relevante for netop en kontrafaktisk konsekvensanalyse:

1. Udvælgelse af evaluatør
2. Tilsyn med gennemførelsen af den kontrafaktiske konsekvensanalyse
3. Rapportering
4. Anvendelse af resultaterne

2.6.1. Udvælgelse af evaluatør

Den første beslutning i forbindelse med udvælgelsen af en evaluatør er, om vedkommende skal være intern eller ekstern. I sidstnævnte tilfælde sker udvælgelsen af evaluatøren gennem en offentlig udbudsprocedure.

Evaluatorens uafhængighed

I artikel 44, stk. 3, i forordningen om fælles bestemmelser anerkendes det, at der er en problemstilling omkring uafhængighed, og det anføres, at:

⁶⁸ Tilgængelig her: https://ec.europa.eu/regional_policy/sources/docgener/evaluation/guide/guide_evaled.pdf.

"Evalueringer udarbejdes af interne eller eksterne eksperter, der er funktionelt uafhængige". Det første kriterie for udvælgelse af en evaluator er derfor dennes reelle uafhængighed af forvaltningsmyndigheden og ESF+-programmets beslutningsprocesser. Det andet kriterie er, om de nødvendige færdigheder og ressourcer til at gennemføre en kontrafaktisk konsekvensanalyse er til rådighed internt. Dette er sandsynligvis den hyppigste hindring for interne evalueringer, fordi teknisk viden til gennemførelse af en kontrafaktisk konsekvensanalyse er meget specifik og sjældent tilgængelig i de fleste forvaltninger.

Når der er valgt en intern eller ekstern evaluator, skal forvaltningen opstille en liste over evalueringens tekniske specifikationer. Ved udvælgelsen af en ekstern evaluator vil specifikationerne indgå i den offentlige udbudsprocedure. I kommissoriet for en kontrafaktisk konsekvensanalyse fortjener følgende aspekter særlig opmærksomhed:

*Hovedelementerne
i kommissoriet*

- **Formålet med evalueringen** og forskningsspørgsmålene skal være klare og konsekvent føre til en konsekvensanalyse og en kontrafaktisk konsekvensanalyse, hvor det er muligt.
- De **tilgængelige data** skal specificeres. Dette er et af de vigtigste spørgsmål, der skal behandles i forbindelse med udformningen og gennemførelsen af en kontrafaktisk konsekvensanalyse. Det fortæller evaluatorene noget om eventuelle problemer med datatilgængelighed og giver dem mulighed for at udarbejde realistiske og nøjagtige forslag.
- **Anmodningen om at gennemføre en kontrafaktisk konsekvensanalyse** kan være eksplicit i kommissoriet, men det bør overlades til evaluatoren at vælge de specifikke metoder, der skal anvendes. På denne måde kan de foreslå originale løsninger og demonstrere deres kompetencer.
- **Det foreslåede team** skal samlet besidde ekspertise på flere områder, navnlig inden for: evaluering og kontrafaktiske konsekvensanalyser (viden om, hvordan man udformer og gennemfører en evaluering), økonometri (brug af de rette metoder) og det politiske område for den vurderede intervention (rekonstruktion af forandringsteorien og fuld fortolkning af resultaterne).
- **Tildelingskriterierne** skal prioritere kvalitet frem for omkostningerne ved forslaget. Kvaliteten af en evaluering afhænger i høj grad af evaluatorens evne til at udforme og udføre den. De retlige rammer for offentlige indkøb passer ofte ikke fuldstændig til behovene i forbindelse med evalueringssudbud, men forvaltningsmyndigheden bør så vidt muligt tilpasse de eksisterende regler for at kunne lægge vægt på kvalitet og teknisk kapacitet.
- I **udvælgelsesprocessen** bør der sidde eksperter i evaluering og kontrafaktiske konsekvensanalyser i udvælgelseskomitéen for at kunne vurdere forslagernes kvalitet og deres forskellige metodologiske og organisatoriske løsninger korrekt.

I nogle medlemsstater er erfaringerne med kontrafaktiske konsekvensanalyser stadig begrænsede. Dette kan skabe problemer i udvælgelsesprocessen, fordi markedet for denne tjeneste er for lille. Det er muligt, at meget få virksomheder eller eksperter vil være i stand til at opfylde de tekniske betingelser i kommissoriet og fremsætte et forslag af god kvalitet. I disse tilfælde kan forvaltningsmyndigheden tilrettelægge åbne seminarer

med akademikere og virksomheder, der er interesseret i evalueringer, som kan give input til forberedelsen af en kontrafaktisk konsekvensanalyse og den efterfølgende indkaldelse af forslag. Seminarerne vil indsamle forslag fra nationale eksperter og give dem tid til at organisere deltagelsen i udbuddet. For at respektere gennemsigtigheden og undgå interessekonflikter bør seminarer være offentlige, og de vigtigste oplysninger bør også stilles til rådighed for dem, der ikke kunne deltage.

2.6.2. Tilsyn med den kontrafaktiske konsekvensanalyse

Efter at have valgt evaluator og iværksat den kontrafaktiske konsekvensanalyse skal forvaltningsmyndigheden føre tilsyn med evalueringsprocessen. Dette omfatter hovedsagelig kontrol og validering af evaluatorens arbejde, tilrettelæggelse af statusmøder med evaluatoren og interessenterne og om nødvendigt indgåelse af aftaler med dataejerne.

Forvaltningsmyndigheden skal have kapacitet til at udføre disse opgaver. Det betyder også, at der skal være ressourcer til at overvåge hele processen og færdighederne til at kontrollere de tekniske løsninger til kontrafaktisk konsekvensanalyse. Forvaltningsmyndigheden afsætter generelt særligt personale til evalueringerne, og hvis de ikke har tilstrækkelig erfaring med kontrafaktiske konsekvensanalyser, bør de støttes af uafhængige eksterne eksperter. Disse eksperter kan ansættes på korte kontrakter, eller deres ydelser kan være gratis, hvis de kommer fra offentlige institutioner. De vil sandsynligvis skulle afsætte mellem 6 og 12 arbejdsdage, hvilket burde være tilstrækkeligt til at gennemgå rapporterne og deltage i nogle møder med forvaltningsmyndigheden, evaluatoren og de vigtigste interessenter.

Interne og eksterne kompetencer i tilsynsprocessen

En styregruppe anvendes ofte til at overvåge evalueringsprocesserne i et program. Denne gruppe består generelt af forvaltningsmyndigheden, embedsmænd, der repræsenterer andre gennemførelsesafdelinger, de vigtigste interessenter (arbejdsmarkedets parter, NGO'er, andre forvaltninger) og visse eksperter fra den akademiske verden og/eller offentlige institutioner (f.eks. det nationale statistiske kontor). Styringsgruppen deltager i udarbejdelsen af evalueringsspørgsmålene og drøfter evalueringens resultater, formidlingen og anvendelsen heraf i de politiske beslutningsprocesser. En mere fleksibel og specialiseret undergruppe kunne også føre tilsyn med evalueringens resultater med forvaltningsmyndigheden.

2.6.3. Rapportering

Rapportering er en vigtig aktivitet i evalueringen og er det vigtigste middel til at formidle resultaterne. Evalueringens rapporter skal være klare, koncise, forståelige for lægfolk og gennemsigtige i deres vurderinger og politiske anbefalinger. De skal også påvise evalueringens pålidelighed ved at specificere den anvendte metode og de anvendte data samt identificere de mulige begrænsninger for analysen.

Generelt vil der blive udarbejdet tre rapporter for hver kontrafaktisk konsekvensanalyse:

- Den **indledende rapport** forelægges normalt kort tid efter kontraktens underskrivelse og beskriver den endelige udformning af metoden i overensstemmelse med drøftelserne med forvaltningsmyndigheden og den indledende screening af de tilgængelige data.
- Der afleveres en **foreløbig rapport** omkring midtvejs i evalueringens processen og den kontrafaktiske konsekvensanalyse og kan

Særskilte rapporter til opfølgning af fremskridt og anvendelse af resultater

anvendes til at vise de indsamlede data, sammensætningen af behandlings- og kontrolgrupper samt prøveudtagning.

- Den **endelige rapport** indeholder analyserne og resultaterne fra den kontrafaktiske konsekvensanalyse. Den indeholder nærmere oplysninger om metoder og data og fremhæver eventuelle politiske konsekvenser, der muligvis vil følge af konklusionerne. Den endelige rapport skal indeholde nogle centrale afsnit: evalueringsspørgsmål, forandringsteori (eller interventionslogik) og beskrivelse af de outcomes, der skal vurderes, den anvendte metode, anvendte data og behandlings- og kontrolgrupperes karakteristika, estimerede virkninger, svar på evalueringsspørgsmålene og deres politiske konsekvenser.

Myndigheden, der har bestilt analysen, gennemgår alle rapporter med bistand fra eksterne eksperter, hvor det er nødvendigt, og kommentarerne drøftes med evaluatoren, som vil tage sig af de ønskede præciseringer og forbedringer inden den formelle godkendelse af rapporten.

En rapport om en kontrafaktisk konsekvensanalyse skal undgå overdrevne tekniske beskrivelser og sikre, at resultaterne er klare og læselige for interessenter og ikke-eksperter. Specifikke bilag kan omfatte de mere tekniske trin i analysen, der sikrer en klar og flydende beskrivelse i hovedrapporten. Metodologiske detaljer er imidlertid vigtige for at vise, at den kontrafaktiske konsekvensanalyse er pålidelig, og de skal også være gennemsigtige. Rapporten skal omfatte identifikation af outcomevariabler, datakvalitet, behandlings- og kontrolgrupperes sammensætning og stikprøvestørrelse, metoden til estimering af virkningerne og de anvendte statistiske test. Desuden kan de anvendte anonymiserede data og metoderne til databehandling i overensstemmelse med den akademiske praksis gøres tilgængelige for forskere, der vil gentage og kontrollere analyserne.

Den endelige rapport for alle kontrafaktiske konsekvensanalyser bør indeholde et standardiseret sæt minimumsoplysninger. Disse oplysninger tjener to formål: 1) at gøre det lettere at sammenligne resultater på tværs af forskellige OP'er og forskellige lande, 2) at indsamle oplysninger, som vil være tilgængelige for fremtidige metaanalyser.

Modellen for indsamling af et standardiseret sæt minimumsoplysninger om hver ESF-evaluering følger anbefalingerne i pilot- og feasibilityundersøgelsen. Ud fra et operationelt synspunkt kan disse oplysninger medtages i et bilag til den endelige rapport og kan også sendes til Kommissionen.

Tabel 4 viser et forslag til en skabelon til indsamling af de krævede minimumsoplysninger.

Oplysninger, der som minimum skal indgå i den endelige rapport om en kontrafaktisk konsekvensanalyse

Tabel 4 Grundlæggende oplysninger, der skal medtages i en rapport om den kontrafaktiske konsekvensanalyse

Oplysningskategori	Specifikke oplysninger, der skal indsamles og indberettes
1. Oplysninger om intervention	<ul style="list-style-type: none"> - Navnet på den evaluerede intervention - Karakteristika ved interventionen (uddannelse, arbejdsformidlinger, tjenester for social inklusion osv.) - Tilgang for ESF+-programmet (prioritetsakser, specifikt mål, aktion) - Forvaltningsmyndigheden og/eller andre gennemførelsesorganer - Udgifter - Varighed/intensitet af den behandling, der ydes ved interventionen - Deltagergruppe efter alder, køn og støtteberettigelsesstatus (modtager af arbejdsløshedsforsikring, langtidsledige, NEET'er) - Foranstaltningens territoriale anvendelsesområde
2. Måling af interventionseffektivitet	<ul style="list-style-type: none"> - Indikator for, om den kontrafaktiske konsekvensanalyse vurderer en positiv og statistisk signifikant behandlingseffekt på de(t) relevante resultat(er) - Den faktiske størrelse af denne effekt - De to ovennævnte mål (+/-) og (størrelse) for de relevante og sammenlignelige outcomes: f.eks. beskæftigelsesfrekvens, indkomst osv. - "Valuta for pengene"-indikatorer (hvis estimeret)
3. Data og metode	<ul style="list-style-type: none"> - Datakilde (spørgeskemaundersøgelse, administrative data) - Observationsperiode - Økonometrisk analysemetode - Tidshorisont (generelt, kort ≤ 12 måneder fra interventionens afslutning, middel > 12 og ≤ 24 måneder, langsigtede virkninger > 24 måneder) - Stikprøvestørrelser for behandlings- og kontrolgrupper - Tilgængelighed/medtagelse af data fra før programmet

2.6.4. Anvendelse af resultaterne

Anvendelsen af resultaterne er vigtig, fordi evalueringen vil have ringe effekt, hvis resultaterne ikke formidles effektivt og når ud til den tilsigtede målgruppe.

Formidling af resultaterne af den kontrafaktiske konsekvensanalyse er af afgørende betydning

Formidling af konklusioner og evalueringsresultater omfatter normalt:

- mindst én skriftlig evalueringsrapport, herunder et sammendrag og et resumé
- mindst én mundtlig præsentation af resultaterne understøttet af slides eller lignende værktøjer
- et teknisk afsnit (eller et bilag) i rapporten med en grundig redegørelse for den anvendte metode, de vigtigste antagelser og den anvendte tilgang til statistiske analyser.

Alle evalueringsrapporter skal offentliggøres. Dette er fastlagt i forordningen om fælles bestemmelser for programmeringsperioden 2021-2027⁶⁹. Desuden skal evalueringsresultaterne forelægges og drøftes i overvågningsudvalget for programmet. Dette er obligatorisk, men er ikke i sig selv nok til at fremme en dybtgående debat, fordi der generelt fokuseres på operationelle prioriteter på møderne i overvågningsudvalgene. Se tekstboksen nedenfor for et eksempel på national praksis med hensyn til præsentation og formidling af evalueringsresultater.

⁶⁹ Forordningen om fælles bestemmelser 2021/1060, artikel 44, stk. 7.

Der findes ikke én generel effektiv kommunikationspolitik til at formidle og indlede en debat om erfaringer. Hver kontekst kræver sin egen strategi. Det er derfor vigtigt at levere en effektiv strategi og navnlig sikre, at interessenter uden for forvaltningsmyndigheden hører om resultaterne. Det er muligt at opdele den kontrafaktiske konsekvensanalyse "målgruppe" i hovedgrupper og identificere de vigtigste kommunikationstyper. Eksempelvis:

Mange målgrupper og mange kommunikationsstrategier

- Politiske beslutningstagere og politiske repræsentanter på højt niveau — de er vant til at se korte dokumenter (politiske oplæg, sammendrag) med de vigtigste resultater og anbefalinger.
- Embedsmænd fra administrationen, der er involveret i gennemførelsen af interventionen eller lignende politikker — seminarer eller workshops kan være et effektivt instrument til at præsentere resultaterne af den kontrafaktiske konsekvensanalyse og tilpasse præsentationens tekniske niveau til graden af interaktion med deltagerne.
- Politiske interessenter (arbejdsmarkedets parter, NGO'er, støttemodtagere osv.) — sammenfatning af rapporterne og årlige møder om evalueringresultater kan være den kombinerede løsning til at inddrage disse aktører i debatten. Navnlig når disse aktører har bidraget til definitionen af evalueringsspørgsmålene, bør rapporteringen af resultaterne af evalueringen kædes sammen med disse spørgsmål.
- Ekspertter og akademikere — videnskabelige konferencer eller seminarer kan være det bedste sted at fremlægge resultaterne af evalueringen. Disse aktører validerer resultaterne af evalueringen ud fra et videnskabeligt synspunkt og styrker dermed deres pålidelighed.

Metaevaluering⁷⁰ er et godt supplerende instrument til formidling af resultaterne af kontrafaktiske konsekvensanalyser. Det gør det muligt at bekræfte og generalisere de vigtigste resultater, som en enkelt kontrafaktisk konsekvensanalyse ikke kan beskrive alene. For at kunne foretage en metaevaluering er det nødvendigt at have et tilstrækkeligt antal pålidelige kontrafaktiske konsekvensanalyser af god kvalitet om sammenlignelige politikker til rådighed. I øjeblikket er metaevalueringer stadig sjældne og ikke udelukkende fokuseret på ESF-interventioner. En generel anvendelse af metaevalueringer ville kræve en mere udbredt anvendelse af kontrafaktiske konsekvensanalyser og deres forventede planlægning for at gøre det lettere efterfølgende at sammenligne dem. I programmeringsperioden 2021-2027 for ESF+ kan der gennemføres metaevalueringer i lande med et stort antal programmer eller gennem tværnationale sammenligninger.

Metaevaluering med henblik på at styrke udbredelsen og generaliseringen

⁷⁰ Det er en systematisk sammenligning, også understøttet af statistiske værktøjer, af flere kontrafaktiske konsekvensanalyser af lignende interventioner. En metaevaluering giver indikationer på effektiviteten af en interventionstype ved at undersøge virkningerne af lignende foranstaltninger i mange forskellige sammenhænge. Metaevalueringer kan imidlertid ikke gå i detaljer med hver enkelt intervention i sammenligningen, og politisk kan de anvendes til bredere politiske retningslinjer og ikke de enkelte detaljer for hver enkelt intervention.

Tekstboks 13 Polens erfaringer med evalueringskonferencer

Den internationale evalueringskonference, der organiseres af ministeriet for udviklingsfonde og regionalpolitik og det polske agentur for virksomhedsudvikling, er et af de centrale elementer i evalueringen af offentlig intervention i Polen. Det er blevet en veletableret platform for drøftelser mellem interessenter med ansvar for politikudformning, politikgennemførelse og effektvurdering.

Konferencen startede i 2005 og er blevet afholdt hvert år i ti år og to gange om året siden 2015, hvor arrangementet i 2021 var den 14. udgave. I øjeblikket afholdes konferencen over to dage. Den ene dag er helliget forskningsresultater og virkninger af interventionerne, og her drøftes politikken. Den anden dag fokuserer på metodologiske spørgsmål og evaluering som en proces.

I løbet af de 14 konferencer har de kontrafaktiske metoder flere gange været genstand for paneldiskussioner eller oplæg fra eksperter, og de er blevet behandlet fra forskellige synsvinkler på begge dage. I 2017 blev der f.eks. fremlagt vurderinger af virkninger på sessioner om effekten af interventionerne, mens der blev afholdt separate drøftelser om fordele og ulemper ved estimeringsteknikker.

Konferencen er også bemærkelsesværdig på grund af dens forskelligartede målgruppe, da den ikke kun henvender sig til forskere eller specialister inden for metodologi, men også til dem, der udformer og gennemfører interventioner. Konferencen har et dobbelt formål: Opbygning af evalueringssystemets kapacitet ved at udvikle medlemmernes færdigheder og viden samt netværkssamarbejde ved at give interessenterne mulighed for at udveksle erfaringer mellem nationale, regionale og udenlandske partnere. Faktisk samler hver udgave af konferencen ca. 300-400 repræsentanter fra forskellige sektorer — den offentlige forvaltning, den akademiske verden, konsulentfirmaer og NGO'er. Det er et stort antal i betragtning af, at den evidensbaserede tilgang til politik endnu ikke er blevet fuldt ud rodfæstet i Polen.

Den internationale evalueringskonference er en del af den offentlige forvaltnings evalueringssystem. Det var muligt at opbygge en betydelig kapacitet takket være uddannelse afholdt af Det Fælles Forskningscenter og Verdensbanken eller finansieret under teknisk bistand i operationelle programmer. Uddannelse i evaluering, herunder den kontrafaktiske tilgang, var ikke kun tilgængelig for ansatte i evalueringenhederne under forvaltningsmyndighederne, men også for institutioner såsom lokale eller regionale arbejdsformidlinger. Konferencen giver lejlighed til at indsamle, drøfte og udveksle aktuelle begivenheder og resultater⁷¹.

⁷¹ Yderligere oplysninger om konferencen findes på: <https://www.ewaluacja.gov.pl/strony/xiv-miedzynarodowa-konferencja-ewaluacyjna/>. Vi vil gerne takke Piotr Strzęboszewski og andre embedsmænd fra ministeriet for udviklingsfonde og regionalpolitik for oplysninger om conferencesystemet.

Kapitel 3. Valg af den rette metode til kontrafaktisk konsekvensanalyse

Dette kapitel beskriver empiriske metoder til kontrafaktisk konsekvensanalyse. Nærmere bestemt drøftes den eksperimentelle fremgangsmåde (randomiserede kontrollerede forsøg) samt de mest almindelige kvasiexperimentelle metoder: matchning (efter sandsynlighed), difference-in-differences, regressionsdiskontinuitetsdesign og instrumentvariabler. Hver metode er kendetegnet ved en specifik måde at danne en kontrolgruppe på til at besvare det kontrafaktiske spørgsmål "Hvad ville der være sket for behandlingsgruppen med hensyn til outcomes, hvis den ikke havde deltaget i/været eksponeret for interventionen?". I praksis kan den valgte udformning skræddersys til den givne kontekst, alt efter hvilken type intervention der er tale om, og hvilke data der er tilgængelige eller kan indsamles⁷².

Det er ikke muligt at give detaljerede retningslinjer for valget af den mest hensigtsmæssige udformning af evalueringen i lyset af de meget forskelligartede omstændigheder, som forvaltningsmyndighederne arbejder under. Når forvaltningsmyndighederne vælger den mest relevante tilgang til en kontrafaktisk konsekvensanalyse tilpasset deres særlige omstændigheder, bør de overveje, hvad der har fungeret godt i tidligere evalueringer, både inden for selve forvaltningsmyndigheden, i medlemsstaterne og i andre medlemsstater — forvaltningsmyndighederne kan lære af, hvad der er opnået før inden for deres program og andre steder, hvor lignende omstændigheder har været gældende. For udveksling af erfaringer i forbindelse med udformning og gennemførelse af evalueringer kan være nyttige informationskilder i denne henseende. Søgning i litteraturen med henblik på evalueringer af lignende interventioner kan også være en vigtig informationskilde til støtte for designprocessen. Ekspertter, som forvaltningsmyndigheden har bestilt, vil også have input om, hvordan man bedst griber evalueringens designet an. Det er vigtigt at huske på, at der kan være betydelig ekspertise og erfaring inden for forvaltningsmyndighederne, som der kan trækkes på.

Valg af den rette tilgang

3.1. Randomisering — den eksperimentelle tilgang

Randomisering i hovedtræk

- Personer, der er berettiget til at deltage, allokeres vilkårligt til en behandlings- eller kontrolgruppe
- Randomisering sikrer, at begge grupper er identiske (i gennemsnit) med hensyn til alle relevante karakteristika
- Kontrolgruppen besvarer derfor det kontrafaktiske spørgsmål, og forskellen i outcomes mellem behandlings- og kontrolgrupperne indikerer foranstaltningens årsagsvirkning.

⁷² Der findes yderligere oplysninger om andre tekniske metodologiske aspekter i [Europa-Kommissionen, 2019](#).

Randomiserede evalueringsdesign kan antage en eller flere former. Her fokuseres der på en enkel todelt tilgang — én behandlingsgruppe og én kontrolgruppe — for at opridse hovedprincipperne. Figur 1 illustrerer et enkelt randomiseret design.

Det centrale er, at randomisering sikrer, at de to grupper i alle henseender er statistisk ens på tidspunktet for randomiseringen. Derefter eksponeres behandlingsgruppen for den intervention, der er genstand for evalueringen, og hvis virkninger eller effekt skal måles.

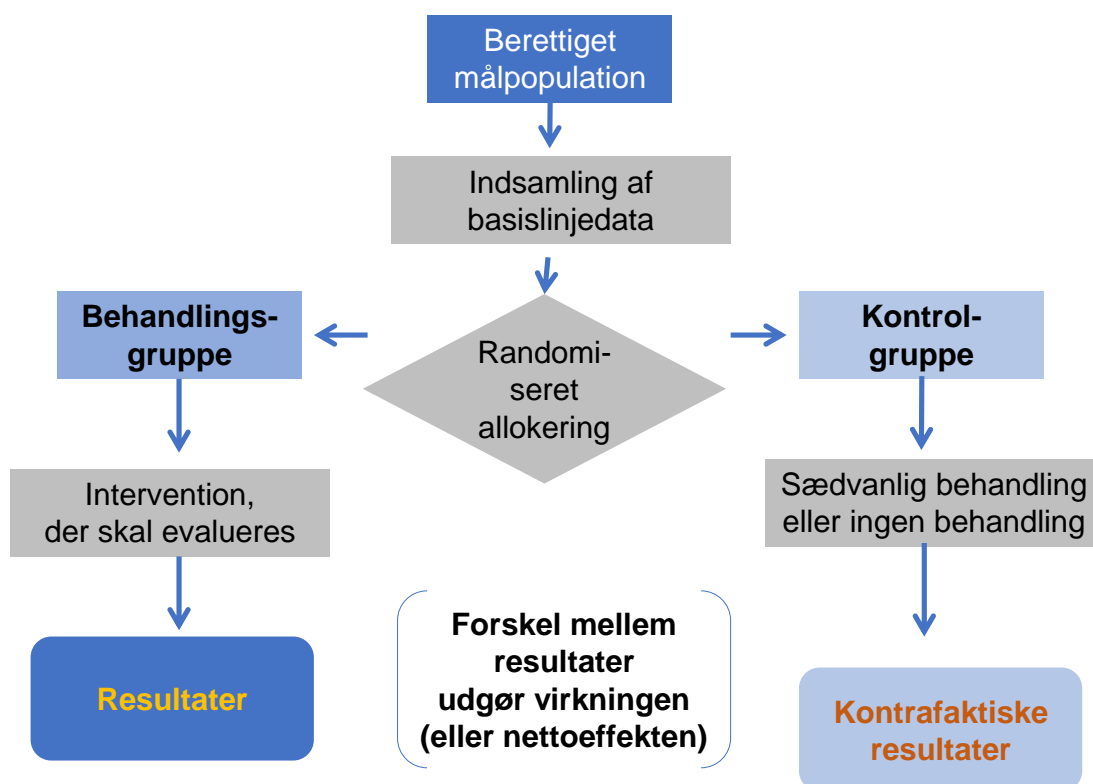
Statistisk ækvivalente grupper

Afhængigt af det centrale politiske spørgsmål kan kontrolgruppen allokeres til slet ikke at modtage behandling, eller behandlingsgruppen kan sammenlignes med en gruppe, der eksponeres for en anden behandling af interesse (kan anses for at repræsentere den sædvanlige behandling), eller der kan være flere behandlingsgrupper sammen med en kontrolgruppe. Der kan f.eks. være interesse i at sammenligne virkningerne af et ESF-finansieret uddannelsesprogram med andre nationalt finansierede uddannelser eller med levering af andre tjenesteydelser til de samme grupper.

Ingen behandling eller anden behandling for kontrolgruppe

Da behandlings- og kontrolgrupper statistisk set er ækvivalente ved randomisering, og eksponeringen for efterfølgende behandlinger kontrolleres, kan forskelle i resultaterne tilskrives den intervention, der evalueres (med forbehold af de sædvanlige statistiske usikkerheder), og alternative forklaringer udelukkes som årsager til eventuelle observerede forskelle (se et eksempel i tekstboksen nedenfor).

Figur 6 Randomiseret kontrol med to grupper



Takket være deres iboende egenskaber, og hvis de gennemføres korrekt, giver randomiserede design mulighed for at frembringe god dokumentation for en interventions virkninger. Af den grund foretrækkes de oftest. De kræver imidlertid tidlig og detaljeret planlægning og kan være vanskelige at udforme og administrere. Desuden vil programforvalterne have betydelige udfordringer

Klar dokumentation...

med at gennemføre dem korrekt. F.eks. kan selve tilstedeværelsen af randomiseringsprocessen ændre sammensætningen af dem, der deltager i en intervention: Det vil sige, at nogle potentielle deltagere måske vil blive afskrækket af idéen om randomisering og nægte at deltage. Desuden opfylder personer, der er genstand for randomisering, måske ikke altid deres tildelingsstatus, ligesom der er en række andre udfordringer, som muligvis ville skulle tackles. I nogle tilfælde kan det være dyrt at udføre randomiserede kontrolforsøg.

Af disse og andre årsager forekommer det usandsynligt, at evalueringer af ESF-finansierede instrumenter og interventioner vil blive foretaget ved hjælp af en randomiseret tilgang. Denne vejledning advarer dog mod at udelukke randomisering uden videre uden nærmere overvejelse. Tilgangen er blevet anvendt i vid udstrækning, og ud over eksemplet fra Det Forenede Kongerige, der drøftes i den følgende tekstboks, er der blandt mange andre også GAIN-forsøgene fra USA⁷³ (der er mange andre eksempler fra Nordamerika), forsøg gennemført i Sverige⁷⁴ i ESF-interventioner i 2014-2020 (se næste tekstboks) samt en undersøgelse foretaget i Tyskland for at vurdere virkningerne af de aktive arbejdsmarkedstjenester, der leveres af private udbydere, i forhold til dem, der leveres af den offentlige arbejdsformidling⁷⁵.

... men vanskelige at udforme

Tekstboks 14 Eksempel på et randomiseret forsøg på et ESF-projekt for unge

Evalueringen af projektet "Ung framtid" i Sverige

Det svenske "Ung framtid" er et ESF-finansieret projekt, der blev gennemført mellem 2015 og 2018. Det havde til formål at støtte unge jobsøgende i alderen 18-24 år, der kom ind på arbejdsmarkedet i Mellemnorrland, det nordlige Mellemsverige og Sydsverige. Projektet intensiverede og individualiserede støtteaktiviteter rettet mod unge jobsøgende og blev gennemført af den svenske arbejdsformidling (Arbetsförmedlingen). Det omfattede individuel planlægning, information, rådgivning og konkret støtte til matchning, uddannelse og andre mulige aktiviteter. Projektet nåede ud til næsten 17 000 unge.

Evalueringen blev bestilt af det svenske ESF-råd hos Arbetsförmedlingen som led i den evalueringsramme, der blev udviklet i forbindelse med et ESF-projekt for 2014-2020, der havde til formål at forbedre de offentlige arbejdsformidlingers evalueringskapacitet (evidensbaseret EU-finansieret projekt for 2014-2020 — se tekstboksen i det sidste kapitel i vejledningen).

I et forsøg, der blev gennemført mellem juni 2017 og januar 2018, og som omfattede otte lokale arbejdsformidlinger ud af i alt 90 deltagere i projektet, blev støtteberettigede unge tilfældigt allokeret til behandlingsgruppen (modtog intensiveret støtte) og kontrolgrupper (modtog almindelig støtte)⁷⁶. 4 689 unge blev tilfældigt allokeret til behandlings- og kontrolgrupper i forsøget, hvoraf 2 972 blev allokeret til kontrolgruppen. Processen for tilfældig allokering resulterede i behandlings- og kontrolgrupper, som var meget ens på allokeringstidspunktet. Som følge heraf kan eventuelle forskelle mellem de to grupper med hensyn til outcomevariablerne (indtræden på arbejdsmarkedet, andel af arbejdsløse og gennemsnitligt antal dage som arbejdsløs), der måles efter deltagelse i ESF-interventionen, tilskrives aktiviteterne i "Ung framtid".

Resultaterne af undersøgelsen viser, at "Ung framtid"-projektet havde en positiv indvirkning på kvinder, men ikke på mænd. For kvinder blev der konstateret positive virkninger i alle outcomevariabler, der blev målt: behandlede kvinders indtræden på arbejdsmarkedet var omkring syv procentpoint højere i de første to måneder efter deltagelsen i projektet, andelen af arbejdsløse faldt mere i den behandlede gruppe end i kontrolgruppen, og endelig var de behandlede kvinder arbejdsløse i færre dage end kvinder i

⁷³ Se [Riccio J., Friedlander, D., Freedman S., 1994.](#)

⁷⁴ Se [Hagglund, P., 2006.](#)

⁷⁵ Se [Krug G., Stephan G., 2011.](#)

⁷⁶ Der blev gennemført en pilotfase mellem april og maj 2017 for at teste flere aktiviteter, herunder styring af randomiseringen, inddragelse af unge deltagere og rapportering.

kontrolgruppen⁷⁷.

Randomiserede design adskiller sig primært fra andre tilgange ved at lægge stor vægt på at kontrollere potentielle skævheder mellem behandlings- og kontrolgrupper gennem forskningsdesign. Denne vægt på design gør tilgangen meget intuitiv, men kræver en forudgående planlægning. Randomiserede design gennemføres ofte bedst i forbindelse med evaluering af nye pilotinterventioner frem for eksisterende. Det skyldes, at de kræver en vis grad af kontrol med, hvordan deltagerne rekrutteres til den intervention, der evalueres. Denne "kontrol" er ofte vanskeligere at opnå i eksisterende programmer end i nye interventioner, der er åbne for nye idéer.

Randomisering gennem forskningsdesign

Gennemførelse af et randomiseret kontrolleret design betyder, at en del af den berettigede målpopulation ikke tilbydes deltagelse, men allokeres til en kontrolgruppe. Denne allokering er helt tilfældig og sker ikke på foranledning af hverken ansøgeren eller interventionens administratorer. Derfor kan de politiske beslutningstagere være tilbøjelige til at modsætte sig randomiserede kontrollerede forsøg af etiske grunde, før de overvejer, om de er relevante ud fra et praktisk og analytisk perspektiv.

Etiske indsigelser

Der er imidlertid et stærkt argument for randomiserede design. Hvis randomisering giver den bedste og mest pålidelige dokumentation for effektiviteten af offentligt finansierede interventioner, er det vigtigt, at det anvendes mere bredt i vurderingen af indvirkningerne af interventionerne. Hvis virkningerne af en bestemt intervention på forhånd er ukendte, er det ikke uetisk at udelukke enkeltpersoner, da vi ikke kan gå ud fra, at den ville have gavnet dem. Desuden anvendes sådanne tilgange i vid udstrækning inden for medicin og andre studieområder som f.eks. uddannelsesforskning. Endelig kan det under visse omstændigheder, hvor de tjenester og den støtte, der ydes ved en intervention, ikke er tilstrækkelige (dvs. at der er flere støtteberettigede personer, end interventionen kan rumme), være det mest etiske middel til at fordele knappe ressourcer at allokere enkeltpersoner tilfældigt fra puljen af dem, der opfylder kriterierne.

... men også stærke argumenter for anvendelse af randomiseret design

3.2. Ikke-randomiserede eller kvasiekperimentelle designs

Hvis randomisering ikke er mulig, findes der et sæt alternative metoder i værktøjskassen til kontrafaktiske konsekvensanalyser, som anvender forskellige tilgange til at oprette en kontrolgruppe til at besvare det kontrafaktiske spørgsmål. Dette underafsnit skitserer først den største udfordring ved kvasiekperimentelle metoder generelt og præsenterer derefter fire tilgange, der er almindeligt anvendt i empirisk praksis i forbindelse med kontrafaktiske konsekvensanalyser.

Når randomisering ikke er mulig

3.2.1. Mål- og kontrolgrupper uden randomisering

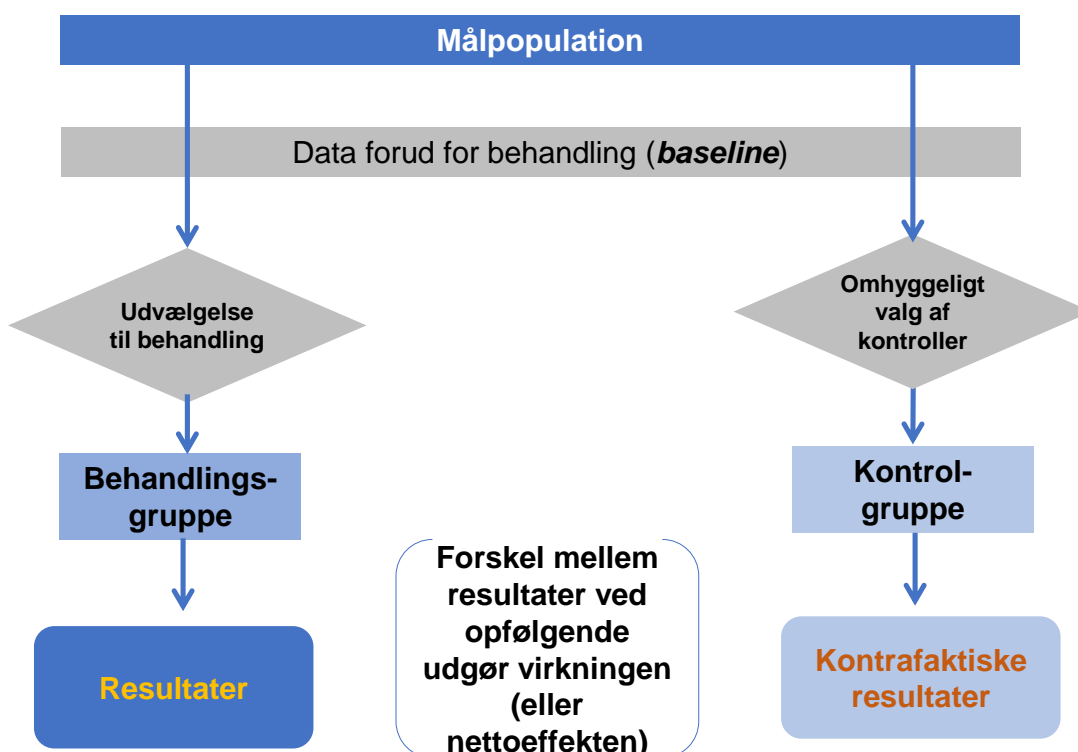
I kvasiekperimentelle designs sammenlignes de målgrupper, der modtager interventionen, med en kontrolgruppe⁷⁸ af ikketilfældigt allokerede mål eller

⁷⁷ Axdorph E, Egebark J., Lundström T., Özcan G., 2019.

⁷⁸ "Kontrolgruppe" er strengt taget det ord, der bruges i eksperimentelle designs, mens man i forbindelse med kvasiekperimentelle designs normalt taler om "sammenligningsgrupper". I praksis anvendes de to imidlertid ofte synonymt. I denne vejledning anvendes "kontrolgruppe" derfor for den gruppe, der anvendes til at estimere den kontrafaktiske situation, uanset designet af den kontrafaktiske konsekvensanalyse.

potentielle mål, der ikke modtager interventionen. Ligesom ved et eksperiment er målet at opnå et skøn uden bias over de ændringer i outcomes, som den pågældende intervention har medført. Da behandlings- og kontrolgrupper ikke dannes tilfældigt, kræver kvasiexperimentelle designs langt mere opmærksomhed på metoder, der tager højde for potentielle forskelle mellem behandlingsgruppens medlemmer og potentielle kontroller, som sandsynligvis vil påvirke beslutningen om at deltage og dermed resultaterne. Det vigtigste er at udvælge en plausibel kontrolgruppe. Undladelse af at udvælge en passende kontrolgruppe og tage højde for de resterende forskelle mellem de to grupper i analysen svækker troværdigheden af estimaterne og kan hindre forsøgene på at udelukke alternative forklaringer på eventuelle observerede virkninger.

Figur 7 Stiliseret kvasiexperimentelt design med behandlings- og kontrolgrupper



I forbindelse med ESF-medfinansierede interventioner består det hyppigst anvendte kvasiexperimentelle evalueringsdesign af to grupper med en basislinje og opfølgning. Sådanne designs omfatter en kontrolgruppe og en behandlingsgruppe som ved randomisering, bortset fra at kontrolgruppen er konstrueret (uden brug af randomisering) fra eksisterende ikkedeltagende grupper, hvilket gør den så sammenlignelig som muligt med den behandlede gruppe.

Kontrol- og behandlingsgrupperne skal ligne hinanden

En god strategi til at finde en valid kontrolgruppe inden for kvasiexperimentelle rammer er at udvælge kontrolpersoner, der er blevet udelukket fra behandlingen, på grundlag af faktorer, der ikke har noget at gøre med deres egenskaber og potentielle resultater. I nogle tilfælde kan der være grund til at tro, at selv om kontrolgrupper ikke er blevet oprettet fuldstændig tilfældigt, kan personer eller virksomheder identificeres efterfølgende, hvis manglende eksponering for behandlingen viser sig at være tilfældigt i forhold til de potentielle resultater. Hvis disse omstændigheder opstår, er de tæt på at være ideelle inden for rammerne af

en kvasiekperimentel tilgang. F.eks. kan visse medlemmer af en interventions målgruppe udelukkes fra deltagelse i interventionen som følge af administrative forsømmelser eller fejl. Det er derfor yderst vigtigt at forstå processen for udvælgelse til behandlingen i forbindelse med oprettelsen af en gyldig kontrolgruppe — dette kan ikke understreges nok.

En troværdig kontrolgruppe kan oprettes på en række måder. For det første kan der anvendes en statistisk matchningsmetode: Her indsamles data fra både behandlede personer og en (typisk meget stor) stikprøve af ikkebehandlede personer. En kontrolgruppe dannes derefter ud fra gruppen af ikkebehandlede personer ved at vælge dem, der er mest ens for personerne i behandlingsgruppen. "Ensartethed" henviser til en række sociodemografiske karakteristika — f.eks. alder, køn, uddannelse, beskæftigelse og arbejdsløshed osv. — målt på et tidspunkt, inden behandlingsgruppen blev optaget i programmet. Som følge heraf bliver ikkebehandlede personer reelt "matchet" med behandlede personer. I praksis kan den potentielt lange liste over sociodemografiske karakteristika opsummeres ved hjælp af værdien af "propensity score", hvilket letter anvendelsen.

Matchning af behandlede og ikkebehandlede personer

3.2.2. Propensity score matching (matchning efter sandsynlighed)

Propensity score matching i hovedtræk

- Matchning imiterer randomisering ved at oprette en efterfølgende kontrolgruppe, der så vidt muligt ligner behandlingsgruppen med hensyn til alle relevante karakteristika.
- I modsætning til randomisering kan kun observerbare karakteristika (alder, køn, uddannelse osv.) matches, mens ikkeobserverbare karakteristika (f.eks. motivation) ikke kan tages i betragtning.
- Metodens validitet afhænger i høj grad af tilgængeligheden af data.

Propensity score matching (PSM) indebærer estimering af en statistisk model for hele stikprøven (behandling og potentielle kontroller), som giver en skønnet sandsynlighed for at deltage for hver enkelt person eller virksomhed — uanset om de rent faktisk deltager eller ej⁷⁹. Behandlede personer eller virksomheder matches derefter — enten med én ikkebehandlet person eller virksomhed eller med flere ikkebehandlede personer eller virksomheder — på grundlag af propensity scoren⁸⁰. Denne procedure identificerer en kontrolgruppe, der efterfølgende kan anvendes til at udlede et skøn over den kontrafaktiske situation. På denne måde sikres det, at skøn over virkningerne tager hensyn til de observerbare forskelle mellem den behandlede gruppe og dem, der fungerer som kontroller, og hvis det antages, at alle relevante forskelle før matchning kan observeres, kan der opnås et skøn uden bias over interventionens virkninger. Hvis udvælgelsen til behandling imidlertid er baseret på uobserverede faktorer, vil der fortsat være tvivl om, hvorvidt matchningen er tilstrækkelig, for så vidt angår dens evne til at fjerne eventuel bias. Den kritiske antagelse, der ligger

Sandsynlighed for at deltage som en metode til definere behandlings- og kontrolgrupper

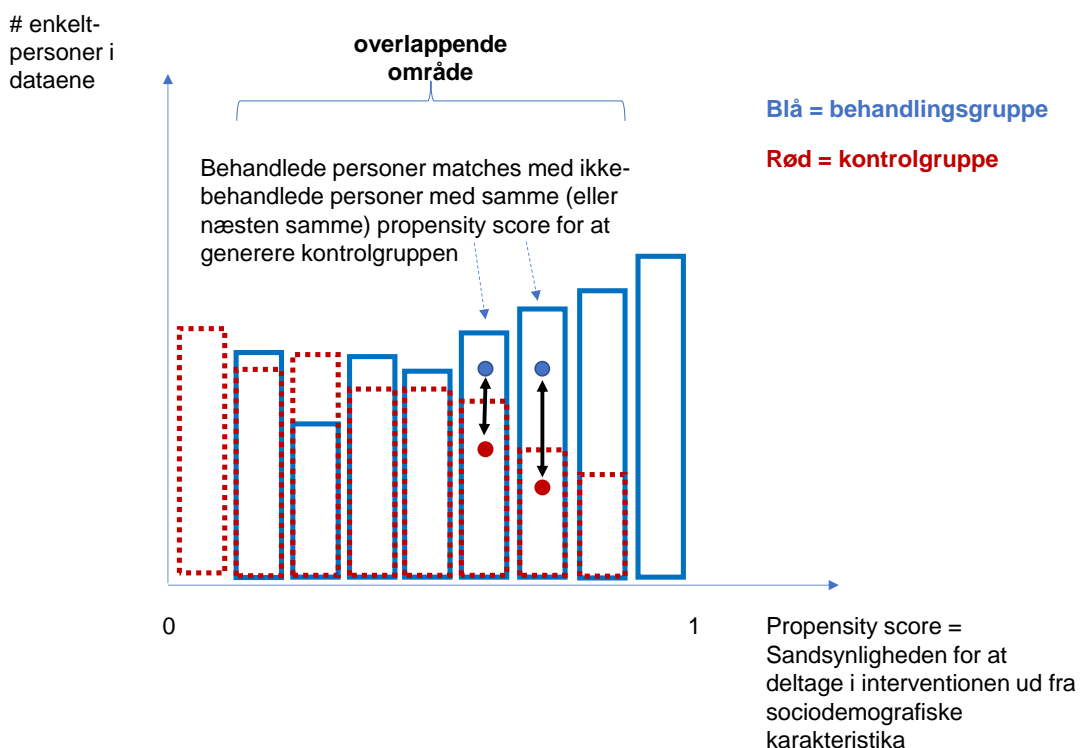
⁷⁹ For at forenkle denne diskussion antages det, at de politiske beslutningstagere ønsker at kende virkningen af behandlingen på dem, der rent faktisk modtog tjenester fra programmet (det vil i mange tilfælde sige en undergruppe af målgruppen, som fik tilbudt muligheden). Dette kaldes en "treatment on the treated"- eller TOT-analyse.

⁸⁰ Der findes en lang række mulige tilgange til propensity score matchning. En enkel oversigt findes i [Caliendo M., Kopeinig S., 2008](#).

til grund for matchningsmetoden, er, at udvælgelsesprocessen er baseret på observerbare data.

Figuren nedenfor viser en intuitiv og forenklet illustration af propensity score matching. Y-aksen repræsenterer antallet af individer i de behandlede og ikkebehandlede grupper sat op efter propensity score på X-aksen. Typisk har behandlede individer relativt højere propensity scores, mens ikkebehandlede personer ofte har lavere propensity scores. Det område, hvor de to gruppers propensity scores overlapper hinanden, kaldes "region of common support".⁸¹ Behandlede cases matches med ikkebehandlede cases i det overlappende område. Der gives to eksempler i diagrammet, men processen gentages i det væsentlige, indtil hver behandlet case svarer til en ikkebehandlet case inden for det overlappende område. I figuren anvendes "nearest neighbour"-matchning. Den "nærmeste nabo" for et medlem af behandlingsgruppen er det kontrolgruppemedlem, der har den nærmeste propensity score. Når der er dannet to grupper, kan deres gennemsnitlige resultater sammenlignes for at få et skøn over virkningerne. I praksis kan propensity score matching blive en meget kompleks proces med en række spørgsmål, der skal overvejes. Mange af disse spørgsmål ignoreres her for at sikre, at hovedprincipperne er klare. Næste tekstboks beskriver et praktisk eksempel på en ESF-evaluering, hvor man brugte matching.

Figur 8 Illustration af propensity score-metoden



Propensity score-metodens troværdighed bygger bl.a. på den antagelse, at udvælgelsen til behandling fuldt ud kan karakteriseres af de observerbare data. Med andre ord, at der ikke er nogen uobserverede forskelle mellem behandlings- og kontrolgrupper, der vedrører resultater og/eller beslutningen om at deltage i interventionen. Troværdigheden af denne antagelse styrkes ved indarbejdelse af en lang række variabler i estimeringen af propensity

Udvælgelse på grundlag af observerbare data

⁸¹ Hvor stor det overlappende område er, har betydning for stikprøvestørrelsen og resultaternes politiske brugbarhed, navnlig hvor et stort antal behandlede cases ligger uden for det overlappende område.

scores, idet udvælgelsen af variabler baseres på forudgående viden og teori. Specifikt anbefales det i forbindelse med interventioner på arbejdsmarkedet i høj grad at inddrage individuel tidligere arbejdsmarkedshistorik i forbindelse med kontrol for potentielle uobserverede forskelle⁸².

Tekstboks 15 Et eksempel på en evaluering baseret på en matchningsmetode⁸³

Virkningerne af ESF-interventioner finansieret i 2014-2020 for langtidsledige i Marche-regionen⁸⁴

En matchningsmetode blev brugt til at evaluere virkningen af en række ESF-interventioner (praktikophold, jobstipendier, erhvervs erfaring i kommuner og uddannelsesvouchere) for de langtidsledige, der finansieres i regionen Marche i Italien. Interventionerne var ikke specifikt rettet mod langtidsledige, men repræsenterede størstedelen af deltagerne i alle de ESF-interventioner, der blev analyseret.

Formålet med evalueringen var at fastslå ESF-foranstaltningernes indvirkning på sandsynligheden for at være i beskæftigelse 6, 9, 12, 15 og 18 måneder efter påbegyndelsen af interventionerne. Der blev anvendt flere outcomevariabler: sandsynligheden for at være i beskæftigelse på et bestemt tidspunkt efter interventionerne, sandsynligheden for at være i beskæftigelse på en tidsubegrænset kontrakt og antallet af arbejdsdage i en periode efter interventionerne.

Der blev anvendt en propensity score-tilgang (nearest neighbour matching) til at identificere forskelle mellem behandlede og ikkebehandlede grupper, og de fire ESF-interventioner, der er nævnt ovenfor, blev analyseret separat. For hver intervention bestod behandlingsgruppen af deltagere, som påbegyndte interventionen inden udgangen af august 2019 (526 for praktikophold, 1 058 for jobstipendier, 236 for erhvervs erfaring i kommuner og 241 for uddannelsesvouchers). Kontrolgruppen bestod af arbejdsløse, der var registreret hos de offentlige arbejdsformidlinger i perioden 2016-2018, med en ledighedsperiode på mindst 12 måneder (77 255 registreringer).

Der blev anvendt en række variabler til beregning af propensity scoren ud fra de sociodemografiske karakteristika ved behandlede og ikkebehandlede, såsom: køn, alder, statsborgerskab, uddannelsesniveau, bopæl, dato for deltagelse i interventionen og beskæftigelsehistorik 36 måneder før interventionen.

Til måling af outcomevariablerne for behandlingsgrupper og kontrolgrupper (før og efter interventionerne) fra COB, som er arkivet over virksomhedernes obligatoriske anmeldelse af ansættelsesaftaler indgivet til de offentlige arbejdsformidlinger, anvendt.

Resultaterne af undersøgelsen var blandede med positive virkninger for praktikophold og jobstipendier, mens der var negative virkninger for erhvervs erfaring i kommunerne og ingen væsentlige virkninger for uddannelsesvouchere⁸⁵.

3.2.3. Difference-in-differences

Difference-in-differences i hovedtræk

- Difference-in-differences er en intuitiv tilgang, der sammenligner forskellen før og efter i outcomes i behandlingsgruppen med outcomes i kontrolgruppen.
- Eftersom ændringen over tid i kontrolgruppen måler, hvad der ville være sket for behandlingsgruppen uden interventionen (det kontrafaktiske scenarie), giver enhver yderligere forskel før/efter i behandlingsgruppens outcomes årsagsvirkningen af interventionen.
- Dette er en enkel metode, der ofte er mulig at udføre i praksis.

⁸² Se [Caliendo, Mahlstedt og Mitnik, 2017](#) og [Kluve, Lehmann og Schmidt, 2008](#).

⁸³ Der findes andre eksempler på anvendelse af metoderne (PSM, RDD og Diff-in-Diff) i praksis i [Europa-Kommissionen, 2020](#).

⁸⁴ Se [Pompili M., Giorgetti I., 2020a](#).

⁸⁵ Som det blev understreget i evalueringsrapporten kan resultaterne for uddannelsesvouchere kun betragtes som foreløbige, eftersom stikprøven er for lille i forhold til deltagerantallet.

Enten alene eller sammen med matchning kan basislinjemål (før behandling) for resultatvariablerne anvendes til at foretage difference-in-differences-estimerer. Her trækkes forskellen i et resultat før og efter behandling i en kontrolgruppe fra den samme forskel, der er observeret i en behandlet gruppe, for at nå frem til et skøn over virkningen af en intervention. Igen er det vigtigt at vælge en plausibel kontrolgruppe. Virkninger beregnet på grundlag af difference-in-differences observeres normalt inden for en regressionsramme, som også tager højde for andre observerede forskelle mellem behandlings- og kontrolgrupper. Desuden kontrollerer denne tilgang for uobserverede forskelle mellem de to grupper, der er faste over tid, og dem, der varierer over tid, men som påvirker både kontrol- og behandlingsgrupper ligeligt (f.eks. faktorer, der påvirker hele økonomien). På grund af denne mulighed for at kontrollere visse aspekter af uobserverede forskelle mellem behandling og kontrol udgør en tilgang baseret på difference-in-differences i de fleste tilfælde en forbedring i forhold til en matchningsstrategi baseret på et tværsnit. Figur 4 illustrerer difference-in-differences-metoden.

*Før og efter
behandlingen*

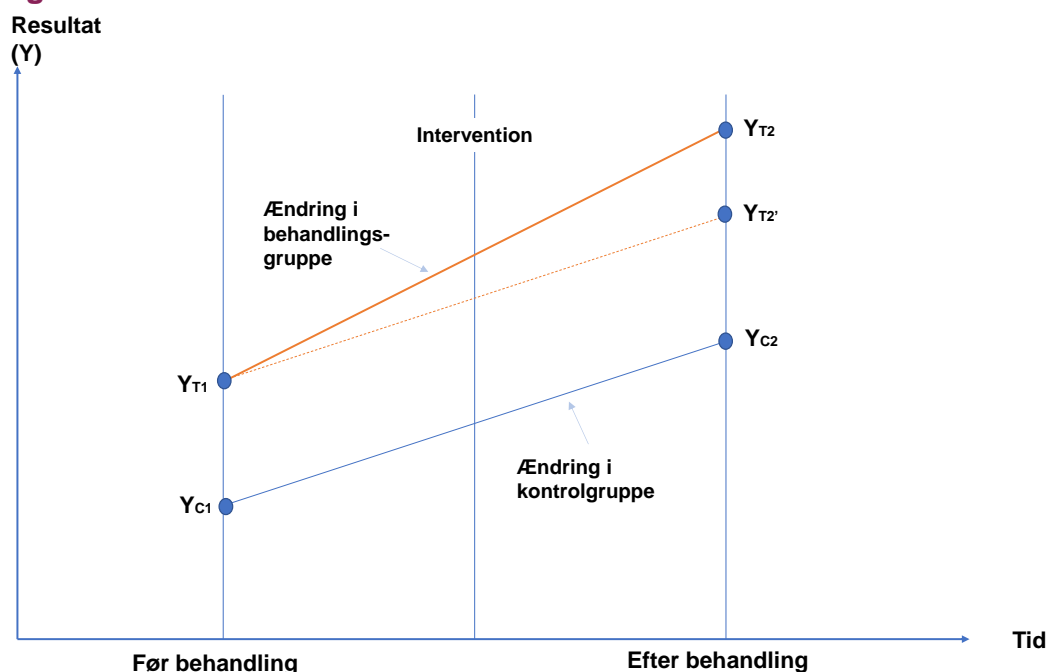
X-aksen repræsenterer tiden og y-aksen en skala, på hvilken resultaterne registreres. Resultatet kan i dette tilfælde være løn. Gennemsnitslønnen for behandlingsgruppen i perioden før behandlingen er YT1, mens den for kontrolgruppen er YC1. I perioden efter behandlingen er lønnen YT2 og YC2 for henholdsvis behandlings- og kontrolgrupperne. Den øverste ubrudte linje repræsenterer således lønudviklingen i behandlingsgruppen, mens den nederste ubrudte linje viser lønudviklingen i kontrolgruppen.

En sammenligning af lønninger i behandlings- og kontrolgrupper i perioden efter behandlingen, dvs. YT2-YC2, ville give et groft skøn over virkningen af interventionen. Dette ville imidlertid være forkert, da det ville ignorere forskelle i lønninger før behandlingen. Én måde at betragte difference-in-differences-metoden på er, at den trækker skønnet over bias før behandlingen fra forskellen i resultaterne efter behandlingen. Som følge heraf justeres forskellen i lønninger efter behandling (YT2 - YC2) ved at fratække forskellen i lønninger før behandlingen (YT1 - YC1), og derfor kan estimeringen af virkningen ifølge difference-in-differences meget enkelt udtrykkes som følger:

$$(YT2 - YC2) - (YT1 - YC1).$$

*Logikken bag den
dobbelte forskel*

Hvis lønforskellene efter behandling ikke justeres for allerede eksisterende forskelle mellem behandlings- og kontrolgrupper, kan der forekomme bias i skønnene. Alternativt kan difference-in-differences-metoden, som tidligere nævnt, betragtes som, at den trækker ændringen i resultaterne i kontrolgruppen fra den ændring, der er observeret i behandlingsgruppen. Den observerede ændring i kontrolgruppen opfattes som den ændring, der ville have fundet sted i behandlingsgruppen, hvis interventionen ikke havde fundet sted.

Figur 9 Illustration af difference-in-differences-metoden

I den enkleste case er den vigtigste antagelse, som difference-in-differences-metoden bygger på, den antagelse, at der er tale om fælles tendenser: Det vil sige, at tidstendenserne i resultaterne inden for behandlings- og kontrolgrupper er ens i fravær af behandling. Denne antagelse kan ikke testes direkte, selv om den i tilfælde, hvor der foreligger flere målinger af resultaterne før behandlingen for både behandlings- og kontrolgrupper, og hvor disse målinger viser parallelle tendenser, støtter metodens plausibilitet. Et eksempel på difference-in-differences-metoden findes i tekstboksen nedenfor.

*Hovedantagelse:
Fælles tendenser
er ens i fravær af
behandling*

Tekstboks 16 Et eksempel på en evaluering baseret på en difference-in-differences-metode

Evaluering af foranstaltninger for ældre arbejdstagere i Lubelskie finansieret af ESF i programmeringsperioden 2007-2013⁸⁶

Hovedformålet med undersøgelsen var at evaluere ESF-interventioner til forlængelse af ældres erhvervsaktive periode på baggrund af ugunstige demografiske forhold i Lubelskie-regionen i Polen. Interventionerne blev finansieret under det operationelle program for menneskelig kapital 2007-2013.

Den behandlede gruppe, en stikprøve på 145 personer, bestod af personer, der var arbejdsløse, da deres deltagelse i interventionerne begyndte. Kontrolgruppen blev taget fra de anonymiserede registerdata fra arbejdsformidlingerne, som indeholdt sociodemografiske data og den registrerede historik vedrørende specifikke personers beskæftigelsesstatus (data fra otte arbejdsformidlinger indeholdt oplysninger om 67 102 personer). Outcomevariablen var "status som registreret i arbejdsløshedsregistret"⁸⁷, og dette kunne observeres både før og efter interventionen, hvilket gjorde det muligt at gennemføre en difference-in-differences-analyse. Beskæftigelsesstatus blev målt på fem tidspunkter: seks måneder før projektdeltagelse, under projektdeltagelsen og 6, 12 og 18 måneder efter projektets afslutning.

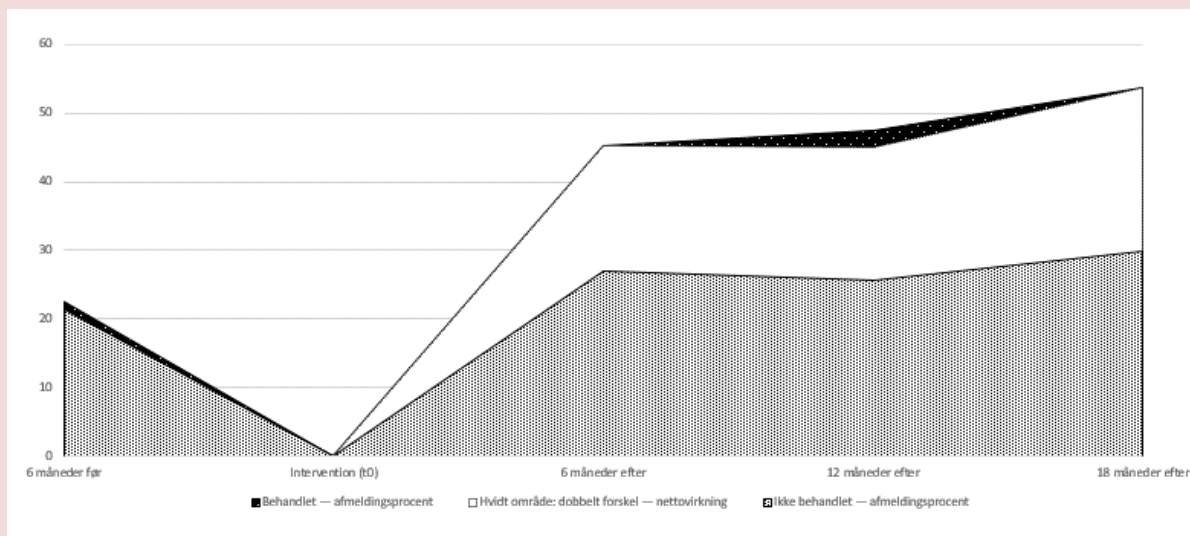
Valget af kontrolgruppe i denne difference-in-differences-tilgang blev understøttet af matchningsmetoder til at identificere en ikkebehandlet person for hver behandlet person på grundlag af en række observerbare

⁸⁶ Se [Re-source Pracownia Badań i Doradztwa, 2015](#).

⁸⁷ Selv om det at skifte status fra "arbejdsløs" i registret ikke er det samme som at være "i beskæftigelse", var en sådan antagelse nødvendig for at kunne gennemføre undersøgelsen.

kovarianser, såsom køn, alder, uddannelsesniveaue, arbejdsløshedsprocent på bopælsstedet.

Analysen viste, at støtten til arbejdsløse over 50 år havde en positiv nettoeffekt på beskæftigelsen: De behandlede arbejdsløse over 50 år havde over 2,5 gange større chance for at afmelde sig som arbejdsløse end en ikkebehandlet. Figuren nedenfor viser udviklingen i outcomevariablen for den behandlede gruppe og kontrolgruppen og den skønnede nettoeffekt efter at have taget højde for den parallelle antagelse i difference-in-differences-tilgangen.



3.2.4. Regressionsdiskontinuitetsdesign

Regressionsdiskontinuitet i hovedtræk

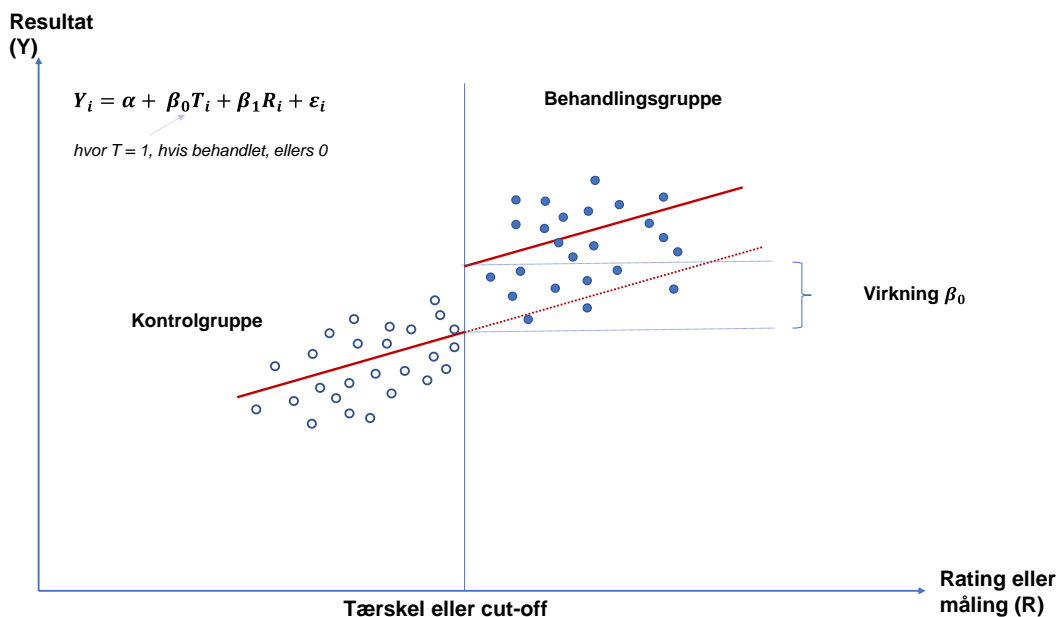
- Et regressionsdiskontinuitetsdesign er baseret på den idé, at en specifik værdi (tærskelværdi eller "cut-off") af en score eller en rating afgør, om en person deltager i interventionen eller ej.
- Personer, der ligger tæt på denne værdi, anses derefter for at være sammenlignelige, med den eneste forskel, at personer på den ene side af tærsklen deltager (behandlingsgruppen), og personer på den anden side ikke deltager (kontrolgruppen).
- En sammenligning af disse grupper giver typisk et præcist og intuitivt mål for interventionens effekt. Metoden kan imidlertid kun bruges sammen med en udvælgelsesproces baseret på en score eller en rating.

Regressionsdiskontinuitet kan anvendes, når adgangen til en intervention bestemmes af et skæringspunkt langs en kontinuerlig rating, skala eller måling. For eksempel kan adgangen til et praktikforløb bestemmes ved resultater i en egnethedsprøve, hvor dem, der scorer over en bestemt tærskel (eller cut-off), vil blive tilbudt praktikforløb, mens dem, der opnår en score under tærsklen, ikke bliver det. For at metoden er gyldig, bør skæringspunktet bestemmes uden at kende de potentielle praktikanters scorere. Ansøgere, der befinder sig umiddelbart omkring skæringspunktet, vil være meget ens, bortset fra at dem, der ligger lige over det, vil deltage i interventionen, mens dem umiddelbart under ikke deltager. Resultaterne for dem over og under

Et skæringspunkt deler behandlings- og kontrolgruppen

skæringspunktet kan sammenlignes for at få et skøn over interventionens virkning ved skæringspunktet.

Figur 10 Illustration af regressionsdiskontinuitetsmetoden



Et regressionsdiskontinuitetsdesign (RDD) kan gennemføres, hvis skæringspunktet enten identificerer behandlingsgruppen fuldstændigt (med fuld overensstemmelse), i hvilket tilfælde der opnås en *skarp* diskontinuitet, eller hvis ikke alle dem på en given side af skæringspunktet under visse betingelser ikke følger deres allokering til behandling ("fuzzy" diskontinuitet).

Skarp eller fuzzy diskontinuitet

Figur 10 ovenfor viser et stiliseret eksempel på et regressionsdiskontinuitetsdesign. Dette er den enkleste illustration af en skarp diskontinuitet. Interventionen har konstante virkninger ved hver værdi af ratingen, og virkningerne vurderes ved hjælp af en lineær regressionsmodel (der er ingen problemer med hensyn til den funktionelle form af virkningsregressionen). I praksis vil analysen altid skulle være betydeligt mere sofistikeret end den, der er vist i figur 10.

Prikkerne i figur 10 repræsenterer individuelle enheder, f.eks. praktikanter. X-aksen registrerer den rating eller det mål, der anvendes til at allokere deltagere til pladser i praktikforløbet. Personer, der scorer til højre for den ubrudte lodrette linje, som angiver tærsklen for denne rating eller dette mål (f.eks. en egnethedsprøve), deltager i praktikforløbet og udgør "behandlingsgruppen". Potentielle deltagere, der scorer under tærsklen for ratingen eller målet, deltager ikke i forløbet og udgør kontrolgruppen.

Det centrale er, at den rating, der anvendes til at tildele målgruppen til enten behandlings- eller kontrolgruppen, er en kontinuerlig kvantitativ variabel målt forud for behandlingen, og at en person får adgang til praktikforløbet baseret på, om vedkommendes score er over eller under et skæringspunkt eller en tærskelværdi, der er fastsat på forhånd.

Resultatet er plottet ind på y-aksen. I det væsentlige identificeres behandlingens virkning ved at anslå en lineær regressionsmodel på dataene (ud fra ovenstående antagelser). Det vil sige regression af resultatvariablen i forhold til ratingmålet sammen med en dummyvariabel (en behandlingsindikator), som definerer, om en score ligger under eller over

skæringspunktet (dvs. om personen allokeres til behandlings- eller kontrolgruppen).

En sådan regressionsligning for virkningen er illustreret i figur 10. Effekten eller virkningen af praktikforløbet i vores eksempel ses i koefficienten på behandlingsindikatoren, dvs. β_0^{88} . Dette viser, om der er et brud eller en diskontinuitet omkring skæringspunktet, som i figur 10 er vist ved et opadgående forløb for regressionslinjen ved tærsklen eller skæringspunktet. I dette eksempel observeres en positiv virkning af praktikforløbet på resultatet.

En alternativ måde at forstå virkningsestimater på er at betragte den stiplede linjeudvidelse til kontrolgruppelinjen som vist i figur 10. Dette kan betragtes som et kontrafaktisk skøn for behandlingsgruppen — forholdet mellem ratingen og resultatmålingen, som ville have været gældende i fravær af interventionen — forskellen mellem denne stiplede linje og tendenslinjen for behandlingsgruppen repræsenterer behandlingseffekten eller -virkningen. Bemærk, at der i fravær af behandling ikke er diskontinuitet i linjen, og vi antager, at resultatet i så fald varierer kontinuerligt med ratingen eller målet.

Regressionsdiskontinuitetsmetoden virker, fordi observationer i behandlings- og kontrolgrupper tæt på skæringspunktet ligner hinanden, bortset fra at dem over skæringspunktet i dette eksempel kommer i praktik, mens dem under ikke gør. Situationen ligner derfor randomisering for observationer tæt på skæringspunktet. Der er dog en betydelig begrænsning. I de fleste anvendelser kan de virkninger, der anslås ved hjælp af RDD-metoden, kun fortælle den politiske beslutningstager om virkningerne for personer, der ligger tæt på skæringspunktet eller tærsklen. Det kan være begrænset, i hvilket omfang der kan foretages generaliseringer af dem, der ligger længere væk fra tærsklen.

RDD kan være en nyttig tilgang, hvor personer allokeres til en intervention på grundlag af behov målt efter en kontinuerlig rating eller score. Analysen kan imidlertid blive kompleks, hvis skæringspunktet er fuzzy, og der er manglende overensstemmelse, og hvor der er problemer med regressionsmodellens funktionelle form. Faktisk er det nødvendigt at anvende en række antagelser, og gyldigheden af disse antagelser kan ikke altid bekræftes.

Den næste tekstboks viser et praktisk eksempel, hvor der blev anvendt en regressionsdiskontinuitetsmetode i en evaluering af en ungdomsgarantiordning.

God intern validitet, men generalisering kan være begrænset

Tekstboks 17 Eksempel på en evaluering baseret på regressionsdiskontinuitet

Erhvervsuddannelse for unge arbejdsløse i Letland: Dokumentation fra et regressionsdiskontinuitetsdesign"⁸⁹

Forskere fra Kommissionens Fælles Forskningscenter anvendte en regressionsdiskontinuitetsstilgang til at vurdere virkningen af et erhvervsuddannelsesprogram på arbejdsmarkedsoutcomes for arbejdsløse unge i Letland, som finansieres under ungdomsgarantiordningen for perioden 2014-2020 og er rettet mod unge mellem 15 og 29 år, som ikke er i beskæftigelse eller under uddannelse (NEET'er).

De anvendte data stammer fra det lettiske statslige arbejdsformidlingskontor, som gav oplysninger om både deltagere og ikkedeltagere, der var registreret som arbejdsløse på bestemte datoer, og de blev matchet med data fra skattemyndighederne, som omfatter oplysninger om personers indkomst på bestemte datoer før og efter programmet (januar 2012 og juni 2017). Efter rensning af datasættene udgjorde den endelige

⁸⁸ I en simpel case ville dette være effekten af behandlingsintentionen ved skæringspunktet (se [Bloom H. S., 2009](#)).

⁸⁹ Se [Bratti M. et al., 2018](#).

stikprøve af den behandlede gruppe 898 individer og kontrolgruppen 10 717.

Evaluatorene støttede sig til en officiel lettisk regel, som giver arbejdsløse under 25 år forrang til programmet. Alder er derfor den løbende variabel, som afgør sandsynligheden for deltagelse i programmet, idet alderen 25 år er den tærskel, under hvilken deltagelsen i programmet stiger drastisk som følge af prioriteringsreglen. Da personer ikke har nogen kontrol over deres alder, kan allokeringen til programmet anses for at være tilfældig omkring tærsklen. Prioriteringsreglen og det forhold, at deltagelse var frivillig, krævede, at forskerne anvendte et fuzzy regressionsdiskontinuitetsdesign (FRDD). Identifikationsstrategien var baseret på, at personer tæt på skæringspunktet (25 år) lignede hinanden, bortset fra deltagelsen i programmet.

Resultaterne viste, at virkningen i form af den sandsynlige fremtidige beskæftigelse og en månedlig indkomst i op til 3,5 år efter deltagelsen i programmet er positiv, men statistisk insignifikant. Der blev dog observeret en positiv virkning af prioriteringsreglen om programdeltagelse. Da validiteten af FRDD er baseret på, at potentielle deltagere ikke kan styre den løbende variabel (alder), blev der gennemført flere test for at bekræfte denne effekt.

3.2.5. Instrumentvariabler

Instrumentvariabler i hovedtræk

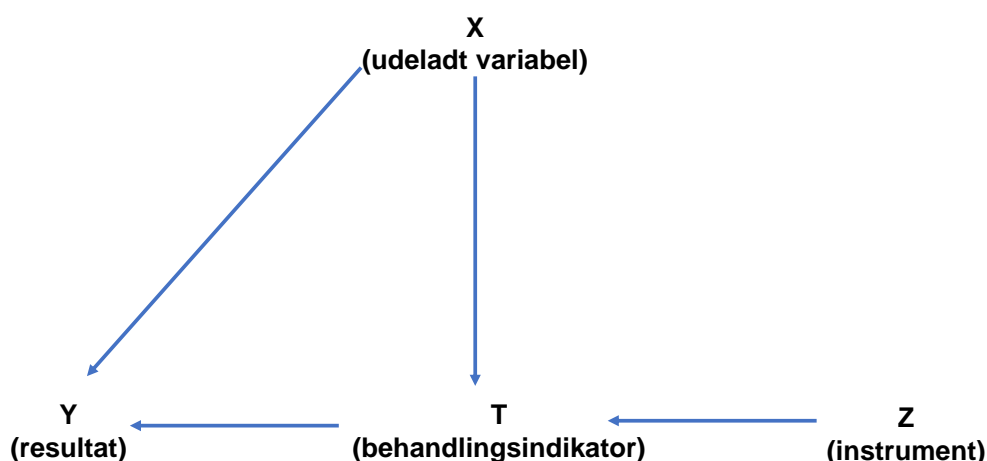
- Udgangspunktet for instrumentvariabler er, at nogle (forudbestemte/eksogene) processer afgør deltagelsen, men det er ikke den egentlige udvælgelsesprocedure.
- Den på forhånd fastlagte proces og dens nøglevariabler kan anvendes til at tage hensyn til enhver form for selvvalg i forbindelse med de enkeltes beslutning om at deltage eller ej.

Ved metoden med **instrumentvariabler** (IV) bør udvælgelsen til behandling i det mindste delvist bestemmes af en eksogen faktor, som ikke har noget at gøre med andre resultater end dem, der opnås gennem behandlingen. Som sådan har den eksogene faktor indflydelse på deltagelsen, men ikke direkte på resultaterne. Sådanne eksogene faktorer kan typisk være administrative fejl eller forsømmelser eller andre tilfældige variationer i den modtagne behandling.

Hvordan eksogene faktorer påvirker deltagelsen

Figur 11 nedenfor viser metoden med instrumentvariabler. Fire variabler vises i et meget forenklet årsagssystem. Variablerne repræsenterer data indsamlet fra en population, der hypotetisk er omfattet af en uddannelsesintervention (både dem, der modtager uddannelse, og dem, der fungerer som kontrol).

Figur 11 Illustration af metoden med instrumentvariabler



"Y" repræsenterer det pågældende resultat. I forbindelse med en uddannelsesintervention kan dette være indkomst. "T" er en indikator, der viser, om en person har deltaget i uddannelse⁹⁰.

"X" er en udeladt variabel, som ikke observeres, men som er knyttet til både "Y" (resultatet) og "T" (behandlingsindikatoren), der udvider idéen om et uddannelsesprogram, f.eks. et basislinjemål for evne. I dette tilfælde hænger evne både sammen med deltagelse i uddannelse og indkomst. F.eks. kan medlemmer af målgruppen med bedre evner vælge at tage en uddannelse og få højere lønninger.

Eksistensen af "X" motiverer søgningen efter et instrument og betyder, at uddannelsens indvirkning på indkomsten hænger sammen med dens eksistens. Med andre ord er der bias i skønnet på grund af forekomsten af X og den omstændighed, at den ikke observeres og ikke kan tages direkte i betragtning i analysen.

Endelig er variabelen "Z" et instrument, der ifølge Morgan og Winship⁹¹ kan betragtes som et chok for "T", som er uafhængigt af "X". Derfor er der ingen linje i figur 11, der forbinder Z med X. Endvidere påvirker Z kun Y gennem T, og der er ingen anden vej, hvorigennem Z kan påvirke Y. Det betyder, at Z kan bruges til at skabe variation i T (behandlingen), som ikke er forbundet med konfounderen X. Derfor kan et mål effekten af T på Y uden bias kun opnås ved at anvende denne variation⁹².

Det er vigtigt at identificere et validt instrument

De enkleste omstændigheder, hvorunder en IV-tilgang kan anvendes, er beskrevet nedenfor, idet der ses bort fra mange af de komplekse forhold, der er forbundet hermed. I praksis er det ofte vanskeligt at finde et overbevisende instrument. Anvendeligheden af forskellige potentielle instrumenter er i høj grad afhængig af konteksten, og de underliggende identificerende antagelser kan generelt ikke testes statistisk (faktisk kan den krævede korrelation mellem Z og T testes statistisk, mens "uafhængigheden" af Z og X ikke kan). En strategi kunne f.eks. være at anvende afstanden fra de centre, hvor uddannelsen afholdes (uddannelsens fysiske placering) til en deltagers hjem, som et instrument til at vurdere virkningen af uddannelse på deltagernes nettoindkomst. Det kan observeres, at deltagere, der bor tættere på

⁹⁰ Med andre ord er der fuld opfyldelse, og alle i behandlingsgruppen deltager i uddannelse.

⁹¹ Morgan S.L., Winship C., 2014.

⁹² Årsagsvirkningen af T på Y beregnes ved forekomsten af et instrument ved at estimere forholdet mellem Z og Y og dividere dette med det estimerede forhold mellem Z og T.

uddannelsesstederne, er mere tilbøjelige til at deltage i en uddannelsesintervention, og at afstanden mellem en deltagers hjem og et uddannelsessted ikke er relateret til andre faktorer for nettoindkomst og deltagelse i uddannelse (f.eks. foranstaltninger vedrørende menneskelig kapital). Den eneste vej, hvorigennem dette afstandsmål kan påvirke nettoindkomsten, er gennem dets indvirkning på uddannelsen⁹³.

Instrumentvariabler kan anvendes i en lang række sammenhænge. Estimerer kan beregnes ved hjælp af forskellige estimeringsmetoder, afhængigt af responsvariablen. Hidtil er denne tilgang ikke blevet anvendt i ESF-evalueringerne. I tekstboks 18 nedenfor illustrerer vi et eksempel på analysen af årsagsvirkningerne af en politik baseret på en IV-tilgang.

⁹³ Fortolkningen af resultaterne af en sådan analyse kan blive mere kompliceret, afhængigt af om instrumentet er forbundet med variation i behandlingseffekten (se [Bryson, et al., 2002](#)).

Tekstboks 18 Eksempel på en evaluering baseret på instrumentvariabler**Beskæftigelsesmæssige virkninger af sprogundervisning for arbejdsløse indvandrere⁹⁴**

Erhvervelse af færdigheder inden for værtslandets sprog er en vigtig faktor for indvandrerne med henblik på at opnå en høj grad af integration på arbejdsmarkedet.

I en nylig rapport (Lang J., 2021) anvender forfatteren en metode med instrumentvariabler til at vurdere årsagsvirkningen af et tysk sprogkursus til erhvervmæssige formål på beskæftigelsesresultaterne for indvandrere, der deltager i kurserne, to år efter tilmeldingen. Programmet blev gennemført af det tyske forbundskontor for migration og flygtninge (BAMF) og finansieret af ESF. ESF-BAMF-programmet var operationelt indtil udgangen af 2017. Kurserne bestod af flere elementer: tyskundervisning, opbygning af faglige kvalifikationer og praktikophold. Personer med indvandrerbaggrund, som var aktive på det tyske arbejdsmarked, var berettiget til at deltage, uanset nationalitet og dato for deres ankomst til Tyskland.

For at afhjælpe den uobserverede heterogenitet med hensyn til sprogfærdigheder anvender forfatteren den eksogene variation i lokal sprogundervisningsintensitet på jobcenterniveau. Dette anses for at være et gyldigt instrument, da jobcentre har vide skønsmæssige justeringer med hensyn til gennemførelsen af forskellige kombinationer af programmer afhængigt af deres "politiske stil", og denne variabel er eksogen i forhold til de jobsøgendes beskæftigelsesoutcomes. Kurser afholdt i 2014 blev analyseret.

Evalueringen bygger på et omfattende sæt administrative data, "Integrierten Erwerbsbiografie" (arbejdsmarkedsregister (IEB)), en fælles database med administrative oplysninger fra det tyske føderale arbejdsmarkedsagentur. IEB indeholder oplysninger om deltagelse i sprogundervisning, detaljerede oplysninger om beskæftigelsehistorik (undtagen selvstændig erhvervsvirksomhed), jobsøgning, overførselsindkomst i arbejdsløshedsperioden og personoplysninger. Andre data (WGH, Werdegangshistorie) blev også anvendt til at få detaljerede oplysninger om deltagelse i sprogkurser leveret af BAMF og til at dække manglende data fra IEB, hvor det var muligt, med oplysninger om selvstændig virksomhed og forældreorlov.

Resultaterne viser, at efter en periode med fastlåsnings effekter, der varer lige over et år generelt, to år efter interventionen, har indvandrere, der deltager i uddannelsesprogrammet, en ni procentpoint højere sandsynlighed for at være i beskæftigelse end kontrolgruppen, og denne sandsynlighed er ikke begrænset til ufaglærte job. Det blev også konstateret, at tidlig sprogundervisning (kort tid efter ankomsten til Tyskland) er gavnlig for integrationen af indvandrere på arbejdsmarkedet.

⁹⁴ Se [Lang, J., 2021](#).

Tabel 5 Sammenligning af hovedtrækkene i de primære metoder til kontrafaktisk konsekvensanalyse

Metode	Hovedtræk	Fordele	Datakrav	Begrænsninger
Randomisering — den eksperimentelle tilgang	<ul style="list-style-type: none"> - Personer, der er berettiget til at deltage, fordeles vilkårligt til en behandlings- og en kontrolgruppe - Randomisering sikrer, at begge grupper er identiske (generelt) med hensyn til alle relevante karakteristika - Kontrolgruppen besvarer derfor det kontrafaktiske spørgsmål, og forskellen i outcomes mellem behandlings- og kontrolgrupperne indikerer foranstaltningens årsagsvirkning 	<ul style="list-style-type: none"> - Hvis de gennemføres korrekt, er skøn over virkningerne "uden bias" - Resultaterne er gennemsigtige og letforståelige - Resultaterne er i mindre grad genstand for forbehold og tvivl - Forskellige designvarianter, der er til rådighed til at tage hensyn til en række politiske sammenhænge og omstændigheder omkring interventionen 	<ul style="list-style-type: none"> - Grundlæggende krav om kontrol af udvælgelse til intervention via randomisering - En fortegnelse over, hvem der er blevet allokeret til hvilken gruppe - Indsamling af basislinjedata er af afgørende betydning - Resultatmålinger skal registreres for både behandlings- og kontrolgrupper 	<ul style="list-style-type: none"> - Ofte opfylder behandlings- og kontrolgruppen reglerne for allokering til behandlingen - Der kræves ofte samtykke fra deltagerne - Randomisering kan påvirke udvælgelsen af dem, der deltager i/ansøger om en intervention - Viden om, hvorvidt de er blevet tildelt behandlingen, kan ændre deltageres adfærd og påvirke resultaterne - Ethiske betænkeligheder - Store krav til planlægning og design - Kan være dyrt (men ikke nødvendigvis)
Matchning (propensity score)	<ul style="list-style-type: none"> - Matchning imiterer randomisering ved at oprette en efterfølgende kontrolgruppe, der så vidt muligt (generelt) ligner behandlingsgruppen med hensyn til alle relevante karakteristika - I modsætning til randomisering kan kun observerbare karakteristika (alder, køn, uddannelse osv.) matches, mens ikkeobserverbare karakteristika (f.eks. motivation) ikke kan tages i betragtning - Metodens validitet afhænger i høj grad af tilgængeligheden af data 	<ul style="list-style-type: none"> - Kræver godt kendskab til udvælgelsesprocesserne, men kræver ikke direkte kontrol med udvælgelsen til interventionen - Kan anvendes retrospektivt, hvis de rigtige data er tilgængelige, og i forskellige sammenhænge - Teknisk set en semiparametrisk estimationsmetode, som kræver færre parameterantagelser (f.eks. ingen standardregressionsantagelse) - Kan bruges til at estimere flere behandlingseffekter 	<ul style="list-style-type: none"> - Nøjagtig identifikation af deltagerne i interventionen - Datakilder, hvorfra der skal udtages stikprøver - Klart koncept for deltagelse og god forståelse af udvælgelse til behandling - Omfattende data, der ideelt set indsamles ved basislinjen for at vælge match - Resultatmål for interventionen for deltagere og kontrol 	<ul style="list-style-type: none"> - Kræver betydelige mængder data, der muliggør en fuldstændig karakterisering af udvælgelsesprocessen - Validiteten afhænger af kontrollernes kvalitet og omhyggelig udvælgelse af dem samt af graden af fælles støtte - Bygger på den antagelse, at udvælgelsen i forbindelse med interventionen i tilstrækkelig grad kan karakteriseres ved observerbare data - De forskellige tilgængelige metoder til matchning kræver en sensitiv analyse - Resultaterne kan være komplekse at forklare og fortolke og potentielt tvetydige
Difference-in-differences (DID)	<ul style="list-style-type: none"> - DID er en intuitiv tilgang, der sammenligner forskellen før og efter i outcomes i behandlingsgruppen med outcomes i kontrolgruppen 	<ul style="list-style-type: none"> - Kontrol af visse aspekter af uobserverede forskelle mellem deltagere og kontroller - Kan anvendes sammen med 	<ul style="list-style-type: none"> - Datakravene svarer til andre metoder, men med det yderligere krav om resultatmål før behandlingen 	<ul style="list-style-type: none"> - Kræver, at der tages udgangspunkt i en antagelse af almindelige tendenser i resultaterne mellem deltagere og kontroller

Metode	Hovedtræk	Fordele	Datakrav	Begrænsninger
	<ul style="list-style-type: none"> - Eftersom ændringen over tid i kontrolgruppen måler, hvad der ville være sket for behandlingsgruppen uden interventionen (det kontrafaktiske scenarie), giver enhver yderligere forskel før/efter i behandlingsgruppens outcomes årsagsvirkningen af interventionen - Dette er en enkel metode, der i mange tilfælde er anvendelig i praksis 	<ul style="list-style-type: none"> matchning - Arbejder med data fra før og efter interventionen, f.eks. paneldata (individuelle data over tid) eller gentagne tværnsnitsdata (data om personer indsamlet på forskellige tidspunkter) 	<ul style="list-style-type: none"> - For at teste de vigtigste forudsætninger er der behov for flere observationer af resultater før behandlingen for både behandlings- og kontrolgrupper 	<ul style="list-style-type: none"> - Analysen kan blive temmelig kompleks og åben for fejlfortolkning - Omfattende data om tiden før behandlingen om resultater, der er nødvendige for at teste antagelsen om almindelige tendenser - Kan bruges til at estimere flere behandlingseffekter⁹⁵
Regressionsdiskontinuitetsdesign (RDD)	<ul style="list-style-type: none"> - RDD er baseret på den idé, at en specifik værdi (tærskelværdi eller "cut-off") af en score eller en rating afgør, om en person deltager i interventionen eller ej - Personer, der ligger tæt på denne værdi, anses derefter for at være sammenlignelige, med den eneste forskel, at personer på den ene side af tærsklen deltager (behandlingsgruppen), og personer på den anden side ikke deltager (kontrolgruppen) - En sammenligning af disse grupper giver typisk et præcist og intuitivt mål for interventionens effekt. Metoden kan imidlertid kun bruges sammen med en udvælgelsesproces baseret på en score eller en rating 	<ul style="list-style-type: none"> - Der findes både skarpe og fuzzy RDD-metoder - Kan give objektive vurderinger af behandlingseffekter på visse betingelser 	<ul style="list-style-type: none"> - Valget af skæringspunkt skal være uafhængig af værdierne for den rating, der gives til hvert medlem af en målgruppe - Der kræves data om enkeltpersoner for både behandlings- og kontrolgrupper med hensyn til rating eller mål, tærskelværdi eller skæringspunkt og resultater 	<ul style="list-style-type: none"> - Denne metode er ikke valid uden et kontinuerligt mål eller en kontinuerlig rating, der afgør behandlingen - Analyser kan blive komplekse og usikre, hvis problemer med den funktionelle form af virkningsregressionen bliver fremtrædende, hvor der er manglende opfyldelse, eller hvor stikprøvens størrelse omkring skæringspunktet er begrænset - Der kan være risici forbundet med fortolkning af resultater og ekstrapolering af generaliseringer
Instrumentvariable (IV)	<ul style="list-style-type: none"> - Ideen bag IV er, at nogle (forudbestemte/eksogene) processer afgør deltagelsen, men det er ikke den egentlige udvælgelsesprocedure - Den på forhånd fastlagte proces kan anvendes til at tage hensyn til enhver form for selvvalg i forbindelse med de enkeltes faktiske beslutning om at deltage eller ej 	<ul style="list-style-type: none"> - Kan give estimater af høj kvalitet eller dokumentation for forekomsten af årsagsvirkninger - Løser problemet med bias pga. udeladte variabler (eller selektionsbias) - Kan anvendes retrospektivt 	<ul style="list-style-type: none"> - Kræver basislinjedata, data om resultater og deltagelse i interventionen, og desuden at et instrument kan identificeres - Et instrument skal være knyttet til deltagelse i interventionen og påvirker kun resultaterne på dette grundlag. Instrumentet bør ikke være forbundet med andre determinanter for resultater 	<ul style="list-style-type: none"> - Kan være vanskeligt at finde et plausibelt instrument - Kan være svært at forklare over for ikkeeksperter - Det er ikke ligetil at fortolke resultaterne. Begrænset mulighed for at teste de identificerende antagelser

⁹⁵ Se [Frolich, M., 2004.](#)

Kapitel 4. Kontrafaktiske konsekvensanalyser — de næste skridt

Denne vejledning har til formål at tilskynde forvaltningsmyndighederne til at udføre flere kontrafaktiske konsekvensanalyser af god kvalitet og støtte dem i dette arbejde. Den giver vejledning til dem, der er ansvarlige for at planlægge og iværksætte konsekvensanalyser af ESF+-medfinansierede interventioner. Hidtil har der været fokus på planlægning og gennemførelse af en kontrafaktisk konsekvensanalyse, og en række centrale spørgsmål, der kræver overvejelse, er blevet drøftet. Programmeringsperioden 2014-2020 fremhævede imidlertid en række andre "bredere spørgsmål" samt de udfordringer, der skal tackles for at styrke ESF+-evalueringerne og anvendelsen af kontrafaktiske konsekvensanalyser.

Denne del af vejledningen indeholder nogle forslag til, hvordan disse "bredere" spørgsmål kan håndteres. Det drøftes navnlig, hvilke skridt der skal tages for at tage hånd om følgende:

Bredere spørgsmål, der skal overvejes

- manglende viden om tilgange til kontrafaktisk konsekvensanalyse inden for forvaltningsmyndighederne og blandt medlemsstaternes politiske beslutningstagere mere generelt
- mangel på eksterne, tilstrækkeligt kvalificerede og erfarne leverandører i medlemsstaterne, som er i stand til at udføre kontrafaktiske konsekvensanalyser
- håndtering af juridiske hindringer, der skal tackles generelt på tværs af kontrafaktiske konsekvensanalyser
- fremskridt i retning af bedre planlægning af kontrafaktiske konsekvensanalyser
- udvidelse af anvendelsesområdet for kontrafaktiske konsekvensanalyser

4.1. Bedre forståelse blandt interessenterne

For programmeringsperioden 2021-2027 fastsættes det i artikel 44, stk. 1, i forordningen om fælles bestemmelser, at "[...] forvaltningsmyndigheden foretager evalueringer af programmerne vedrørende et eller flere af følgende kriterier: effektivitet, produktivitet, relevans, sammenhæng og merværdi på EU-plan med det formål at forbedre kvaliteten af udformningen og gennemførelsen af programmerne. Evalueringerne kan også omfatte andre relevante kriterier, såsom inklusion, ikkeforskelsbehandling og synlighed, og kan omfatte mere end ét program". Som beskrevet i de foregående afsnit betyder dette, at der skal være en passende evalueringskapacitet til rådighed i både den forvaltningsmyndighed, der iværksætter en kontrafaktisk konsekvensanalyse, og hos de evaluatore, der ansøger om kontrakten.

I nogle tilfælde registrerede Evaluation Helpdesk en mangel på kapacitet til at udføre kontrafaktiske konsekvensanalyser — dvs. udbud og støtte til gennemførelsen — i nogle forvaltninger. Dette gjorde det vanskeligt for evaluatore at gennemføre kontrafaktiske konsekvensanalyser, fordi der ikke

på forhånd var blevet sikret klare evalueringsspørgsmål, grundlæggende datatilgængelighed og velunderbygget gennemførelsesplanlægning.

Nogle gange, især i små lande, mangler der kapacitet til at gennemføre en kontrafaktisk konsekvensanalyse på rådgivningsmarkedet, og det samme gælder levering af teknisk ekspertise. Der er derfor et stort behov for at stimulere både efterspørgslen efter og udbuddet af kontrafaktiske konsekvensanalyser. Udbuddet kan øges, efterhånden som forvaltningsmyndighederne og medlemsstaterne begynder at iværksætte kontrafaktiske konsekvensanalyser eller begynder at italesætte deres behov for sådanne undersøgelser. Hvor hurtigt der reageres på den øgede efterspørgsel efter kontrafaktiske konsekvensanalyser, afhænger af allerede eksisterende færdigheder, erfaring og institutioner i medlemsstaterne, der er i stand til at foretage sådanne undersøgelser. Stimulering af efterspørgslen kan imidlertid også delvis opnås ved at forbedre kendskabet til og forståelsen af metoder til kontrafaktisk konsekvensanalyse blandt dem, der arbejder i forvaltningsmyndighederne.

Stimulering af efterspørgslen efter og udbuddet af kontrafaktiske konsekvensanalyser

En løsning på dette problem er, at forvaltningsmyndighederne afholder kurser i metoder til kontrafaktisk konsekvensanalyse for deres personale. Kurserne bør fokusere på fordelene for forvaltningsmyndighederne ved at indføre disse metoder. Desuden bør der lægges vægt på spørgsmål om ansvarlighed og at finde ud af, hvad der virker. Bilag 2 indeholder et forslag til kursusbeskrivelse.

Afholdelse af kurser i metoderne

4.2. Kapacitetsudvikling

Et andet spørgsmål, der blev rejst i perioden 2014-2020, og som er nævnt i afsnittene ovenfor, var behovet for at udvikle kapacitet til at udføre kontrafaktiske konsekvensanalyser hos medlemsstaternes forskere, akademiske miljøer og konsulenter. I nogle tilfælde stod det klart, at de færdigheder, der var nødvendige for at gøre dette, var tilgængelige i medlemsstaterne, men at de pågældende personer havde haft vanskeligt ved at anvende dem i en evalueringssammenhæng (f.eks. begrænset adgang til brugbare data eller problemer med at identificere en rimelig kontrolgruppe).

Der kan tages en række skridt til at udvikle udbuddet af evalueringstjenester. Mange af de rejste spørgsmål gælder såvel kontrafaktiske konsekvensanalyser som evalueringer mere generelt. Der kan normalt tages tre skridt til at styrke udbuddet af evalueringer:

Styrkelse af institutioner og opbygning af praksisfællesskaber

- opbygge forbindelser med akademiske institutioner, navnlig universiteter
- udvikle og styrke et uafhængigt fællesskab af konsulenter
- støtte udviklingen af et professionelt evalueringsmiljø.

Universiteter

Det er vigtigt at udvikle forbindelser til universiteter af to grunde. For det første kan akademisk personale ved universiteterne besidde de færdigheder og den viden, der er nødvendig for at gennemføre kontrafaktiske konsekvensanalyser. For eksempel har mange mikroøkonomer, økonomologer, kvantitative sociologer eller psykologer de nødvendige kvalifikationer. I mange medlemsstater kan de nødvendige færdigheder være til rådighed, men de, der har dem, har ikke tidligere tænkt på at anvende dem

Udvikling af akademiske færdigheder

til evaluering af interventioner. Der kan være et manglende incitament for dem til at gøre det, hvilket der skal gøres noget ved.

I nogle medlemsstater er der tradition for, at akademiske forskere deltager aktivt i anvendt politisk forskning. I denne sammenhæng vil akademikere være vant til at arbejde med regeringer og forvaltningsmyndigheder. I andre medlemsstater, hvor universiteter og akademikere ikke beskæftiger sig lige så meget med anvendt forskning, kan der være behov for en kulturændring. En vellykket metode til at udvikle et leverandørgrundlag inden for universitetssektoren er, at medlemsstaternes myndigheder og forvaltningsmyndighederne kan finansiere udgifterne til forskningscentre, der beskæftiger sig med metoder til konsekvensanalyse.

Universiteter og akademikere kan også spille en rolle i uddannelsen af den næste generation af evaluatore. I tæt samarbejde med universiteterne kan det være muligt at tilskynde dem til at medtage metoder til programevaluering i deres studieplaner og som led i denne udvikling sikre, at metoderne til kontrafaktiske konsekvensanalyser er omfattet af undervisningsprogrammerne. I nogle medlemsstater kan universiteter også spille en rolle i forbindelse med løbende faglige udviklingskurser om konsekvensvurdering og kontrafaktisk konsekvensanalyse. Dette kan rettes mod politiske beslutningstagere, tekniske specialister inden for forvaltningsmyndighederne samt andre potentielle leverandører såsom uafhængige konsulenter. Medlemsstaterne kan overveje at yde finansiering til en sådan uddannelse.

Uddannelse af næste generation

Uafhængige konsulenter

For nogle omfattende evalueringer findes der et internationalt marked. Dette er bestemt tilfældet for store kontrafaktiske konsekvensanalyser. Mange medlemsstater ønsker imidlertid at udvikle national kapacitet til at gennemføre kontrafaktiske konsekvensanalyser. Én strategi til at opnå dette kan være ved at etablere strategiske alliancer mellem potentielle indenlandske leverandører og internationale konsulentfirmaer.

Udvikling af markedet

Nedenfor findes en række forslag til udvikling af et indenlandsk leverandørgrundlag til at foretage kontrafaktiske konsekvensanalyser, som kan anvendes af forvaltningsmyndigheder (eller andre organer), der iværksætter kontrafaktiske konsekvensanalyser:

- kræve, at bud fra konsortier eller partnerskaber altid skal omfatte visse lokale konsulenter
- skalere evalueringskontrakter på en måde, så nye, nationale aktører på evalueringsmarkedet kan foretage relativt små evalueringer med lav risiko
- sikre, at de tekniske og finansielle krav, der er forbundet med at byde på evalueringer, ikke er for restriktive, og gøre det muligt for nye leverandører at deltage
- fremhæve tekniske kriterier og kriterier for knowhow i forbindelse med udvælgelsen snarere end komplekse administrative procedurer, som mindre erfarne konsulenter måske ikke er bekendt med
- afholde informationsmøder med potentielle konsulenter for at besvare spørgsmål og tilskynde dem til at afgive bud i et konkurrencepræget miljø
- støtte netværkssamarbejde mellem relativt isolerede evalueringskonsulenter med henblik på at fremme opbygning af teams,

dannelse af konsortier og andre faglige netværk og sammenslutninger, også på internationalt plan

- anerkende hos den myndighed, der har bestilt analysen, at det kan være nødvendigt at anvende en mere praktisk ledelsestilgang for nye leverandører for at fremskynde deres erhvervelse af viden og erfaring.

Tekstboks 19 Eksempel på et projekt, der har til formål at styrke en kultur og kapacitet for kontrafaktiske konsekvensanalyser

I programmeringsperioden 2014-2020 blev der gennemført et ESF-projekt i Sverige med det formål at forbedre de offentlige arbejdsformidlingers kultur og evalueringskapacitet. Projektet, kaldet "evidensbaserede EU-finansierede projekter", løb fra 2016 til 2018 og involverede ledere og andet personale fra offentlige arbejdsformidlinger i workshopper og andre uddannelsesaktiviteter inden for evaluering.

Aktiviteterne fokuserede ikke på tekniske spørgsmål, men på betydningen af en evaluering, de vigtigste trin i en konsekvensanalyse, vigtigheden af gode data og kvalitetssikring af en evaluering. Der blev oprettet et websted, og undervisningsmateriale blev uploadet hertil.

De vigtigste resultater af projektet var styrkelse af ledernes viden om evalueringer, deres åbenhed over for at foretage evalueringer og deres samarbejdsvilje. Konkret svarede den interviewede person, at "projekternes indvirkning på antallet af gennemførte konsekvensanalyser" endnu ikke er stor, men der er sket tydelige fremskridt i betragtning af, at der i den foregående programmeringsperiode 2007-2013 ikke blev gennemført nogen kontrafaktiske konsekvensanalyser i Sverige.

Evalueringen af projektet "Ung fremtid", som anvendte en randomiseret tilgang, er et konkret eksempel i praksis, som inddrog ledere fra de offentlige arbejdsformidlinger i evalueringsprocessen⁹⁶.

Fagligt fællesskab

Det er vigtigt at oprette faglige evalueringsgrupper i medlemsstaterne med plads til drøftelse af metoder til kontrafaktisk konsekvensanalyse og udveksling af erfaringer. Udviklingen af faglige fællesskaber er vigtig for gensidig støtte og læring samt for opretholdelse af kvalitetsstandarder. En nyttig strategi kunne være at udvikle forbindelser med de relevante nationale evalueringssammenslutninger og tilskynde dem til at fremme kontrafaktiske konsekvensanalyser i ESF+-interventioner i forbindelse med kurser, specifikke konferencer eller seminarer eller oplysningsmøder.

Udvikling af faglige fællesskaber

Deling af erfaringer

Kommissionen har et ønske om, at der udføres grundigere konsekvensanalyser af ESF+, og der er i vid udstrækning blevet slået til lyd for kontrafaktiske konsekvensanalyser i perioden 2014-2020⁹⁷. Det er for nuværende ikke muligt at bekræfte, at anvendelsen af kontrafaktiske konsekvensanalyser er steget, men de anvendes stadig i et begrænset omfang i en række forvaltningsmyndigheder og medlemsstater. Udveksling af erfaringer med metoderne til kontrafaktisk konsekvensanalyse er et af de vigtigste midler til at udvikle kapacitet og støtte udbredelsen af kontrafaktiske konsekvensanalyser i hele EU. Eksisterende fora for gensidig læring såsom peerevalueringer af politikker for beskæftigelse og social inklusion og

Anvendelse af eksisterende fora

⁹⁶ Oplysningerne er baseret på et interview med en embedsmand fra den svenske Arbetsförmedlingen.

⁹⁷ I forordningen om fælles bestemmelser for perioden 2014-2020 (forordning (EU) nr. 1303/2013) — bilag XI kræves det, at der foreligger "et system af resultatindikatorer, som er nødvendige for at vælge de tiltag, der mest effektivt bidrager til at opnå de ønskede resultater, overvåge fremskridt mod resultater og foretage en evaluering af virkningerne". Dette krav var en del af forhåndsbetingelserne for programmeringsperioden 2014-2020.

praksisfællesskaber inden for ESF+ bør anvendes til dette formål. Initiativer fra individuelle forvaltningsmyndigheder eller medlemsstater, såsom internationale konferencer eller seminarer, kan også fremme udveksling af erfaringer.

4.3. Håndtering af juridiske hindringer

Et af de største og væsentligste problemer, som forskere, der udfører kontrafaktiske konsekvensanalyser, støder på i medlemsstaterne, er at få adgang til data. Navnlig møder forskere ofte juridiske begrænsninger, der har til formål at beskytte fortroligheden for de personer, der er repræsenteret i datasæt. Som det fremgår af ovenstående med hensyn til den generelle forordning om databeskyttelse, ligger svaret på disse spørgsmål i at gennemføre bredere reformer og indgå aftaler, der gør det muligt løbende at stille relevante data til rådighed for evaluatorene på en kontrolleret måde.

Fjernelse af juridiske hindringer for dataadgang

Analytiske versioner af administrative datasæt kan udarbejdes regelmæssigt på grundlag af administrative data, som medlemsstaternes myndigheder ligger inde med, dokumenteres og opbevares i et arkiv med kontrolleret adgang. Godkendte leverandører kan udtrække data fra sådanne arkiver med tilladelse. Hvis det ikke er muligt at opnå specifikt samtykke, vil data blive fuldt anonymiseret med krypterede personlige identifikatorer for at overholde den generelle forordning om databeskyttelse. Der er oprettet dataarkiver som dette i flere lande. Hvis adgang imidlertid stadig er en hindring på grund af forskellige fortolkninger af reglerne om beskyttelse af privatlivets fred, bør et nationalt initiativ på regeringsniveau fremme aftaler og systemer, der kan levere data til forskningsformål på relativt kort tid.

Oprettelse af analytiske datasæt

Hvis der fortsat er betænkeligheder med hensyn til fortroligheden af personoplysninger, bør det overvejes at oprette datalaboratorier. Her vil evaluatorene, der arbejder med administrative datasæt, kun få adgang til registre på sikre steder, hvor adgangen til data overvåges og kontrolleres nøje. Data skal behandles og analyseres på stedet, og kun resultaterne af eventuelle analyser kan forlade lokalerne.

Oprettelse af datalaboratorier

4.4. Udvikling i retning af mere fremadrettede tilgange

Et fælles træk ved det begrænsede antal kontrafaktiske konsekvensanalyser, der hidtil har været gennemført af ESF-finansierede interventioner, er, at de har været retrospektive snarere end fremadrettede. Det betyder, at ekspertevaluatorer er blevet bedt om at foretage evalueringer af interventioner, der er blevet udviklet uden at tage hensyn til evalueringen, og hvor der i nogle omstændigheder kun har været begrænset eller slet ingen planlægning af en konsekvensanalyse. Det betyder, at evaluatorene har måttet konstruere datakilder på tidskrævende, dyre og ikke så optimale måder, idet de har anvendt de data, der tilfældigvis har været tilgængelige, i stedet for datakilder, der er konstrueret med henblik på konsekvensvurdering.

En fremadrettet tilgang ville derimod inddrage evaluatorene i planlægningen af en kontrafaktisk konsekvensanalyse så tidligt som muligt og ville gøre det muligt at påvirke interventionerne (enten nye eller eksisterende), ofte på ret subtile måder, og dermed gøre dem mere egnede til kontrafaktisk konsekvensanalyse. Planlægning af en kontrafaktisk konsekvensanalyse på forhånd kan gøre forskellen mellem at kunne foretage en grundig evaluering

Fremadrettede tilgange inddrager evaluatorene fra begyndelsen

eller slet ikke være i stand til at foretage en. Inddragelse af tilstrækkeligt uddannet internt personale eller eksterne eksperter tidligt i en interventions levetid, eller når der træffes finansieringsbeslutninger, betyder, at:

- passende registrering kan integreres i gennemførelsen af programmer og interventioner
- de nødvendige datakilder kan identificeres tidligt, og adgangs- og databeskyttelsesspørgsmål kan behandles i god tid
- indsamling af basisdata kan specificeres, og der kan gennemføres undersøgelser, hvis det er nødvendigt
- praktiske spørgsmål i forbindelse med, hvordan deltagerne rekrutteres til interventioner, kan løses på en sådan måde, at rekrutteringsprocesserne i højere grad er i overensstemmelse med en streng evaluering.

Inddragelse af evaluatore, der er uddannet i kontrafaktiske konsekvensanalyser (interne evaluatore i forvaltningsmyndighederne eller eksterne eksperter) i processen med at udvikle nye ESF+-interventioner eller i beslutninger vedrørende eksisterende interventioner, vil gøre det muligt at påbegynde planlægningen af konsekvensanalyser i begyndelsen af programperioden og høste betydelige fordele i forbindelse med evidensbaseret politisk beslutningstagning.

4.5. Udvidelse af anvendelsesområdet for kontrafaktiske konsekvensanalyser

Et sidste, meget vigtigt spørgsmål vedrører den politiske anvendelse af kontrafaktiske konsekvensanalyser og omfatter forskellige dimensioner: De omfattede politikker, analyseoutcomes og fuldstændigheden af fortolkningen af resultaterne.

I perioden 2014-2020 fokuserede kontrafaktiske konsekvensanalyser næsten udelukkende på aktiv arbejdsmarkedspolitik og beskæftigelsesrelaterede virkninger. Dette kan sandsynligvis forklares ved, at administrative data om beskæftigelse generelt er lettere at få adgang til, og disse data udgør den vigtigste outcomevariabel for analysen af mange aktive arbejdsmarkedspolitikker.

Politikker for uddannelse og social inklusion har ikke eller kun meget sjældent været underkastet en kontrafaktisk konsekvensanalyse på trods af deres betydning for ESF's strategi og tildeling af midler. Nogle af de få eksempler fra Spanien, Polen og Portugal er anført i tekstboks 20 nedenfor.

For at modvirke denne mangel på kontrafaktiske konsekvensanalyser og forbedre kendskabet til disse vigtige politikker er der behov for en række kombinerede initiativer:

- Planlægning og forberedelse af kontrafaktiske konsekvensanalyser om social inklusion eller uddannelsesforanstaltninger på forhånd med henblik på at vurdere tilgængeligheden af data i tide, aktivering af det nødvendige samarbejde mellem forskellige aktører for at øge denne tilgængelighed og oprettelse af behandlings- og kontrolgrupper på et tidligt tidspunkt.
- Fremme af dataejerers inddragelse i udformningen af kontrafaktiske konsekvensanalyser. En bredere inddragelse i de kontrafaktiske konsekvensanalyser af de forvaltninger, der er ansvarlige for uddannelse

Større fokus på kontrafaktiske konsekvensanalyser i politikkerne for social inklusion og uddannelse

Inddragelse af yderligere aktører og udvikling af passende datakilder

og socialpolitik, hvor de adskiller sig fra forvaltningsmyndigheden, kan skabe en stærkere og mere generel vilje til at evaluere og stille data til rådighed.

- Stimulering af institutionelle aftaler og softwareværktøjer for at sikre, at administrative data om sociale forhold og uddannelse kan anvendes til kontrafaktiske konsekvensanalyser. Med henblik herpå kan det være vigtigt at inddrage det nationale statistiske kontor som "brobygningsinstitution" og formidler, da disse kontorer har de nødvendige færdigheder og allerede udarbejder mange administrative datasæt med henblik på udarbejdelse af nationale statistikker.

Støtte til relevant dataindsamling fra deltagere i behandlings- og kontrolgrupper ved begyndelsen af interventionen, så der er et konsistent datasæt "før og efter". Hvis administrative data ikke foreligger eller er relevante, kan en specifik undersøgelse give de nødvendige data, men skal starte ved begyndelsen af interventionen og også inddrage kontrolgruppen for at sikre de nødvendige sammenligninger.

Tekstboks 20 Eksempler på evalueringer på uddannelsesområdet

ESF-interventioner mod skolefrafald gennemført i Asturien⁹⁸

Under en generel evaluering af det operationelle program Principado de Asturias ESF 2014-2020 blev der foretaget en specifik konsekvensanalyse af Diversificación Curricula y de Mejora del Aprendizaje y del Rendimiento (PMAR). PMAR er en ESF-foranstaltning, der gennemføres under IP10.i, og som har til formål at bekæmpe skolefrafald. Under PMAR er eleverne inddelt i specifikke grupper (fra 8 til 15 elever, men af forskellig størrelse, afhængigt af omstændighederne) for at studere inden for lingvistik, samfundsvidenskab, naturvidenskab og matematik samt fremmedsprog (de resterende fag ligger på stamholdet). Foranstaltningen blev gennemført i det akademiske år 2016-2017 på det andet og tredje år af sekundærtrinnet (ESO). Elever, der har taget mindst ét fag om på et hvilket som helst tidspunkt, og dem, der har gennemført det 1. ESO, men som ikke kan gå videre til næste år, er alle berettigede til at deltage. Den endelige udvælgelse foretages på grundlag af den individuelle vurdering af studerende (akademisk og psykopedagogisk), der foretages af lærerne.

I alt deltog 1 053 elever, 512 fra det andet år af ESO og 541 fra det tredje år. Disse elever blev sammenlignet med andre med indlæringsvanskeligheder⁹⁹ med henblik på at udvælge kontroller, der med hensyn til kriterier var sammenlignelige med de studerende i behandlingsgruppen (i alt 3 852 personer).

Der blev anvendt en matchningsmetode ved hjælp af de administrative data fra SAUCE-databasen (Sistema Informático para la Administración Unificada de Centros Educativos) i anonymiseret form fra undervisningsministeriet. Disse data blev anvendt til at beregne deres propensity score¹⁰⁰ og måle outcomevariablen, som er oprykning til næste akademiske år. Evalueringen viste en positiv og statistisk signifikant effekt af interventionerne: Elever på det andet år i behandlingsgruppen har en 18 procentpoint højere succesrate end kontrolgruppen, og for tredjeårselever er forskellen ca. 16 procentpoint. I begge grupper er virkningen mere positiv for kvinder end for mænd.

ESF-interventioner til forbedring af erhvervsuddannelserne i Podlaskie¹⁰¹

Undersøgelsen undersøgte virkningerne af projektet "Good Profession — Great Life", som havde til formål at fremme erhvervsuddannelse i Podlaskie-regionen i Polen. Projektet var rettet mod personalet i grundskolen, elever og forældre, og der blev gennemført en række aktiviteter både for institutioner og

⁹⁸ Se [Diaz J.M. et al., 2019](#).

⁹⁹ Disse blev identificeret som studerende, der i det akademiske år 2015-2016, var på det første og andet år af sekundærtrinnet, som ikke havde bestået et eller flere fag, og som havde gået om mindst én gang i deres skoletid.

¹⁰⁰ Følgende variabler blev anvendt til at matche de to grupper: akademiske resultater i det foregående akademiske år, skoletype (offentlig eller privat), deltagelse på anden supplerende uddannelse i det foregående akademiske år (ja eller nej), køn, gennemsnitlig indkomst i bopælskommunen, bopæl, fødeland (Spanien eller resten af verden).

¹⁰¹ Se [Zub M. et al., 2020](#).

mennesker (skræddersyede markedsførings- og kommunikationsstrategier for skoler, uddannelse af lærere, samarbejde med virksomheder og lokale ledere i industrien, mentorordninger for elever og forældre, erhvervsfaglig og uddannelsesmæssig rådgivning for elever). Projektet blev gennemført fra 2017 til 2019.

Den behandlede population bestod af elever, der deltog i det evaluerede projekt. Ud af 9 403 elever blev 1 500 udvalgt ved stratificeret tilfældig stikprøvetagning, og 200 elever deltog i spørgeskemaet i undersøgelsen. Oplysninger om elever, der deltog i projektet, blev hentet fra overvågningsdataene. Kontrolgruppen blev angivet som elever med karakteristika, der svarede til elever i den behandlede gruppe, som ikke deltog i projektet. En kontrolgruppe på ca. det dobbelte af behandlingsgruppens antal blev planlagt og udvalgt tilfældigt fra det nationale folkeregister (PESEL — et register, der tildeler hver enkelt polsk borger et individuelt nummer). På grund af vanskeligheder forårsaget af covid-19 kunne den planlagte metode imidlertid ikke gennemføres. I sidste ende blev sneboldsteknikken (kædehenvisning) anvendt som en primær stikprøvemetode, hvilket resulterede i 401 interviews, hvoraf 384 blev anvendt i undersøgelsen.

Data om outcomevariabler¹⁰² for både behandlings- og kontrolgrupperne blev indsamlet via spørgeskemaundersøgelser ved hjælp af personlige interview, telefonsamtaler eller webinterview. Undersøgelsen anvendte vægtning af propensity score som kontrafaktisk metode til at vurdere virkningen (kun på et generelt niveau, ikke for specifikke undergrupper). De kontrafaktiske analyser, der blev foretaget i undersøgelsen, gav ikke entydige resultater. Resultaternes statistiske signifikans var utilfredsstillende.

Evaluering af systemet for tilskud til videregående uddannelser for mindre privilegerede studerende i Portugal¹⁰³

I en anden evaluering, som endnu ikke er offentliggjort officielt, undersøges et tilskudsprogram for studerende fra lavindkomsthusholdninger, der har til formål at fremme adgangen til videregående uddannelse og øge deltagelsen. Tilskuddet støttes af ESF (siden 2007-2013) i den nordlige og centrale region og i Alentejo-regionen og af staten i de øvrige regioner. Berettigede studerende er studerende med midler under en bestemt tærskel (indkomst på mellem 6 800 og 7 900 EUR pr. person) og et minimumsantal point, der er opnået i det foregående studieår for studerende, der ikke er på deres første studieår. De samme regler for deltagelse gælder, uanset om de er tilmeldt offentlige eller private universiteter. Tilskudsbeløbet er proportionalt med husstandsindkomsten.

Siden 2011 har programmet hvert år haft deltagelse af ca. 70 000 studerende, men evalueringen dækker perioden fra 2012 og frem på grund af tilgængeligheden af data fra dette år. Evalueringen fokuserer desuden på studerende, der ansøger om tilskuddet for første gang og er indskrevet på deres første år af en uddannelse (bachelor eller master). Ret til deltagelse afhænger kun af indkomstkriteriet¹⁰⁴.

Der blev anvendt to administrative datasæt: data om studerende, der ansøgte mellem 2012 og 2018 (leveret af generaldirektoratet for videregående uddannelse i Portugal (DGES)), som blev lagt sammen med en entydig identifikator med et andet datasæt med oplysninger om uddannelsesforløb og fremskridt (leveret af generaldirektoratet for statistik og videnskab (DGEEC)). Den endelige stikprøve, der blev analyseret, består af 156 002 studerende, hvoraf 130 602 blev behandlet, og 25 400 ikke blev det. Indkomsttærsklen (den løbende variabel) gjorde det muligt for evaluatoren at anvende et regressionsdiskontinuitetsdesign, og der blev målt flere outcomevariabler: På kort sigt den procentdel af de studerende, der stadig var indskrevet ved udgangen af det første år, og de point, der var opnået ved udgangen af det første år. På længere sigt sandsynligheden for at gennemføre uddannelsen, det opnåede resultat og antallet af år til gennemførelse.

I den indledende periode viser tilskuddet positive virkninger i form af indskrivningsprocent, men ikke i form af opnåede resultater. I den længere periode bliver de studerende i højere grad færdig med uddannelsen, på kortere tid og med højere karakterer end kontrolgruppen. Virkningerne var større for mænd, studerende bosiddende i områder, der støttes af ESF, og studerende, der er indskrevet på en kandidatuddannelse og på statslige universiteter.

¹⁰² Følgende outcomevariabler (baseret på svar i spørgeskemaundersøgelsen) blev anvendt: indskrevet på en teknisk skole på tidspunktet for undersøgelsen, indskrevet på en erhvervsskole på første trin på tidspunktet for undersøgelsen, indskrevet på enten en teknisk skole eller en erhvervsskole på første trin på tidspunktet for undersøgelsen, ansøgning til et gymnasium, ansøgning til teknisk skole, ansøgning til erhvervsskole på første trin, ansøgning til skole, der tilbyder erhvervsuddannelse (enten en teknisk skole eller en erhvervsskole på første trin), skolen opfylder elevens forventninger i meget høj grad, erhverv, der svarer fuldstændig til elevens interesser.

¹⁰³ Se Guthmuller S., Meroni E.C., endnu ikke offentliggjort.

¹⁰⁴ Der er planer om at analysere den fulde stikprøve af studerende (ikke kun dem på første år), men denne er endnu ikke afsluttet.

Det er tidligere blevet understreget, at metaevalueringerne er nyttige og gode til at give forklaringer. En generel stigning i antallet af kontrafaktiske konsekvensanalyser og dermed en mulig stigning i metaevalueringerne ville styrke debatten om specifikke foranstaltninger og samtidig inddrage flere forvaltningsmyndigheder og medlemsstater i denne debat. Som nævnt ovenfor skal disse evalueringers metodologiske kompleksitet overvindes med forbedret planlægning af kontrafaktiske konsekvensanalyser på nationalt og internationalt plan.

*Flere
metaevalueringer*

I ESF-interventioner målte kontrafaktiske konsekvensanalyser generelt effekter med kvantitative eller binære (ja/nej) variabler såsom indkomst- eller beskæftigelsesstatus. "Bløde outcomes", der vedrører deltageres selvopfattelse og evner, er imidlertid vigtige mål for en vellykket intervention. F.eks. er "beskæftigelsesegnethed", selvværd eller erhvervede kompetencer en forudsætning for at finde et job eller for at være aktiv på arbejdsmarkedet. I mange sammenhænge ville analyser af bløde outcomes være endnu mere nyttige end tallene for tilbagetrækning fra arbejdsmarkedet og ville kaste lys over matchningen mellem interventioner og individuelle behov.

*Overvejelser om
"bløde outcomes"*

Kontrafaktiske konsekvensanalyser kan anvendes på disse bløde outcomes på betingelse af, at oplysningerne indsamles før og efter interventionen i både behandlings- og kontrolgrupper i overensstemmelse med interventionens forandringsteori. Disse undersøgelser skal planlægges tidligt og involvere alle relevante aktører, navnlig modtagerne, som er i direkte og konstant kontakt med deltagerne. Et eksempel på en kontrafaktisk konsekvensanalyse, der vurderer virkningerne på "bløde outcomes" findes i tekstboks 21 nedenfor.

Tekstboks 21 Et eksempel på vurdering af virkningerne på "bløde outcomes" i Tyskland

Forbedrer jobskabelsesordninger de langtidslediges sociale integration og trivsel?¹⁰⁵

Undersøgelsen undersøger virkningerne af en jobskabelsesordning rettet mod en sårbar gruppe af personer, langtidsledige, som har modtaget offentlige ydelser i mindst fire år og har helbredsproblemer eller børn eller begge dele. Programmet kørte fra 2015 til 2018 og omfattede ca. 20 000 deltagere. Foranstaltningen støttede op til 36 måneders almindelige ansættelseskontrakter med 30 timer om ugen (hovedsagelig med offentlige arbejdsgivere eller velgørenhedsorganisationer).

Eftersom den subsidierede beskæftigelsesforanstaltning, der varede op til 36 måneder, udtrykkeligt havde til formål at fremme den sociale integration af målgrupperne, forsøgte forfatterne at vurdere virkningerne i denne henseende.

Et omfattende datasæt, som integrerede administrative data ("Integrierten Erwerbsbiografie" (arbejdsmarkedsregister)) med en panelundersøgelse blandt programdeltagere og kontrolpersoner, blev anvendt til analysen. Panelundersøgelsen gjorde det muligt at indsamle oplysninger om subjektive mål for livskvalitet: tilfredshed med livet, mental sundhed, socialt tilhørsforhold og social status. Der blev anvendt Likert-vægte til at indsamle oplysninger om disse mål i overensstemmelse med andre undersøgelser (såsom den nationale undersøgelse PASS) for at opnå sammenlignelige resultater. De vigtigste trin i undersøgelsen var følgende: Behandlede og ubehandlede personer blev identificeret og matchet i de administrative data, de resulterende behandlings- og kontrolgrupper blev interviewet (tre bølger), det endelige datasæt blev opbygget, dataene blev rensset og tilfælde med manglende data blev fjernet. Stikprøven bestod af 2 531 matchede par i fase 1, 1 191 i fase 2 og 450 i fase 3 (ud af omkring 62 000 personer i alt, heraf 12 400 behandlede og 49 600 kontroller).

Med udgangspunkt i propensity score matching og måling af virkningerne 7, 18 og 29 måneder efter optagelsen i programmet konstaterede forfatterne, at interventionerne havde en positiv indvirkning på velfærdsmålene, men i forskellig grad: For eksempel steg tilfredsheden med livet betydeligt, mens den

¹⁰⁵ Se [Ivanov B. et al., 2020](#).

sociale status kun blev moderat forbedret. Virkningerne har dog tendens til at aftage i løbet af programmerne, hvilket potentielt kan forklares med en stigning i både antallet af deltagere, der forlader programmet, og kontrolpersonerne, der finder et job. For de mest sårbare personer var de positive virkninger højere.

Som understreget i begyndelsen af denne vejledning er kontrafaktisk konsekvensanalyse et effektivt analytisk instrument til at vurdere, "hvad" ESF+-interventionen har gjort. For at forstå, "hvordan" eller "hvorfor" de målte virkninger er opstået, skal der andre instrumenter til.

*Kontrafaktisk
konsekvensanalyse
integreret med TBE*

I nogle tilfælde, hvor interventionerne er velkendte og allerede er blevet grundigt evalueret, er en relativt begrænset analyse af gennemførelsen og nogle interview med støttemodtagerne tilstrækkelig til at identificere de vigtigste elementer, der ligger til grund for de målte virkninger. I mere komplekse, mindre kendte eller innovative interventioner er det ofte nødvendigt at integrere den kontrafaktiske konsekvensanalyse i andre evalueringer.

En gennemførelses- eller procesevaluering, der er gennemført før den kontrafaktiske konsekvensanalyse, kan vise, hvordan interventionen forløb, problemerne med gennemførelsen, og hvordan de forskellige aktører samarbejdede for at sikre en vellykket intervention. Resultaterne af denne evaluering kan anvendes som input til den kontrafaktiske konsekvensanalyse med angivelse af, hvilke virkninger der bør undersøges, hvordan kontrolgruppen skal sammensættes i forhold til deltagelseskarakteristika, og den rette tidsplan for en kontrafaktisk konsekvensanalyse.

Der kan gennemføres en teoribaseret evaluering (TBE) parallelt med den kontrafaktiske konsekvensanalyse. De to evalueringer kan give et gensidigt grundlag for hinanden: TBE indeholder forklaringer vedrørende de kontekstuelle, sociale og individuelle mekanismer, der påvirkede virkningerne og muliggjorde dem, mens den kontrafaktiske konsekvensanalyse beregner nettovirkningerne og præciserer, i hvilket omfang interventionen har givet væsentlige outcomes. I en konsekvensanalyse anbefales det at integrere kontrafaktisk konsekvensanalyse og TBE for at give et samlet overblik over virkningerne og årsagerne hertil.

Glossarer

Akronymer

ALMP	Aktiv arbejdsmarkedspolitik
CBA	Cost-benefit-analyse
CIE	Kontrafaktisk konsekvensanalyse
CPR	Forordningen om fælles bestemmelser
GD EMPL	Generaldirektoratet for Beskæftigelse, Sociale Anliggender, Arbejdsmarkedsforhold og Inklusion
GD REGIO	Generaldirektoratet for Regionalpolitik
DID	Difference-in-difference(s)
Kommissionen	Europa-Kommissionen
EP	Evalueringsplan
EFRU	Den Europæiske Fond for Regionaludvikling
ESF	Den Europæiske Socialfond
EU	Den Europæiske Union
Den generelle forordning om databeskyttelse	Den generelle forordning om databeskyttelse (EU) 2016/679
IB	Bemyndiget organ
IEB	Integrierten Erwerbsbiografie (tysk arbejdsmarkedsregister)
IV	Instrumentvariabel
JLD	Jobseeker longitudinal dataset (irsk arbejdsmarkedsregister)
LFS	Arbejdsstyrkeundersøgelse
MA	Forvaltningsmyndighed(er)
MS	Medlemsstat(er)
NGO'er	Ikkestatslige organisationer
OP	Operationelt program
PES	Offentlig arbejdsformidling
PSM	Propensity score matching (matchning efter sandsynlighed)
RCT	Randomiseret kontrolleret forsøg
RDD	Regressionsdiskontinuitetsdesign
SMV'er	Små og mellemstore virksomheder
UBI	Ungdomsbeskæftigelsesinitiativet

Definitioner

Udtryk	Definition
Basislinjedata	Data om variabler, der måles, før en enhed (person eller virksomhed) deltager i en intervention. I mange tilfælde vil der blive indsamlet mål for interventionsresultater før behandlingen for både behandlings- og kontrolgrupper.
Støttemodtager	I henhold til artikel 2 i Rådets forordning (EU) 1060/2021 (forordningen om fælles bestemmelser), er en støttemodtager "et offentligt eller privat organ, en enhed med eller uden status som juridisk person eller en fysisk person, der har ansvaret for at iværksætte eller både iværksætte og gennemføre operationer". I forbindelse med finansielle instrumenter (undertiden anvendt til f.eks. mikrokreditter til selvstændig virksomhed i forbindelse med ESF+) er en støttemodtager det organ, der gennemfører holdingfonden eller, hvis der ikke findes nogen holdingfondstruktur, det organ, der gennemfører den særlige fond eller, hvis forvaltningsmyndigheden forvalter det finansielle instrument, forvaltningsmyndigheden.
Kontrolgruppe	En gruppe personer, virksomheder eller andre enheder, som ligner behandlingsgruppen så meget som muligt, men som ikke bliver behandlet, og hvorfra der indhentes kontrafaktiske skøn over resultaterne. Strengt taget anvendes udtrykket "kontrolgruppe" for eksperimentelle metoder (se "Randomisering" nedenfor), mens udtrykket "sammenligningsgruppe" anvendes til kvasiexperimentelle metoder, men i praksis anvendes de to i flæng.
Kontrafaktisk analyse	En sammenligning af, hvad der faktisk skete, og hvad der ville være sket uden interventionen, med hensyn til resultaterne. Da forskellen mellem faktiske og kontrafaktiske resultater definerer foranstaltningens årsagsvirkning, omfatter den kontrafaktiske analyse alle metoder, der har til formål at vurdere andelen af observerede ændringer, som kan tilskrives den evaluerede intervention.
Difference-in-differences (DID)	I sin enkleste form trækkes forskellen i et resultat før og efter behandling i en kontrolgruppe fra den samme forskel, der er observeret i en behandlet gruppe, for at nå frem til et skøn over virkningen af en intervention. Virkninger, der beregnes på grundlag af difference-in-differences, udledes normalt inden for regressionsrammerne.
Effektivitet	Henviser til opfyldelse af mål og evalueres ved at sammenligne det, der er opnået, med, hvad der var planlagt (eller med en udgangssituation), eller ved at sammenligne, hvad der observeres, efter at foranstaltningen har fundet sted, med hvad der ville være sket uden foranstaltningen (kontrafaktisk situation).
Produktivitet	Produktivitet defineres som opnåelse af et givet output til minimale omkostninger eller tilsvarende som maksimering af output for et givet ressourceniveau. Det kan fastslås ved hjælp af cost-benefit- eller omkostningseffektivitetsanalyser.
Evalueringsplan	I henhold til artikel 44 i Rådets forordning (EU) 1060/2021 (forordningen om fælles bestemmelser) skal medlemsstaten eller forvaltningsmyndigheden udarbejde en evalueringsplan (EP), der kan omfatte mere end ét program, og som skal forelægges overvågningsudvalget senest ét år efter afgørelsen om godkendelse af programmet.
Ekstern evaluering	Evaluering foretaget af en ekstern og uafhængig evaluator på grundlag af en udbudsprocedure.
Virkning	I forbindelse med kontrafaktiske konsekvensanalyser henviser virkningerne til nettovirkninger, der defineres som forskellen mellem den gennemsnitlige behandling og de kontrafaktiske resultater. I denne vejledning anvendes udtrykket "virkninger" synonymt med "nettoeffekter".
Kontrafaktisk konsekvensanalyse	En type konsekvensanalyse, der forsøger at identificere foranstaltningernes årsagsvirkninger ved at anslå de gennemsnitlige kontrafaktiske resultater og trække disse fra de gennemsnitlige observerede resultater blandt de behandlede enheder. Skøn over kontrafaktiske resultater opnås typisk fra kontrolgrupper, der udvælges omhyggeligt, så de svarer til den behandlede gruppe.
Instrumentvariabelmetode (IV)	Ved anvendelse af denne metode bør udvælgelsen til behandling i det mindste delvist bestemmes af en eksogen faktor (eller et instrument), som ikke har noget at gøre med andre resultater end dem, der er opnået gennem behandlingen. Den eksogene faktor har således indflydelse på deltagelsen, men ikke direkte på resultaterne.

Udtryk	Definition
Intern evaluering	Evaluering foretaget internt, dvs. direkte bestilt af en uafhængig offentlig institution eller enhed (fra forvaltningsmyndigheden eller bemyndigede organer) uden en udbudsprocedure eller i form af en udvidet overvågnings- og analyseproces.
Interventioner	Henviser generelt til operationer i ESF's operationelle programmer eller projekter, der medfinansieres af ESF.
Matchning	En metode, hvor interventions- og kontrolstikprøver matches med hinanden på grundlag af deres observerede karakteristika.
Ikkerandomiseret eller kvasiekperimentelt design	Metoder til kontrafaktisk konsekvensanalyse, hvor kontrolgrupper oprettes ved hjælp af andre metoder end randomisering.
Outcome	De sandsynlige eller opnåede virkninger på kort og mellemlang sigt af en interventions output ¹⁰⁶ . Outcome svarer til "resultat", men anvendes oftere i konsekvensanalyser.
Output	Et output er det, der produceres/leveres direkte gennem gennemførelsen af en ESF-operation, målt i fysiske eller pengemæssige enheder. Output måles hovedsagelig i antallet af støttede personer, støttede enheder, leverede varer og tjenesteydelser og gennemførte projekter.
Deltagere	Deltager er en fysisk person, der drager direkte fordel af en operation, uden at være ansvarlig for at iværksætte eller gennemføre en sådan operation. I ESF-sammenhæng henviser det til personer, der støttes af ESF-interventioner.
Procesevaluering	Procesevaluering fokuserer på programgennemførelsen, herunder, men ikke begrænset til, hvordan tjenesteydelserne leveres, forskelle mellem den tilsigtede målgruppe og den målgruppe, der betjenes, adgang til programmet og forvaltningspraksis.
Propensity score matching (matchning efter sandsynlighed) (PSM)	PSM indebærer estimering af en statistisk model for hele stikprøven (behandling og potentielle kontroller), som giver en skønnet sandsynlighed for at deltage for hver person eller virksomhed, uanset om de rent faktisk deltager eller ej. Behandlede personer eller virksomheder matches derefter — enten med én ikkebehandlet person eller virksomhed eller med mange ikkebehandlede personer eller virksomheder — på grundlag af propensity scoren.
Randomisering	Det er en metode, hvor medlemmer af en målgruppe tilfældigt allokeres til en række behandlings- eller kontrolbetingelser. Randomisering sikrer, at grupperne statistisk set er ens i alle henseender på det tidspunkt, hvor de randomiseres.
Regressionsdiskontinuitetsdesign (RDD)	Denne metode kan anvendes, når adgangen til en intervention bestemmes af et skæringspunkt langs en kontinuerlig rating, skala eller måling. Denne tilgang bygger på den omstændighed, at dem, der befinder sig umiddelbart omkring skæringspunktet, vil være meget ens, men at de personer, der befinder sig på den ene side af skæringspunktet, deltager, mens dem på den anden side ikke deltager. Resultaterne for dem over og under skæringspunktet kan sammenlignes for at vurdere interventionens virkning.
Relevans	Relevans henviser til hensigtsmæssigheden af de eksplicite mål for en intervention med hensyn til de socioøkonomiske problemer, som interventionen skal løse ¹⁰⁷ .
Resultat	Virkningerne på deltagere eller enheder som følge af en operation, f.eks. med hensyn til deres beskæftigelsessituation, indkomst, resultater i standardiserede uddannelsestest, overskud osv. Virkningerne kan måles på kort eller længere sigt. I forbindelse med ESF+ anses den kortsigtede periode ved måling af resultatindikatorer for at være umiddelbart efter deltagelsen (fire uger), mens den længere periode betragtes som seks måneder efter deltagelsen ¹⁰⁸ . I forbindelse med konsekvensanalysen betyder længere sigt imidlertid ofte en længere periode, 24 måneder eller mere efter interventionerne.

¹⁰⁶ Se [OECD, 2010](#).

¹⁰⁷ Se [Europa-Kommissionen, 2013](#).

¹⁰⁸ Se [Europa-Kommissionen, 2021a](#) — og [Europa-Kommissionen, 2018](#).

Udtryk

Definition

Behandlingsgruppe

En gruppe personer, virksomheder eller andre enheder, som drager fordel af eller er eksponeret for en intervention (dette kan være et tilbud om behandling eller faktisk modtagelse heraf).

Bibliografi

- Ashenfelter, O. (1978). Estimating the effect of training programmes on earnings, *Review of Economics and Statistics*, 6, side 47-57. <https://www.jstor.org/stable/1924332?origin=crossref>
- Axdorph E, Egebark J., Lundström T., Özcan G. (2019) Effekter av förstärkta förmedlingsinsatser för unga arbets sökande — Resultat från utvärderingen av Ung framtid. Arbetsförmedlingen
- Baran J. et al. (2016). Badanie efektów wsparcia zrealizowanego na rzecz osób młodych w ramach Programu Operacyjnego Wiedza Edukacja Rozwój. I Raport Wskaźnikowy. <http://files.evaluationhelpdesk.eu/Evaluations/PLE49.pdf>
- Baran J. et al. (2017). Badanie efektów wsparcia zrealizowanego na rzecz osób młodych w ramach Programu Operacyjnego Wiedza Edukacja Rozwój II Raport wskaźnikowy. <http://files.evaluationhelpdesk.eu/Evaluations/PLE196.pdf>
- Baran J. et al. (2018). Badanie efektów wsparcia zrealizowanego na rzecz osób młodych w ramach Programu Operacyjnego Wiedza Edukacja Rozwój III Raport wskaźnikowy. <http://files.evaluationhelpdesk.eu/Evaluations/PLE197.pdf>
- Baran J. et al. (2018a). Badanie efektów wsparcia zrealizowanego na rzecz osób młodych w ramach Programu Operacyjnego Wiedza Edukacja Rozwój. II Raport Tematyczny. <http://files.evaluationhelpdesk.eu/Evaluations/PLE221.pdf>
- Bazzoli M., De Poli S., Rettore E., Schizzerotto A. (2018). Are Vocational Training Programmes Worth Their Cost? Evidence from a Cost-Benefit Analysis. In *Politica Economica/Journal of Economic Policy* Vol. XXXIV(3), pagg. 215-240. <https://www.rivisteweb.it/doi/10.1429/92119>
- Bloom, H. S. (1995). Minimum detectable effects: A simple way to report the statistical power of experimental designs, *Evaluation Review*, 8(2), 225-246. <https://cpb-us-e2.wpmucdn.com/sites.uci.edu/dist/1/1159/files/2021/03/Bloom-MDES-Eval-Rev-1995-Bloom.pdf>
- Bloom, H. S. (2009) Modern regression discontinuity analysis, MDRC Working Papers on Research Methodology, New York: MDRC. <https://www.mdrc.org/publication/modern-regression-discontinuity-analysis>
- Boockmann B. et al. (2017). Evaluation des ESF-Bundesprogramms zur Eingliederung langzeitarbeitsloser Leistungsberechtigter nach dem SGB II auf dem allgemeinen Arbeitsmarkt. Im Auftrag des Bundesministeriums für Arbeit und Soziales. <http://files.evaluationhelpdesk.eu/Evaluations/DEE189.pdf>
- Boockmann B. et al. (2018). Evaluation des ESF-Bundesprogramms zur Eingliederung langzeitarbeitsloser Leistungsberechtigter nach dem SGB II auf dem allgemeinen Arbeitsmarkt. Im Auftrag des Bundesministeriums für Arbeit und Soziales. <http://files.evaluationhelpdesk.eu/Evaluations/DEE52.pdf>
- Boockmann B. et al. (2019). Evaluation des ESF-Bundesprogramms zur Eingliederung langzeitarbeitsloser Leistungsberechtigter nach dem SGB II auf dem allgemeinen Arbeitsmarkt. Im Auftrag des Bundesministeriums für Arbeit und Soziales. <http://files.evaluationhelpdesk.eu/Evaluations/DEE94.pdf>
- Boockmann B. et al. (2021). Evaluation des ESF-Bundesprogramms zur Eingliederung langzeitarbeitsloser Leistungsberechtigter nach dem SGB II auf dem allgemeinen Arbeitsmarkt. Im Auftrag des Bundesministeriums für Arbeit und Soziales. <http://files.evaluationhelpdesk.eu/Evaluations/DEE202.pdf>
- Borik V. et al. (2015). The net effects of graduate work experience and the promotion of self-employment. Technical Report. <http://files.evaluationhelpdesk.eu/Evaluations/SKE8.pdf>

- Bratti M. et al. (2018). Vocational Training for Unemployed Youth in Latvia: Evidence from a Regression Discontinuity Design IZA DP No. 11870. <http://ftp.iza.org/dp11870.pdf>
- Bratu C. et al. (2014). Knowledge gaps in evaluating labour market and social inclusion policies identified 39 ESF counterfactual impact evaluations in the 2007-2013 programming period. doi:10.2788/083390. <https://op.europa.eu/en/publication-detail/-/publication/d0e73217-2a40-47af-a5c3-a700dcd47da3/language-en>
- Bredgaard, T. (2015). Evaluering af, hvad der virker for hvem i aktive arbejdsmarkedspolitikker. European Journal of Social Security, 17(4), 436-452. <https://doi.org/10.1177/138826271501700403>
- Bryson, A., Dorsett, R. og Purdon, S. (2002) The use of propensity score matching in the evaluation of active labour market policies, Department for Work and Pensions, Working Paper Number 4. [http://eprints.lse.ac.uk/4993/1/The use of propensity score matching in the evaluation of active labour market policies.pdf](http://eprints.lse.ac.uk/4993/1/The_use_of_propensity_score_matching_in_the_evaluation_of_active_labour_market_policies.pdf)
- Caliendo M., Kopeinig S., (2008), Some practical Guidance for the implementation of propensity score matching. Journal of Economic Surveys; Volume 22, Issue 1 — Pages 31-72. <https://doi.org/10.1111/j.1467-6419.2007.00527.x>
- Caliendo, M., Mahlstedt, R., og Mitnik, O. A. (2017). Unobservable, but unimportant? The relevance of usually unobserved variables for the evaluation of labour market policies. Labour Economics, 46, 14-25. <https://doi.org/10.1016/j.labeco.2017.02.001>
- Card, D., Ibarra, P. og Villa, J. M. (2011). Building in an evaluation component for active labour market programs: a practitioner's guide, Discussion Paper No. 6085, Bonn, Tyskland: IZA. <http://ftp.iza.org/dp6085.pdf>
- Card D., Kluve J., Weber A. (2017). What works? A meta-analysis of recent active labor market program evaluations. Journal of the European Economic Association, Volume 16, Issue 3, June 2018, Pages 894–931, <https://doi.org/10.1093/jeaa/jvx028>
- Centre for Disease Control and Prevention CDC (2013). Good Evaluation Questions: Checklist to Help Focus Your Evaluation. Department of Health & Human Services — USA. http://www.cdc.gov/asthma/program_eval/AssessingEvaluationQuestionChecklist.pdf
- Danmarks Statistik (2017). Effektmåling og monitorering 2016 — strukturfondsindsatsen i Syddanmark. <http://files.evaluationhelpdesk.eu/Evaluations/DKE7.pdf>
- Danmarks Statistik (2018). Effektmåling af den Virksomhedsrettede Strukturfondsindsats 2007-2013. <http://files.evaluationhelpdesk.eu/Evaluations/DKE10.pdf>
- Diaz J. M. et al. (2019). Evaluación del PO-FSE 2014/2020 del Principado de Asturias para el informe anual a presentar en 2019. <http://files.evaluationhelpdesk.eu/Evaluations/ESE101.pdf>
- Ecorys, Iseri Europa (2020) Study supporting the evaluation of ESF support to education and training (Thematic Objective 10) for the Commission, Directorate-General for Employment, Social Affairs and Inclusion. <https://op.europa.eu/en/publication-detail/-/publication/d0c1a558-077d-11eb-a511-01aa75ed71a1/language-en/format-PDF/source-173162502>
- Europa-Kommissionen (2007): Indicative Guidelines on evaluation methods: evaluation during the programming period. Working paper no. 5. GD Regionalpolitik. https://ec.europa.eu/regional_policy/sources/docoffic/2007/working/wd5_ongoing_en.pdf
- Europa-Kommissionen (2013): EVALSED: The resource for the evaluation of Socio-Economic Development. Opdateret version. https://ec.europa.eu/regional_policy/sources/docgener/evaluation/guide/guide_evalsed.pdf.
- Europa-Kommissionen (2017) Better Regulation Toolbox, Tool #47 *Evaluation criteria and questions* https://ec.europa.eu/info/sites/default/files/file_import/better-regulation-toolbox-

47_en_0.pdf

Europa-Kommissionen (2018) Monitoring and Evaluation of European Cohesion Policy European Social Fund. Guidance document.

<https://ec.europa.eu/sfc/en/system/files/ged/ESF%20monitoring%20and%20evaluation%20guidance.pdf>

Europa-Kommissionen (2019) Advanced counterfactual evaluation methods. Guidance document. doi:10.2767/464242.

<https://ec.europa.eu/social/main.jsp?catId=738&langId=en&pubId=8254&furtherPubs=yes>

Europa-Kommissionen (2020) Counterfactual impact evaluation of European Social Fund interventions in practice. Guidance document for Managing Authorities. doi:10.2767/55495.

<https://op.europa.eu/en/publication-detail/-/publication/f82c5fb8-bb40-11ea-811c-01aa75ed71a1/language-en>

Europa-Kommissionen (2020) How to use administrative data for European Social Funds counterfactual impact evaluations. A step-by-step guide for managing authorities.

<https://op.europa.eu/en/publication-detail/-/publication/d96feed3-f30c-11ea-991b-01aa75ed71a1>

Europa-Kommissionen (2021) Europa-Parlamentets og Rådets forordning (EU) 2021/1060 af 24. juni 2021 om fælles bestemmelser for Den Europæiske Fond for Regionaludvikling, Den Europæiske Socialfond Plus, Samhørighedsfonden, Fonden for Retfærdig Omstilling og Den Europæiske Hav-, Fiskeri- og Akvakulturfond og om finansielle regler for nævnte fonde og for Asyl-, Migrations- og Integrationsfonden, Fonden for Intern Sikkerhed og instrumentet for finansiell støtte til grænseforvaltning og visumpolitik

Europa-Kommissionen (2021a): Common Indicators Toolbox Working document. Juni 2021

https://ec.europa.eu/sfc/en/system/files/2021/ged/toolbox_june_2021_final.pdf

Fondazione G. Brodolini, Metis GmbH, Applica, Ockham IPS (2020). Study for the Evaluation of ESF Support to Employment and Labour Mobility (Thematic Objective 8), for the European Commission, Directorate- General for Employment, Social Affairs and Inclusion.

<https://ec.europa.eu/social/BlobServlet?docId=22899&langId=en>

Frolich, M (2004) Programme evaluation with multiple treatments, Journal of Economic Surveys, 18(2), side 181-224.

<https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.620.9209&rep=rep1&type=pdf>

Guthmuller S., Meroni E.C., Evaluation of the higher education grant system for less privileged students in Portugal. JRC's tekniske rapport. Endnu ikke offentliggjort.

Hagglund, P. (2006). A description of three randomised experiments in Swedish labour market policy, Institute for Labour Market Policy Evaluation, Report 2006:4.

<https://www.ifau.se/globalassets/pdf/se/2006/r06-04.pdf>

Heckman, J. J., Ichimura, H., Smith, J. og Todd, P. (1998) Characterizing selection bias using experimental data. In Econometrica, Vol. 66, nr. 5. s. 1017-1098.

http://jenni.uchicago.edu/papers/Heckman_Ichimura_etal_1998_Econometrica_v66_n5_r.pdf

Hunger K. og Sattler K. (2017). Evaluationsbericht zum spezifischen Ziel A 1.1. im Rahmen der Evaluierung des Operationellen Programms des Europäischen Sozialfonds in Baden-Württemberg 2014-2020. <http://files.evaluationhelpdesk.eu/Evaluations/DEE26.pdf>

Holland, P. W. (1986). Statistics and Causal Inference, Journal of the American Statistical Association, 81 (396), 945-960 DOI: 10.1080/01621459.1986.10478354.

<http://people.umass.edu/~stanek/pdffiles/causal-holland.pdf>

IAW Institut für Angewandte Wirtschaftsforschung — ISG Institut für Sozialforschung und Gesellschaftspolitik, GmbH (2015). Evaluation der Modellprojekte "Bürgerarbeit". Endbericht.

<http://files.evaluationhelpdesk.eu/Evaluations/DEE7.pdf>

- ICF, Cambridge Econometrics and Eurocentre (2020) Study supporting the 2020 evaluation of promoting social inclusion, combating poverty and any discrimination by the European Social Fund (Thematic Objective 9) for the European Commission, Directorate-General for Employment, Social Affairs and Inclusion. <https://op.europa.eu/en/publication-detail/-/publication/8788ec85-2308-11eb-b57e-01aa75ed71a1>
- Indecon (2016). Indecon Evaluation of JobBridge Activation Programme. <http://files.evaluationhelpdesk.eu/Evaluations/IEE2.pdf>
- Instytut Badań Strukturalnych, Imapp, IQS. (2015). Badanie efektów wsparcia zrealizowanego na rzecz osób młodych w ramach Programu Operacyjnego Wiedza Edukacja Rozwój. I Raport Tematyczny. <http://files.evaluationhelpdesk.eu/Evaluations/PLE34.pdf>
- Ires Piemonte. (2019). Misura di sostegno all'imprenditorialità. <http://files.evaluationhelpdesk.eu/Evaluations/ITE61.pdf>
- Isfol (2016). Primo rapporto di valutazione del piano italiano Garanzia giovani. <http://files.evaluationhelpdesk.eu/Evaluations/ITE13.pdf>
- Ismeri Europa (2018). I Rapporto tematico di valutazione — I risultati di PIPOL. <http://files.evaluationhelpdesk.eu/Evaluations/ITE36.pdf>
- Ismeri Europa — Ecorys — Institute for Employment Studies. (2019). Pilot and feasibility study on the sustainability and effectiveness of results for European Social Fund participants using counterfactual impact evaluations. doi:10.2767/39339. <https://op.europa.eu/it/publication-detail/-/publication/84cc9eb9-b33d-11e9-9d01-01aa75ed71a1>
- Ivanov B. et al., 2020, Do Job creation schemes improve the social integration and well-being of the long-term unemployed?, Labour Economics. <https://doi.org/10.1016/j.labeco.2020.101836>.
- Kalinowski H. (2020). Badanie efektów wsparcia zrealizowanego na rzecz osób młodych w ramach Programu Operacyjnego Wiedza Edukacja Rozwój V Raport wskaźnikowy. <http://files.evaluationhelpdesk.eu/Evaluations/PLE296.pdf>
- Kalinowski H. et al., 2020 Badanie efektów wsparcia zrealizowanego na rzecz osób młodych w ramach Programu Operacyjnego Wiedza Edukacja Rozwój Raport końcowy. <http://files.evaluationhelpdesk.eu/Evaluations/PLE339.pdf>
- Kluve, J., Lehmann, H., and Schmidt, C. M. (2008). Disentangling treatment effects of Active Labor Market Policies: The role of labor force status sequences. Labour Economics, 15(6), 1270-1295. <https://doi.org/10.1016/j.labeco.2007.12.002>
- Krug, G and Stephan, G. (2011) Is contracting-out intensified placement services more effective than in-house production? Evidence from a randomized field experiment, LASER Discussion Papers — Paper No. 5, <http://doku.iab.de/externe/2011/k110912303.pdf>
- Lammers M. and Kok L. (2021). Are active labor market policies (cost-)effective in the long run? Evidence from the Netherlands. I Empirical Economics 60:1719–1746. <https://doi.org/10.1007/s00181-019-01812-3>
- Lang, J. (2021) Employment effects of language training for unemployed immigrants. J Popul Econ (2021). <https://doi.org/10.1007/s00148-021-00832-7>
- Martini, A. (2009). Counterfactual impact evaluation: what it can (and cannot) do for cohesion policy, udarbejdet for den 6. europæiske konference om evaluering af samhørighedspolitikken, Warszawa, 30. november. https://ec.europa.eu/regional_policy/archive/conferences/evaluation2009/abstracts/martini.doc
- Morgan, S. L. and Winship, C. (2014) Counterfactual and causal inference: Methods and principles for social research. 2nd edition. Cambridge and New York: Cambridge University Press.

- OECD (2010), Glossary of Key Terms in Evaluation and Results Based Management, Paris. <https://www.oecd.org/dac/evaluation/2754804.pdf>
- Openfield (2019). Analiza skuteczności i efektywności dotacji na założenie działalności gospodarczej udzielonych w ramach 8 osi priorytetowej RPO WM — komponent 3. <http://files.evaluationhelpdesk.eu/Evaluations/PLE312.pdf>
- Palczyńska M. et al. (2019) Badanie efektów wsparcia zrealizowanego na rzecz osób młodych w ramach Programu Operacyjnego Wiedza Edukacja Rozwój IV Raport wskaźnikowy. <http://files.evaluationhelpdesk.eu/Evaluations/PLE274.pdf>
- Pomatto G. (2017). L'attuazione del Buono per Servizi al Lavoro nella Regione Piemonte. Ires Piemonte. <http://files.evaluationhelpdesk.eu/evaluations/ITE55.pdf>
- Pomatto G. (2019). Buoni per servizi al lavoro nella Regione Piemonte: qualità percepita dai destinatari e meccanismi dell'attuazione. Ires Piemonte. <http://files.evaluationhelpdesk.eu/evaluations/ITE77.pdf>
- Pompili M., Giorgetti I. (2020). Rapporto di Placement. Servizio di attività di valutazione del POR FSE 2014/2020 — Regione Marche. <http://files.evaluationhelpdesk.eu/Evaluations/ITE115.pdf>
- Pompili M., Giorgetti I. (2020a). Rapporto tematico "Disoccupazione di lunga durata". Servizio di attività di valutazione del POR FSE 2014/2020 — Regione Marche. <http://files.evaluationhelpdesk.eu/Evaluations/ITE214.pdf>
- Poy S. (2019). Gli effetti occupazionali del buono per servizi al lavoro nella Regione Piemonte: prime evidenze. Misura per disoccupati da almeno 6 mesi, anno 2017. Ires Piemonte. <http://files.evaluationhelpdesk.eu/evaluations/ITE123.pdf>
- Poy S (2020). Nuove evidenze sull'effetto occupazionale del buono per servizi al lavoro. Target persone disoccupate da almeno 6 mesi. Ires Piemonte. <http://files.evaluationhelpdesk.eu/evaluations/ITE245.pdf>
- Re-source Pracownia Badań i Doradztwa (2015). Ocena działań na rzecz wydłużenia aktywności zawodowej osób starszych w kontekście niekorzystnej sytuacji demograficznej w województwie lubelskim. <http://files.evaluationhelpdesk.eu/evaluations/PLE126.pdf>
- Riccio J., Friedlander, D., Freedman S. (1994). GAIN: Benefits, Costs, and Three-Year Impacts of a Welfare-to-Work Program, MDRC, NYC <https://www.mdrc.org/publication/gain-benefits-costs-and-three-year-impacts-welfare-work-program>
- Scheller F. og Seidel K. (2020). Zweiter Evaluationsbericht zum spezifischen Ziel A 1.1: Teilnehmer*innenperspektive und Wirkungsanalyse. Im Rahmen der Evaluierung des Operationellen Programms des Europäischen Sozialfonds in Baden-Württemberg 2014–2020. <http://files.evaluationhelpdesk.eu/Evaluations/DEE171.pdf>
- Shadish, W. R., Cook, T. D. og Campbell, D. T. (2002). Experimental and quasi-experimental designs for generalised causal inference, Boston, US: Houghton Mifflin Company.
- Stern E. et al. (2012). Broadening the range of designs and methods for impact evaluations, Department for International Development of the UK, Working Paper 38. <https://assets.publishing.service.gov.uk/media/57a08a6740f0b6497400059e/DFIDWorkingPaper38.pdf>
- WK Kellog Foundation (2004). Logic Model Development Guide. <https://www.wkcf.org/resource-directory/resources/2004/01/logic-model-development-guide>
- Zub M. et al. (2020). Ocena wpływu wsparcia RPOWP na popularyzację szkolnictwa zawodowego w województwie podlaskim. Raport końcowy. <http://files.evaluationhelpdesk.eu/Evaluations/PLE359.pdf>

Bilag

Bilag 1. Yderligere oplysninger

Følgende er anbefalet læsning for forvaltningsmyndighedens personale, der er interesseret i mere detaljerede oplysninger om de spørgsmål, der behandles i denne vejledning. Der findes omfattende litteratur om evaluering. Formålet med denne liste er at fremhæve pålidelige større diskussioner, der umiddelbart giver nyttige oplysninger til planlægningen af kontrafaktiske konsekvensanalyser. Efter hver henvisning gives der en kort beskrivelse af de fleste kilder.

Generel evaluering

- Rossi, Peter H., Mark W. Lipsey, og Gary T. Henry. (2018). *Evaluation: A Systematic Approach*. 8th edition. Thousand Oaks, CA: SAGE Publications.

Den klassiske lærebog om evalueringspraksis og -metoder. Indeholder metoder og eksempler.

- Gertler, Paul J., Martinez, Sebastian, Premand, Patrick, Rawlings, Laura B., Vermeersch, Christel M. J. (2016). *Impact Evaluation in Practice, Second Edition*. Washington D.C.: Den Interamerikanske Udviklingsbank og Verdensbanken. ©Verdensbanken
<https://openknowledge.worldbank.org/handle/10986/25030> (tilgængelig på engelsk, portugisisk og spansk)

Ligesom denne vejledning begynder denne håndbog med den klassiske evaluering (kontrollerede randomiserede forsøg) og beskriver derefter alternativer. Selv om den er skrevet for programledere i lavindkomstlande, er den relevant og umiddelbart anvendelig i forbindelse med EU's medlemsstater.

- White, H., & Raitzer, D. A. (2017). *Impact evaluation of development interventions: A practical guide*. Den Asiatiske Udviklingsbank
<https://www.adb.org/sites/default/files/publication/392376/impact-evaluation-development-interventions-guide.pdf>

Ligesom Gertler et al. (2016), jf. ovenfor, indeholder denne vejledning en omfattende gennemgang af generelle metoder til kontrafaktisk konsekvensanalyse, samtidig med at den også kommer ind på evaluering af virkninger af interventioner i lav- og mellemindkomstlande. Den er mere teknisk end Gertler et al. (2016), men indeholder ikke desto mindre mange praktiske oplysninger, der kan være interessante for ESF-programforvaltere.

- Europa-Kommissionen (2019) *Advanced counterfactual evaluation methods*. Guidance document. doi:10.2767/464242. <https://op.europa.eu/en/publication-detail/-/publication/11968bbb-fac9-11e9-8c1f-01aa75ed71a1/language-en>

Dokumentet præsenterer nyere og mere avancerede metoder til kontrafaktisk effektevaluering, såsom sekvensanalyse, dynamisk matchning og syntetisk kontrol.

- Csillag Marton, Kreko Judi og Scharle Agota. (2020). *Counterfactual evaluation of youth employment policies*. Prepared under the "Youth Employment PartnerSHIP" project. <http://yepartnership.ibs.org.pl/content/uploads/2021/02/Methodological-guide.pdf> (tilgængelig på engelsk, spansk, ungarsk (under oversættelse), italiensk og polsk)

Dette er en trinvis introduktion til den kontrafaktiske evaluering af arbejdsmarkedspolitikkerne for unge med fokus på anvendelsen af administrative data. Der redegøres for problemstillinger på grundlag af de praktiske problemer, der er opstået i fire kontrafaktiske konsekvensanalyser i forbindelse med tilskud til ansættelse af unge i Spanien, Ungarn, Italien og Polen.

- HM Treasury (Det Forenede Kongerige) (2020). The Magenta Book: Guidance for evaluation. London: The Agency. <https://www.gov.uk/government/publications/the-magenta-book>

"Magenta" -bogen indeholder detaljerede oplysninger om evalueringsmetoder. Disse dokumenter er interessante som eksempler på et internt statsligt evalueringsspektiv.

Randomiserede kontrollerede forsøg

- White H. (2013) An introduction to the use of randomised control trials to evaluate development interventions, Journal of Development Effectiveness, 5:1, 30-49, DOI: 10.1080/19439342.2013.764652
<https://www.tandfonline.com/doi/pdf/10.1080/19439342.2013.764652>

- White, H., Sabarwal S. og T. de Hoop, (2014). Randomized Controlled Trials (RCTs), Methodological Briefs: Impact Evaluation 7, UNICEF Office of Research, Florence.
https://www.unicef-irc.org/publications/pdf/brief_7_randomized_controlled_trials_eng.pdf

To ikke-tekniske introduktioner til logikken bag RCT med en drøftelse af flere design og kritik af RCT.

- Glennerster, R., og Takavarasha, K., (2013). Running Randomized Evaluations: A Practical Guide, Princeton University Press, Princeton, NJ.

Bogen er en trinvis vejledning i, hvordan man udformer og gennemfører RCT'er inden for sociale programmer. Den bygger på de konkrete RCT'er, som Abdul Latif Jameel Poverty Action Lab har gennemført.

Difference-in-differences

- Card, David, Pablo Ibararán, og Juan Miguel Villa. (2011). Building in an Evaluation Component for Active Labor Market Programs: A Practitioner's Guide. IZA Discussion Paper No. 6085. Bonn: Forschungsinstitut zur Zukunft der Arbeit. <http://ftp.iza.org/dp6085.pdf>

Sammenligner Diff-in-Diff med kontrollerede randomiserede forsøg.

- Lechner, M. (2011) The Estimation of Causal Effects by Difference-in-Difference Methods. Foundations and Trends in Econometrics. Vol. 4, No. 3 (2010) 165-224 DOI: 10.1561/0800000014. https://michael-lechner.eu/ml_pdf/journals/2011_Lechner_DiD_2011_ECO%20403%20Lechner_darf%20aus%20Netz.pdf

I dette dokument drøftes DID-tilgangen indgående og nogle af de vigtigste spørgsmål i forbindelse med anvendelsen heraf. Udvidelser af DID, såsom ikke-lineære applikationer og propensity score matching med DID, præsenteres også.

- Fredriksson A. and de Oliveira G. M. (2019) Impact evaluation using Difference-in-Differences. RAUSP Management Journal Vol. 54 No. 4, s. 519-532. DOI 10.1108/RAUSP-05-2019-0112 <https://www.emerald.com/insight/content/doi/10.1108/RAUSP-05-2019-0112/full/pdf?title=impact-evaluation-using-difference-in-differences>

En oversigt over DID-metoderne med praktiske anbefalinger.

- Card, David og Alan B. Krueger. (1994). Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania. American Economic Review, 84 (4), 774-775. <https://davidcard.berkeley.edu/papers/njmin-aer.pdf>

Det klassiske eksempel på anvendelse af en difference-in-difference-teknik.

Instrumentvariabler

- Morgan, Stephen L. og Christopher Winship. (2014). Counterfactuals and Causal Inference: Methods and Principles for Social Research. 2nd edition. Cambridge and New York: Cambridge University Press.

Dette er en noget teknisk gennemgang af metoder til kontrafaktisk konsekvensanalyse med sociologisk terminologi. Kapitel 9, "Instrumental Variable Estimators of Causal Effects" (s. 291-324) giver et overblik over logikken bag og procedurerne for IV-estimerer.

- Kuhn, Andreas, Jean-Philippe Wuellrich og Josef Zweimüller. 2010. Fatal Attraction? Access to Early Retirement and Mortality. IZA Discussion Paper No.5160. Bonn: Forschungsinstitut zur Zukunft der Arbeit. <http://ftp.iza.org/dp5160.pdf>

Anvender regionale variationer i ændring af pensionsalderen i Østrig som instrumentvariabel i en undersøgelse af virkningen af førtidspensionering på arbejdstagernes sundhed.

- Galiani, S., Rossi, M. A., og Schargrodsky, E. (2011). Conscriptio and crime: Evidence from the Argentine draft lottery. *American Economic Journal: Applied Economics*. <https://www.aeaweb.org/articles?id=10.1257/app.3.2.119>

Innovativ og illustrativ brug af et randomiseret lotteri (til værnepligt) som instrumentvariabel. Meget læseværdig undersøgelse, der er prototypisk for en forskningsgren, der anvender lotteriprocedurer som instrumenter.

Matchning

- Heinrich, Carolyn, Alessandro Maffioli og Gonzalo Vázquez. 2010. A Primer for Applying Propensity Score Matching. Impact-Evaluation Guidelines Technical Notes No. IDB-TN-161. Washington: Den Interamerikanske Udviklingsbank. <https://publications.iadb.org/publications/english/document/A-Primer-for-Applying-Propensity-Score-Matching.pdf>

Ligesom regressionsdiskontinuitetsvejledningen nedenfor er denne skrevet til velinformede evalueringsledere.

- Caliendo M., Kopeinig S., (2008), Some practical Guidance for the implementation of propensity score matching. *Journal of Economic Surveys*; Volume 22, Issue 1 — Pages 31-72. <https://doi.org/10.1111/j.1467-6419.2007.00527.x>

En klassisk tekst om de valg, der skal træffes i forbindelse med gennemførelsen af PSM (med hensyn til estimering, matchningsalgoritmer, kvalitetsvurdering af matchningen, følsomhed af estimerede behandlingsvirkninger).

- Harris, H. og Horst, S. J. (2016) A Brief Guide to Decisions at Each Step of the Propensity Score Matching Process, *Practical Assessment, Research, and Evaluation*: Vol. 21, Article 4. DOI: <https://doi.org/10.7275/yq7r-4820>. Findes på: <https://scholarworks.umass.edu/pare/vol21/iss1/4>.

Svarer til den foregående artikel.

Regressionsdiskontinuitetsdesign

- Jacob, Robin, Pei Zhu, Marie-Andrée Somers og Howard Bloom. (2012). A Practical Guide to Regression Discontinuity. New York: MDRC. https://www.mdrc.org/sites/default/files/RDD%20Guide_Full%20rev%202016_0.pdf

Usædvanligt tilgængelig og grundig drøftelse af regressionsdiskontinuitetsmetoden, som indeholder en omhyggeligt udvalgt bibliografi.

- Lee, D. S. and Lemieux T. (2009), Regression discontinuity designs in economics, NBER Working Paper No. 14723, National Bureau of Economic Research, Cambridge, MA, 2009. http://www.nber.org/papers/w14723.pdf?new_window=1

Dokumentet er en slags "brugervejledning" til regressionsdiskontinuitet (RD), der beskriver metodens logik og illustrerer forskellige metoder til estimering af regressionsdiskontinuitetsdesign.

Bilag 2. Forslag til kursus om kontrafaktisk konsekvensanalyse

Et introduktionskursus i kontrafaktiske konsekvensanalyser for forvaltningsmyndigheder og embedsmænd kan omfatte følgende:

- Introduktion til evalueringsmetoder: Procesevalueringer (hvorfor de er vigtige, og hvordan de vedrører kontrafaktiske konsekvensanalyser) og konsekvensanalyser (kontrafaktiske og teoribaserede tilgange)
- Hvad er kontrafaktiske konsekvensanalyser? Hvilke evalueringsspørgsmål kan gode kontrafaktiske konsekvensanalyser besvare? For hvilke evalueringskriterier sikrer kontrafaktiske konsekvensanalyser den vigtigste dokumentation? Hvorfor er kontrafaktiske konsekvensanalyser vigtige?
- Hvordan fungerer de? (Behandlingsgrupper og kontrolgrupper, sammenligninger "før og efter", inferensprincipper)
- Oversigt over metoder til kontrafaktisk konsekvensanalyse (karakteristika, styrker og svagheder ved hver metode)
 - o Randomiseret kontrolleret forsøg
 - o Matchning
 - o Difference-in-difference(s)
 - o Regressionsdiskontinuitetsdesign
 - o Instrumentvariabler
 - o Angivelse af andre mulige metoder (f.eks. teknikker til rumlige analyser)
- Datakrav
 - o Krav til behandlings- og kontrolgrupper
 - o Mulige kilder (ESF-overvågning, administrative datasæt, undersøgelser)
 - o Regler i den generelle forordning om databeskyttelse
- Oversigt over de enkelte trin i gennemførelsen:
 - o Planlægning af kontrafaktiske konsekvensanalyser (evalueringsplan, gennemførligheden af en kontrafaktisk konsekvensanalyse)
 - o iværksættelse af kontrafaktiske konsekvensanalyser (kommissorium, udvælgelseskriterier og -metoder)
 - o Styring af kontrafaktiske konsekvensanalyser (kontrol af leverancer og interaktion med evaluatoren)
 - o Formidling af resultater fra kontrafaktiske konsekvensanalyser (målgruppetyper og kommunikationskanaler).

Et kursus, der er struktureret som ovenfor, bør omfatte praktiske eksempler, der skal udvikles på workshoper, og vil løbe over 2-3 dage. En tilgang til at afholde et kursus som dette ville være at præsentere en problembaseret læringsmetode og anvende en politisk foranstaltning under programmet som den case, der skal udvikles.

Bilag 3. Kontrafaktiske konsekvensanalyser — Eksempler nævnt i vejledningen

Tabel 6 Karakteristika for de kontrafaktiske konsekvensanalyser, der er nævnt i vejledningen som eksempler

Land	Titel — år	Vurderet politik og observeret behandlingsperiode	Metode	Behandlingsgruppe og kontrolgruppe — definition og størrelse	Udfaldsvariable	Observation periode	Data
Sverige	Virkningerne af intensiveret støtte til unge jobsøgende — Resultater af evalueringen af Ung framtid (2019)	Intensiveret støtte fra den offentlige arbejdsformidling til jobsøgende Juni 2017 — januar 2018	Randomiseret og kontrolleret forsøg (RCT)	BG: Unge i alderen 18-24 tilfældigt allokeret til at modtage den intensiverede støtte KG: Unge i alderen 18-24 tilfældigt allokeret til at modtage almindelig støtte BG: 2 792 — CG: 1 897	- andel af arbejdsløse - % af udskrivning til arbejde - gennemsnitligt antal dage med arbejdsløshed	12 måneder efter start	- Offentlige beskæftigelsesregistre for outcomevariable - Undersøgelse af medarbejdere på arbejdsformidling og formidlere med henblik på at måle støtteintensiteten
Italien (Marche)*	Virkningerne af ESF-interventioner finansieret i 2014-2020 for langtidsledige i Marche-regionen (2020)	Praktikophold, erhvervs erfaring, jobstipendier, uddannelsesvouchers til arbejdsløse (herunder langtidsledige) 2017 — juli 2019	Propensity score matching (matchning efter sandsynlighed) (PSM)	BG: Langtidsledige, der er behandlet KG: Langtidsledige i perioden 2016-2018 registreret hos de offentlige arbejdsformidlinger BG: 526 for praktikophold, 1 058 for jobstipendier, 236 for erhvervs erfaring i kommunerne og 241 for uddannelsesvouchers — KG: 77 255	- sandsynlighed for at være i beskæftigelse et bestemt tidsrum efter interventionerne - sandsynlighed for at være i beskæftigelse på en tidsubegrænset kontrakt - antal arbejdsdage i en periode efter interventionerne	6, 9, 12, 15 og 18 måneder efter påbegyndelse af interventionerne	- ESF-overvågningsdata for behandlingsgruppe - Administrative data fra offentlige arbejdsformidlinger om registrerede personer Administrative data om ansættelseskontrakter aktiveret i regionen (Comunicazioni Obbligatorie — COB)
Polen (Lubelskie)*	Evaluering af foranstaltninger for ældre arbejdstagere i Lubelskie finansieret af ESF i programmeringsperioden 2007-2013 (2015)	Flere typer aktive arbejdsmarkedspolitikker 2007-2013	PSM + Differences-in-differences (DID)	BG: Stikprøve af arbejdsløse 50+, som modtog støtte KG: Arbejdsløse i alderen 50+, som ikke deltog i nogen foranstaltning, og som var tilmeldt et regionalt arbejdsformidlingskontor på det tidspunkt, hvor medlemmerne af behandlingsgruppen deltog i det evaluerede program BG: 145 — KG: 67 102	- sandsynlighed for, at arbejdsløse vil blive udskrevet fra arbejdsløshedsregisteret	6, 12, 18 måneder efter start på interventionerne	- Administrative data fra offentlige arbejdsformidlinger

Land	Titel — år	Vurderet politik og observeret behandlingsperiode	Metode	Behandlingsgruppe og kontrolgruppe — definition og størrelse	Udfaldsvariabler	Observation periode	Data
Letland	Erhvervsuddannelse for unge arbejdsløse i Letland: Dokumentation fra et regressionsdiskontinuitetsdesign (2018)	Erhvervsuddannelse rettet mod NEET'er i alderen 15-29 2014-2015	Regressions diskontinuitetsdesign (RDD)	BG: Arbejdsløse, der har deltaget i erhvervsuddannelse senest et år efter registreringsdatoen KG: Arbejdsløse, der er registreret mellem juni 2013 og december 2015, og som ikke deltog i nogen foranstaltning BG: 898 — KG: 10 717	- sandsynlighed for at være i beskæftigelse et bestemt tidsrum efter interventionerne - indkomst	12-36 måneder efter start	- Administrative data fra Letlands statslige arbejdsformidlingskontor — oplysninger om både deltagere og ikkedeltagere, der er registreret som arbejdsløse på bestemte datoer - Administrative data fra skattemyndighederne — oplysninger om beskæftigelsesforhold på forskellige datoer og om personers indkomst
Tyskland	Beskæftigelsesmæssige virkninger af sprogundervisning for arbejdsløse indvandrere (2021)	Sprogkurser (+ arbejdserfaring for nogle af de behandlede) for jobsøgende indvandrere 2014	Instrumentvariabler (IV)	BG: Migrantdeltagere, der påbegyndte sprogundervisning til erhvervsformål i 2014 KG: Stikprøve af ikkedeltagere med mindst én periode med ikketyk statsborgerskab, der ikke deltager i aktive arbejdsmarkedspolitiske foranstaltninger eller integrationskurser BG: 8 968 — KG: 26 463	- sandsynlighed for at være i beskæftigelse (fast beskæftigelse, fast fuldtidsbeskæftigelse, fast beskæftigelse >6 måneder, beskæftigelse med indkomst over fattigdomsrisikotærsklen, faglært beskæftigelse) - daglig og kumuleret indkomst fra fast beskæftigelse - kumulerede dage i fast beskæftigelse	24 måneder efter start på interventionerne	- Oplysninger fra arbejdsmarkedsregisteret ("Integrierten Erwerbsbiografie", IEB) til identifikation af begge grupper og til udfaldsvariabler - Data indsamlet retrospektivt under møder med jobsøgende og sagsbehandlere ("Werdegangshistorie", WGH), for yderligere oplysninger om perioder med selvstændig virksomhed eller forældreorlov og om uddannelse og beskæftigelse i udlandet
Italien (Trento)	Er erhvervsuddannelsesprogrammerne prisværdige? Dokumentation fra en cost-benefit-analyse (2018)	Erhvervsuddannelseskurser for arbejdsløse 2010-2011	PSM	BG: Arbejdsløse deltagere i erhvervsuddannelser finansieret af provinsmidler og af ESF KG: Personer, der er registreret som arbejdsløse på startdatoen for individuelle uddannelseskurser i de lokale offentlige arbejdsformidlingsregistre BG: 818 (provinskurser), 114 (ESF-kurser) — KG: 17 236 (provinskurser), 1 152 (ESF-kurser)	- sandsynlighed for at være i beskæftigelse - indkomst	3, 6, 12, 18, 24, 36 måneder efter start på interventionerne	- ESF-overvågningsdata for behandlingsgruppe - Administrative data fra offentlige arbejdsformidlinger om registrerede personer - Administrative data om ansættelseskontrakter aktiveret i regionen (Comunicazioni Obbligatorie — COB) - Administrative data om selvangivelser fra skattemyndighederne - Administrative data fra provinsens undervisningsdepartement

Land	Titel — år	Vurderet politik og observeret behandlingsperiode	Metode	Behandlingsgruppe og kontrolgruppe — definition og størrelse	Udfaldsvariabler	Observation periode	Data
Tyskland	Forbedrer jobskabelsesordning er de langtidslediges sociale integration og trivsel? (2020)	Jobskabelsesordning (jobtilskud i op til 36 måneder) for sårbare personer (langtidsledige) 2015 — juni 2017	PSM	BG: Stikprøve af deltagere, som bevarede hele spørgeskemaet KG: Ikkedeltagere registreret hos offentlige arbejdsformidlinger og med samme støtteberettigelseskriterier for kontrolgruppen, som besvarede hele spørgeskemaet BG: 2 531 i runde 1, 1 191 i runde 2, 450 i runde 3 — KG: som ovenfor	- sandsynlighed for at være i beskæftigelse - subjektive mål for tilfredshed med livet, mental sundhed, social tilknytning, social status	7, 18, 29 måneder efter start på programmet	- Oplysninger fra arbejdsmarkedsregisteret ("Integrierten Erwerbsbiografien", IEB) til identifikation af begge grupper og til outcomevariabler - Undersøgelse af subjektive outcomevariabler (bløde outcomes)
Spanien (Asturien)*	Evaluerer af det operationelle program Asturien ESF (2019)	Undervisningsmetoder for grupper af elever, der er i risiko for at forlade skolen for tidligt (andet og tredje år på sekundærtrinnet) 2016-2017	PSM	BG: Elever, der deltager i interventionerne KG: Elever, der er sammenlignelige med hensyn til kriterier for støtteberettigelse BG: 1 053 — KG: 3 852	- sandsynlighed for at gå videre til næste skoleår	12 måneder efter start	- Administrative data om uddannelse: SAUCE-databasen (undervisningsministeriet)
Portugal	Evaluerer af systemet for tilskud til videregående uddannelser for mindre privilegerede studerende i Portugal (2020) — endnu ikke offentliggjort	Tilskud til studerende på videregående uddannelser fra lavindkomstfamilier 2021-2018	RDD	BG: Studerende, der ansøger om det første akademiske år på kandidat- eller bachelordannelse KG: Studerende, der ikke er støtteberettigede, fordi deres indkomst ligger over den fastsatte tærskel BG: 130 602 — KG: 25 400	- sandsynligheden for at være indskrevet ved udgangen af det første år - antal point opnået ved udgangen af det første år - sandsynlighed for at gennemføre uddannelsen - antal år til gennemførelse af uddannelsen - afsluttende karakter	12-48 måneder efter tilskuddet	- Administrative data om studerende, der ansøger om tilskud (DGES) - Administrative data med oplysninger om uddannelse og fremskridt (DGEEC)
Polen (Podlaskie)*	Vurdering af indvirkningen af støtte fra det operationelle program Podlaskie på udbredelsen af erhvervsuddannelse i 2014-2020 (2020)	ESF-projektet "Good Profession — Great Life", der har til formål at fremme erhvervsuddannelse og målrettede grundskoler på sekundærtrinnet 2017-2019	PSM	BG: Stikprøve af deltagende elever KG: Studerende med egenskaber, der ligner dem i den behandlede gruppe, men som ikke deltager i projektet BG: 200 — KG: 384	- indskrivning på en teknisk skole på tidspunktet for undersøgelsen - indskrivning på erhvervsskole på første trin på tidspunktet for undersøgelsen - indskrivning på enten teknisk skole eller erhvervsskole på første trin på tidspunktet for undersøgelsen - ansøgning til et gymnasium, til teknisk skole, til erhvervsskole på første trin, til skole, der tilbyder erhvervsuddannelse	12-36 måneder efter interventioner	- Overvågningsdata - Det nationale folkeregister (PESEL) til udvælgelse af kontrolgruppen - Undersøgelse til måling af outcomevariablerne

Land	Titel — år	Vurderet politik og observeret behandlingsperiode	Metode	Behandlingsgruppe og kontrolgruppe — definition og størrelse	Udfaldsvariable	Observation periode	Data
Italien (Piemonte)*	Beskæftigelsesmæssige virkninger af vouchere til arbejdsformidlingsordningen (Buono servizi lavoro — BSL finansieret af OP Piemonte ESF, 2014-2020) (2019)	Vouchers til arbejdsformidlingstjenester (orientering + uddannelse + trin) for sårbare arbejdsløse med en arbejdsløshedsperiode på mindst seks måneder 2018	PSM	BG: Arbejdsløse, der er blevet behandlet KG: Arbejdsløse, der er registreret hos offentlige arbejdsformidlinger med lignende karakteristika (30 år og arbejdsløse i seks måneder) BG: 8 125 — KG: 130 000	- sandsynlighed for at være i beskæftigelse - sandsynlighed for at være i beskæftigelse på en tidsubegrænset kontrakt	6, 12, 16 måneder efter interventioner	- ESF-overvågningsdata for behandlingsgruppe - Administrative data fra offentlige arbejdsformidlinger om registrerede personer - Administrative data om ansættelseskontrakter aktiveret i regionen (Comunicazioni Obbligatorie — COB)
Tyskland*	Evaluering af programmet for integration af langtidsløse i Tyskland i 2014-2020 (2019)	Integrationsforanstaltninger for langtidsløse i 24 måneder eller mere (jobtilskud) 2015-2017	DID (behandlingsintention blev estimeret)	BG: Langtidsløse, der potentielt er støtteberettigede i programperioden (fra 2015) KG: Personer, der opfylder kriterierne forud for programmets gennemførelse (2010-2012) BG og KG: 134 515 personer for 237 874 arbejdsløshedsperioder (35 % for BG og 65 % for KG)	- sandsynlighed for at være i beskæftigelse (beskæftigelse med bidrag til sociale ordninger)	24 måneder efter interventioner	- Administrative overvågningsdata fra forbundskontoret for administration (ZUWES-databasen) - Oplysninger fra arbejdsmarkedsregisteret ("Integrierten Erwerbsbiografie", IEB) for outcomevariable
Tyskland (Baden-Württemberg)	Anden evalueringsrapport om specifikt mål A1.1 under ESF-OP Baden-Württemberg, 2014-2020 (2020)	Foranstaltninger til integration af arbejdsløse, navnlig langtidsløse og andre sårbare personer Marts 2016 — december 2017	PSM	BG: Deltager identificeret i IEB-databasen KG: Personer, der opfylder kriterierne i samme periode BG: 1 578 (ud af i alt 1 800) — KG: na	- sandsynlighed for at være i beskæftigelse (beskæftigelse med bidrag til sociale ordninger) - sandsynligheden for fortsat arbejdsløshedsunderstøttelse	15 måneder efter interventioner	- Overvågningsdata - Oplysninger fra arbejdsmarkedsregisteret ("Integrierten Erwerbsbiografie", IEB) for outcomevariable

Bemærk: BG = behandlingsgruppe, KG = kontrolgruppe

Bemærk: * Evalueringer, der er identificeret i Helpdesk.

SÅDAN KONTAKTER DU EU

Personligt

Der findes flere hundrede Europe Direct-informationscentre i hele EU. Find dit nærmeste center på: https://europa.eu/european-union/contact_da

Pr. telefon eller e-mail

Europe Direct er en tjeneste, der besvarer spørgsmål om EU. Kontakt Europe Direct:

- på gratisnummeret: 00 800 6 7 8 9 10 11 (visse operatører tager betaling for disse opkald)
- på det almindelige telefonnummer +32 22999696 eller
- pr. e-mail: https://europa.eu/european-union/contact_da

SÅDAN FINDER DU OPLYSNINGER OM EU

Online

Oplysninger om EU er tilgængelige på alle EU's officielle sprog på Europawebstedet: https://europa.eu/european-union/index_da

EU-publikationer

Du kan downloade eller bestille EU-publikationer gratis eller mod betaling på: <https://op.europa.eu/da/publications>. Du kan bestille flere eksemplarer af de gratis publikationer ved at kontakte Europe Direct eller dit lokale informationscenter (se https://europa.eu/european-union/contact_en).

EU-retten og relaterede dokumenter

Du kan nemt få adgang til EU's juridiske oplysninger (herunder al EU-ret siden 1952) på alle officielle EU-sprog på EUR-Lex: <http://eur-lex.europa.eu>

Åbne data fra EU

På EU's portal for åbne data (<http://data.europa.eu/euodp/en/>) har du adgang til datasæt fra EU. Dataene kan downloades og genanvendes gratis til både kommercielle og ikkekommercielle formål.

