



European
Commission

UTFORMNING OCH UTFÖRANDE AV KONTRAFAKTISKA KONSEKVENSBEDÖMNINGAR

PRAKTISK VÄGLEDNING FÖR ESF:s
FÖRVALTNINGSMYNDIGHETER

EUROPEISKA KOMMISSIONEN

Generaldirektoratet för sysselsättning, socialpolitik och inkludering
Direktorat G – Fonder, programplanering och genomförande
Enhet G5 – Bättre lagstiftning

Kontakt: Linda Adamaite

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

*Europeiska kommissionen
B-1049 Bryssel*

UTFORMNING OCH UTFÖRANDE AV KONTRAFAKTISKA KONSEKVENSBEDÖMNINGAR

**PRAKTISK VÄGLEDNING FÖR ESF:s
FÖRVALTNINGSMYNDIGHETER**

Manuskriptet färdigställdes i december 2021.

2021 års upplaga av vägledningen ”Utformning och utförande av kontrafaktiska konsekvensbedömningar – Praktisk vägledning för ESF-förvaltningsmyndigheter” har skrivits av Jochen Kluge (Humboldt-Universität zu Berlin) samt Andrea Naldini och Marco Pompili (Ismeri Europa). Detta är en anpassning av 2013 års upplaga av vägledningen (ISBN 978-92-79-28238-6, DOI 10.2767/94454), som utarbetades av Stephen Morris (Policy Evaluation and Research Unit, Manchester Metropolitan University), Herta Tödting-Schönhofer (Metis GmbH, Wien) och Michael Wiseman (George Washington Institute of Public Policy).

Detta dokument har utarbetats för Europeiska kommissionen, men det återspeglar endast författarnas åsikter. Kommissionen tar inte ansvar för eventuella konsekvenser av att dokumentet vidareutnyttjas. Mer information om EU finns på internet (<http://www.europa.eu>).

Luxemburg: Europeiska unionens publikationsbyrå, 2021

© Europeiska unionen, 2021



Vidareutnyttjandet av Europeiska kommissionens handlingar regleras av kommissionens beslut 2011/833/EU av den 12 december 2011 om vidareutnyttjande av kommissionens handlingar (EUT L 330, 14.12.2011, s. 39). Om inget annat anges är det tillåtet att vidareutnyttja denna handling inom ramen för licensen Creative Commons Attribution 4.0 International (CC-BY 4.0) (<https://creativecommons.org/licenses/by/4.0/>). Detta innebär att vidareutnyttjande är tillåtet förutsatt att källa anges på lämpligt sätt liksom eventuella ändringar.

Tillstånd för användning och reproduktion av delar som inte ägs av EU kan behöva sökas direkt från respektive upphovsrättsinnehavare.

Omslag: © Shutterstock, 2021

PDF ISBN 978-92-76-40725-6

doi:10. 10.2767/02762

KE-02-21-992-EN-N

Innehåll

INLEDNING: VÄGLEDNINGENS BAKGRUND OCH SYFTE	5
KAPITEL 1. BEGREPP OCH STRATEGIER	11
1.1. VÄSENTLIGA ASPEKTER AV KONTRAFAKTISKA ANALYSER	11
1.2. VARFÖR ÄR KONTRAFAKTISKA BEDÖMNINGAR VIKTIGA?	12
1.3. VARFÖR ÄR KONTRAFAKTISKA KONSEKVENSBEDÖMNINGAR TEKNISKT KOMPLICERADE?	13
1.4. ÖVERSIKT AV OLIKA METODER FÖR ATT UTFORMA OCH GENOMFÖRA KONTRAFAKTISKA KONSEKVENSBEDÖMNINGAR..	15
1.5. HUR KONTRAFAKTISKA KONSEKVENSBEDÖMNINGAR KAN INTEGRERAS I EN BREDARE UTVÄRDERINGSRAM.....	16
KAPITEL 2. PRAKTISKA ÖVERVÄGANDEN VID UTFORMNINGEN AV KONTRAFAKTISKA KONSEKVENSBEDÖMNINGAR	23
2.1. VÄLJA INSATSER SOM SKA KONSEKVENSBEDÖMAS	26
2.1.1. <i>Prioritera insatser för konsekvensbedömning</i>	28
2.1.2. <i>Välja insatser som är lämpade för kontrafaktiska konsekvensbedömningar</i>	30
2.2. FASTSTÄLLA UTVÄRDERINGSFRÅGOR OCH UTFALLSVARIABLER	34
2.2.1. <i>Vilka syften och mål har insatsen?</i>	34
2.2.2. <i>Vad är syftet med utvärderingen?</i>	34
2.3. DATA FÖR ATT FASTSTÄLLA KONTROLLGRUPPEN OCH MÄTA UTFALLSVARIABLERNA	39
2.3.1. <i>Finns de uppgifter som behövs tillgängliga eller är det möjligt att få tillgång till dem?</i>	39
2.3.2. <i>Hur identifierar man "behandlingsgruppen"?</i>	51
2.3.3. <i>Faktorer som ska övervägas vid fastställandet av kontrollgrupper</i>	52
2.3.4. <i>Vilka uppgifter bör ingå i utvärderingsschemat?</i>	56
2.3.5. <i>Vilka är de viktigaste hindren i data- och resultatanalysen?</i>	58
2.3.6. <i>Checklista för att kontrollera om en kontrafaktisk konsekvensbedömning är ett genomförbart alternativ</i> .	60
2.4. METODEN FÖR DEN KONTRAFAKTISKA KONSEKVENSBEDÖMNINGEN.....	61
2.5. TIDTABELL OCH BUDGET	62
2.5.1. <i>Vilka resurser finns tillgängliga?</i>	62
2.5.2. <i>När bör insatsen utvärderas?</i>	67
2.6. GENOMFÖRANDE AV KONTRAFAKTISKA KONSEKVENSBEDÖMNINGAR	69
2.6.1. <i>Välja utvärderare</i>	69
2.6.2. <i>Övervakning</i>	71
2.6.3. <i>Rapportering</i>	71
2.6.4. <i>Använda resultaten</i>	73
KAPITEL 3. HUR VÄLJER JAG EN LÄMPLIG METOD FÖR MIN KONTRAFAKTISKA KONSEKVENSBEDÖMNING?	76
3.1. RANDOMISERING – DEN EXPERIMENTELLA METODEN	76
3.2. ICKE-RANDOMISERADE ELLER KVASIEXPERIMENTELLA METODER.....	79
3.2.1 <i>Mål- och kontrollgrupper utan randomisering</i>	79
3.2.2 <i>Benägenhetspoäng (Propensity score matching)</i>	81
3.2.3 <i>Skillnaden-mellan-skillnader</i>	83
3.2.4 <i>Regressionsdiskontinuitetsdesign</i>	86
3.2.5 <i>Instrumentvariabler</i>	89
KAPITEL 4. FRÄMJA ANVÄNDNINGEN AV KONTRAFAKTISKA KONSEKVENSBEDÖMNINGAR	95

4.1. FÖRBÄTTRA FÖRSTÅELEN BLAND BERÖRDA AKTÖRER.....	95
4.2. KAPACITETSUPPBYGGNAD	96
4.3. LÖSA RÄTTSLIGA HINDER	99
4.4. EN FRAMÅTBlickande STRATEGI	99
4.5. UTVIDGNING AV TILLÄMPNINGsomRÅDET FÖR KONTRAFAKTISKA KONSEKVENSBEDÖMNINGAR.....	100
ORDLISTOR.....	105
FÖRKORTNINGAR	105
DEFINITIONER	106
BILAGOR	114
BILAGA 1. LÄS MER	114
BILAGA 2. FÖRSLAG TILL UPPLÄGG FÖR ONLINEKURSER OM KONTRAFAKTISKA KONSEKVENSBEDÖMNINGAR	118
BILAGA 3. KONTRAFAKTISKA KONSEKVENSBEDÖMNINGAR – EXEMPEL SOM NÄMNS I VÄGLEDDNINGEN.....	119

TEXTRUTOR

RUTA 1 EXEMPEL PÅ EN KONTRAFAKTISK KONSEKVENSBEDÖMNING FÖR JÄMFÖRELSE AV FÖRDELAR OCH KOSTNADER	22
RUTA 2 UTVÄRDERING AV KONTRAFAKTISKA KONSEKVENSBEDÖMNINGAR INOM EN BREDARE RAM.....	26
RUTA 3 URVALSFRÅGOR FÖR ATT VÄLJA INSATSER SOM SKA GENOMGÅ EN KONTRAFAKTISK KONSEKVENSBEDÖMNING.....	27
RUTA 4 DE VANLIGASTE TYPERNA AV INSATSER OCH MÅLGRUPPER SOM VÄLJS FÖR KONTRAFAKTISKA KONSEKVENSBEDÖMNINGAR INOM RAMEN FÖR ESF.....	30
RUTA 5 FASTSTÄLLA KONTROLLGRUPPER.....	33
RUTA 6 EXEMPEL PÅ DATA SOM ANVÄNDS FÖR KONTRAFAKTISKA KONSEKVENSBEDÖMNINGAR.....	41
RUTA 7 EXEMPEL PÅ INTEGRERADE DATABASER FÖR KONTRAFAKTISKA KONSEKVENSBEDÖMNINGAR	42
RUTA 8 EU:S REGELVERK OM BEHANDLING AV PERSONUPPGIFTER	46
RUTA 9 DATASKYDD OCH DATAUTBYTE.....	50
RUTA 10 POLITISKA FRÅGOR I SAMBAND MED ETT UTBILDNINGSPROGRAM.....	55
RUTA 11 TOLKA NETTOEFFEKTER	55
RUTA 12 OSÄKERHETER I TOLKNINGEN AV RESULTAT.....	60
RUTA 13 DEN POLSKA ERFARENHETEN AV UTVÄRDERINGSKONFERENSER.....	75
RUTA 14 EXEMPEL PÅ EN RANDOMISERAD PRÖVNING AV ETT ESF-PROJEKT FÖR UNGDOMAR	78
RUTA 15 EXEMPEL PÅ EN UTVÄRDERING SOM TILLÄMPAR EN MATCHNINGSMETOD	83
RUTA 16 EXEMPEL PÅ EN UTVÄRDERING SOM TILLÄMPAR METODEN SKILLNADEN-MELLAN-SKILLNADER.....	85
RUTA 17 EXEMPEL PÅ EN UTVÄRDERING SOM TILLÄMPAR EN REGRESSIONSDISKONTINUITETSMETOD.....	89
RUTA 18 EXEMPEL PÅ EN STUDIE SOM TILLÄMPAR EN METOD MED INSTRUMENTVARIABLER	91
RUTA 19 EXEMPEL PÅ ETT PROJEKT SOM INRIKTAS PÅ ATT STÄRKA KULTUREN OCH KAPACITETEN NÄR DET GÄLLER KONTRAFAKTISKA KONSEKVENSBEDÖMNINGAR	98
RUTA 20 EXEMPEL PÅ UTVÄRDERINGAR PÅ UTBILDNINGSSOMRÅDET.....	101
RUTA 21 EXEMPEL PÅ BEDÖMNING AV EFFEKTER PÅ "MJUKA UTFALL" I TYSKLAND	103

FIGURFÖRTECKNING

FIGUR 1 OLIKA ÄNDAMÅL MED OCH OLIKA TYPER AV UTVÄRDERINGAR	17
FIGUR 2 ILLUSTRATION AV LOGIKMODELLEN ELLER "FÖRÄNDRINGSTEORIN"	19
BILD 3 ORDNINGSFÖLJD FÖR DE VIKTIGASTE STEGEN I EN KONTRAFAKTISK KONSEKVENSBEDÖMNING.....	24
DIAGRAM 4 MINSTA DETEKTERBARA EFFEKTSTORLEKAR VID OLIKA URVALSSTORLEKAR	59
FIGUR 5 FÖRENKLAD TIDSPÄN FÖR RESULTAT FRÅN ETT UTBILDNINGSPROGRAM.....	68
FIGUR 6 RANDOMISERAD KONTROLL MED TVÅ GRUPPER	77
FIGUR 7 STILISERAD KVASIEXPERIMENTELL DESIGN MED BEHANDLINGS- OCH KONTROLLGRUPPER	80
DIAGRAM 8 ILLUSTRATION AV METODEN MED BENÄGENHETSPOÄNG	82
DIAGRAM 9: FÖRKLARING AV SKILLNADEN-MELLAN-SKILLNADER-METODEN.....	85
DIAGRAM 10 ILLUSTRATION AV REGRESSIONSDISKONTINUITETSMETODEN	87
FIGUR 11 ILLUSTRATION AV METODEN MED INSTRUMENTVARIABLER.....	90

TABELLFÖRTECKNING

TABELL 1 REKOMMENDERAT INNEHÅLL I ETT UTVÄRDERINGSSHEMA.....	25
TABELL 2 DATATYPER OCH KÄLLOR.....	41
TABELL 3 STRUKTUR FÖR DE HUVUDSAKLIGA KOSTNADERNA FÖR KONTRAFAKTISKA KONSEKVENSBEDÖMNINGAR	66
TABELL 4 GRUNDLÄGGANDE INFORMATION SOM SKA INGÅ I ETT INFORMATIONSBLAD OM DEN KONTRAFAKTISKA KONSEKVENSBEDÖMNINGEN	73
TABELL 5 JÄMFÖRELSE AV VIKTIGA EGENSKAPER HOS DE VANLIGASTE METODERNA FÖR KONTRAFAKTISKA KONSEKVENSBEDÖMNINGAR	92
TABELL 6 KÄNNETECKEN HOS DE KONTRAFAKTISKA KONSEKVENSBEDÖMNINGAR SOM NÄMNS SOM EXEMPEL I VÄGLEDNINGEN	119

Inledning: vägledningens bakgrund och syfte

Programperioden 2021–2027 inleddes med de dramatiska erfarenheterna av covid-19-pandemin och den därpå följande ekonomiska krisen. Europeiska unionen ökade sina finansiella och politiska insatser för att bidra till återhämtningen av de nationella ekonomierna och öka sysselsättningen. Inom ramen för detta spelar Europeiska socialfonden+ (ESF+) en viktig roll genom att ge ett omfattande stöd till arbetslösa, med fokus på ungdomar och kvinnor, stödja insatser mot barnfattigdom och främja bättre utbildning och social inkludering av svagare samhällsgrupper inom EU. De olika målen för ESF+ och behovet av att nå snabba resultat på områdena sysselsättning och social inkludering kräver en effektiv resursfördelning. Evidensbaserade politiska strategier är viktigare än någonsin, och utvärdering är ett grundläggande instrument för att vägleda den offentliga politiken.

Under programperioden 2014–2020 bidrog Europeiska kommissionen till medlemsstaternas och de förvaltande myndigheternas utvärderingskapacitet och förespråkade användning av kontrafaktiska konsekvensbedömningar. De positiva resultaten av kontrafaktiska konsekvensbedömningar är ett konkret bevis på att sådana konsekvensbedömningar förstärker Europeiska socialfondens effekter. Ett stort antal kontrafaktiska konsekvensbedömningar lades in i utvärderingsplanerna inom ramen för genomförandet av ESF-programmen. I många fall uppstod svårigheter under förberedelserna och genomförandet av de kontrafaktiska konsekvensbedömningarna, medan de i andra fall begränsades till isolerade erfarenheter som inte integrerades i en systematisk utvärderingsram.

Denna vägledning är avsedd för förvaltningsmyndigheter och andra organ med ansvar för genomförandet av insatser och program som finansieras av ESF+, och syftar till att stödja planering, utformning och utförande av kontrafaktiska konsekvensbedömningar. Tidigare erfarenheter beaktas i vägledningen och den innehåller praktiska råd om några av de viktigaste frågor som bör övervägas i utformningen av kontrafaktiska konsekvensbedömningar. Denna vägledning är således en uppdatering av den föregående vägledningen för programperioden 2014–2020, och i denna upplaga läggs mer tonvikt på problem som har uppstått i det praktiska genomförandet av kontrafaktiska konsekvensbedömningar. Metodologiska aspekter diskuteras också, och vägledningen har förenklats där så har varit möjligt, med uppdaterade exempel som valts ut från utvärderingarna av ESF för programperioden 2014–2020.

Kontrafaktiska konsekvensbedömningar handlar om viktiga frågor som möjliggör evidensbaserade politiska beslut: Vilka kausala verkningar får insatserna och "vad fungerar"? Här ger kontrafaktiska konsekvensbedömningar belägg för om ESF-finansierade insatser faktiskt är orsaken till förändringar i deltagarnas förhållanden och

En vägledning för förvaltningsmyndigheter som uppdaterats med erfarenheterna från programperioden 2014–2020

Kontrafaktiska konsekvensbedömningar visar "vad som fungerar"

uppnåendet av insatsens mål. Väl genomförda kontrafaktiska konsekvensbedömningar ger belägg för "nettoeffekten"¹ eller verkan av en insats, vilket ger beslutsfattare möjlighet att utesluta alternativa förklaringar till förändringar i deltagarnas förhållanden eller resultat som kan ha observerats. När kontrafaktiska konsekvensbedömningar ger uppskattningar av förekomsten eller omfattningen av en nettoeffekt har uppskattningarna nödvändigtvis en osäkerhetsgrad, beroende på metodologisk noggrannhet och tillgänglig information. Den typ av belägg som fås fram genom kontrafaktiska konsekvensbedömningar ger beslutsfattarna möjlighet att bedöma insatsernas ändamålsenlighet, jämföra olika insatser och bedöma deras relativa resultat. De ger dessutom viktiga bidrag till kostnads-nyttoanalyser och kostnadseffektivitetsanalyser.

Denna vägledning offentliggörs vid en tidpunkt då ESF+ står inför helt nya utmaningar. Med tanke på den enorma ökningen av EU:s medel för investeringar och sysselsättning enligt EU:s återhämtningsinstrument Next generation EU, är det mycket viktigt att beslutsfattarna mäter och förstår effekterna av de insatser som de ansvarar för. Offentliga medel måste anslås till mer produktiva och ändamålsenliga insatser för att påskynda återhämtningen och minska de sociala obalanserna. De som ansvarar för utbetalningen av medel från ESF+ ska därför motivera sina val genom att visa att deras insatser är ändamålsenliga och ger värde för människorna. Det bästa sättet att åstadkomma detta är att öka antalet kontrafaktiska konsekvensbedömningar av god kvalitet.

ESF+ är EU:s viktigaste instrument för att stödja sysselsättning och social inkludering. Under programperioden 2014–2020 användes nästan 125 miljarder euro från ESF för aktiva arbetsmarknadspolitiska åtgärder, utbildningspolitiska åtgärder och åtgärder för social inkludering, vilka genomfördes via de operativa programmen i de 28 medlemsländerna. Som anges i förordningen om allmänna bestämmelser (förordning (EU) nr 1303/2013) är syftet med utvärderingar att "förbättra kvaliteten på programmets utformning och genomförande samt att bedöma deras effektivitet och verkan".

Prestationer och resultat kommer att fortsätta att granskas under programperioden 2021–2027². Detta kräver en förstärkning av de nuvarande övervaknings- och utvärderingssystemen och kapaciteten för dessa, inbegripet insamling av uppgifter. Utvärderingsplaner är fortfarande obligatoriska och ökad tonvikt läggs vid konsekvensbedömningar. Eftersom det finns ett antal olika metoder för att mäta effekterna av de insatser som får stöd från ESF+ måste de förvaltande myndigheterna besluta vilken metod, eller kombination av metoder, som är lämpligast för att uppfylla de lagstadgade kraven. En noggrann kvantifiering av insatsernas effekter omfattar även en kontrafaktisk analys.

Resultatorientering och utvärderingar av hög kvalitet

¹ Nettoeffekten, eller verkan, är skillnaden mellan total effekt (eller bruttoeffekt) och de resultat som skulle ha uppnåtts utan insatsen. Nettoeffekten kan även vara negativ när insatsen är mindre effektiv än marknadsdynamiken. Se s.9 för en närmare beskrivning.

² Europaparlamentets och rådets förordning (EU) 2021/1060 av den 24 juni 2021 om fastställande av gemensamma bestämmelser för Europeiska regionala utvecklingsfonden, Europeiska socialfonden+, Sammanhållningsfonden, Fonden för en rättvis omställning och Europeiska havs-, fiskeri- och vattenbruksfonden samt finansiella regler för dessa och för Asyl-, migrations- och integrationsfonden, Fonden för inre säkerhet samt instrumentet för ekonomiskt stöd för gränsförvaltning och viseringspolitik.

Inriktningen på goda resultat och resultatorientering är ett viktigt inslag i de nya förordningarna. Utvärderingsstrategier och tekniker av hög kvalitet är en grundläggande förutsättning för att förvärva viktig kunskap som visar medlemsländerna vilka insatser som fungerar och vilka inte. Det är därför viktigt att förbättra kvaliteten på utvärderingarna och inhämta tillförlitliga belägg för insatsernas mervärde.

Utgångspunkten för att inhämta belägg för politiska insatsers ändamålsenlighet är i princip enkel: Följande krav gäller:

- Identifiera det problem som ska åtgärdas.
- Fastställa vilka instrument som ska användas för att åtgärda problemet.

Skapa en metod som kopplar ihop instrumenten och resultaten.

För att utvärdera ett finansieringssystem eller finansieringsinstrument är det viktigt att ha tydliga och mätbara indikatorer, både för de input som tillämpas och för problemrelaterade output och resultat. Det är vanligt att fastställa mål både för output och resultat, och jämföra de faktiska resultaten med målen. Övervakningen syftar till att spåra input, output och resultat över tid och ge återkoppling till ledningen. Den underliggande insatsteorin visar ofta på delresultat som även de kan bli föremål för övervakning. Att fastställa om en viss insats fungerar utifrån detta är däremot inte enkelt.

Behovet av resultat ...

Det finns ett behov av att komplettera befintliga utvärderingsmetoder med strategier som ger mycket starkare belägg för insatsernas nettoeffekter eller genomslag. Resultatmätning handlar om att redovisa de använda medlen. Kontrafaktiska konsekvensbedömningar visar om en insats är ändamålsenlig eller ej. Kontrafaktiska konsekvensbedömningar syftar till att fastställa ett orsakssamband mellan insatser och resultat, men det kan krävas ytterligare teoretiska metoder och processutvärderingsmetoder för att identifiera de underliggande orsakssambanden och bidra till att säkerställa att effekter som nåtts på en plats utgör en evidensbas för att genomföra insatsen på en annan plats.

... och belägg för nettoeffekter

Under programperioden 2014–2020 införde samtliga medlemsländer och förvaltningsmyndigheter en utvärderingsplan, med en beskrivning av mål, verksamheter, organisatoriska aspekter och resultat av utvärderingarna. Utvärderingsplanerna omfattar olika typer av utvärderingar, exempelvis övergripande löpande utvärderingar av programmen, och även tematiska utvärderingar som syftar till att besvara specifika utvärderingsfrågor eller inriktas på särskilda delar eller investeringsprioriteringar i ett operativt program. En analys av utvärderingsplanerna för 177 ESF-program vid utgången av 2018³ visade att omkring 132 kontrafaktiska konsekvensbedömningar var inplanerade under denna programperiod. Denna siffra visar att man har gjort framsteg och att kontrafaktiska metoder har getts större utrymme under programperioden 2014–2020 jämfört med 2007–2013⁴.

Medlemsländernas erfarenheter av kontrafaktiska konsekvensbedömningar

Enligt den projektkatalog som hjälpcentralen för utvärdering driver⁵ hade 1 795 utvärderingar slutförts fram till juni 2021, varav 1 001 rörde

Användningen av kontrafaktiska

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

⁴ Se [Bratu C. m.fl., 2014](#).

⁵ Hjälpcentralprojektet, som finansieras av GD Regionalpolitik och GD Sysselsättning, socialpolitik och inkludering, samlar in

program inom ESF/sysselsättningsinitiativet för unga eller ESF/sysselsättningsinitiativet för unga och Eruf-program (675 respektive 326). Av de 1 001 utvärderingarna som omfattade ESF var 323 konsekvensbedömningar och av dessa tillämpades en kontrafaktisk metod i 106 utvärderingar: 82 av of 234 konsekvensbedömningar rörde program inom ESF/sysselsättningsinitiativet för unga och 24 av 89 konsekvensbedömningar avsåg ESF/sysselsättningsinitiativet för unga och Eruf-program⁶.

konsekvensbedömningar ökar

Under de första åren av den innevarande programperioden inriktades de kontrafaktiska konsekvensbedömningarna av program inom ESF/sysselsättningsinitiativet för unga och eller Eruf-program på den föregående programperioden, medan nästan alla kontrafaktiska konsekvensbedömningar sedan 2018 bedömer effekterna av insatser under programperioden 2014–2020⁷. Tjugo av tjugosju medlemsländer genomförde kontrafaktiska konsekvensbedömningar av program som omfattades av ESF, vilket även i detta fall visar på en mer omfattande användning av denna metod jämfört med tidigare.

Även om medlemsländerna tillämpar olika metoder för kontrafaktiska konsekvensbedömningar var detta den metod som oftast användes för att bedöma effekterna av aktiva arbetsmarknadspolitiska åtgärder (utbildning, incitament, stöd vid arbetssökande, yrkeserfarenhet). Den användes däremot i mycket mindre utsträckning inom andra områden, t.ex. insatser för att stödja utbildningsinstitutioner och elever och socialpolitiska åtgärder, t.ex. insatser för att bekämpa fattigdom eller social utestängning.

Om man ser till metoderna är benägenhetspoäng den vanligast använda tekniken, medan andra metoder är mindre vanliga. Kontrafaktiska konsekvensbedömningar inriktas ofta på effekter på kort eller medellång sikt (sex eller tolv månader), medan långsiktiga effekter bedöms mer sällan.

Analysen av de kontrafaktiska konsekvensbedömningar som genomfördes inom ramen för hjälpcentralprojektet visade på vissa begränsningar, t.ex. personer i kontrollgrupperna var inte tillräckligt lika personerna i stödgrupperna, uppgifterna var inte tillförlitliga, t.ex. när de hämtades från efterhandsutvärderingar, intervjuerna byggde på respondenternas förmåga att korrekt minnas händelser, små urvalsstorlekar användes som gav statistiskt obetydliga resultat osv. I vissa fall innehöll utvärderingsrapporterna inte tillräckligt med information om metodsteg, val och begränsningar. Dessa uppgifter är nödvändiga för att informera andra personer än de som har utfört bedömningen om effekterna av de undersökta åtgärderna och komplettera deras kunskap om effekterna.

men det finns fortfarande begränsningar på många håll

Sammanfattningsvis har framsteg gjorts under de senaste åren, men kontrafaktiska konsekvensbedömningar kräver såväl teknisk sakkunskap som politisk vilja. Denna vägledning förespråkar användningen av

Vägledning för yrkesverksamma

information om utvärderingar som publicerats på de förvaltande myndigheternas webbplatser sedan den 1 januari 2015. Utvärderingarna kartläggs av ett nätverk av sakkunniga på nationell nivå och sammanfattas och bedöms utifrån kvalitet och resultatens tillförlitlighet.

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

⁶ Hjälpcentralprojektet granskade och sammanfattade 90 av 106 program.

⁷ Av de kontrafaktiska konsekvensbedömningar som utfördes fram till 2018 avser 29 av 33 den föregående programperioden.

kontrafaktiska konsekvensbedömningar och belyser några av de frågor som de förvaltande myndigheterna behöver överväga för att ta fram effektiva kontrafaktiska konsekvensbedömningar. Förutom de praktiska aspekterna av kontrafaktiska konsekvensbedömningar lyfter vi även fram mer allmänna problem som kan behöva lösas för att öka konsekvensbedömningarnas kvalitet. Denna vägledning innehåller ett antal grundläggande rekommendationer som förhoppningsvis är till hjälp för de förvaltande myndigheterna, men ett av huvudbudskapen är att det krävs noggrann planering av databehandlingen (särskilt administrativa data), tydliga syften och mål osv. för att undvika eventuella problem i genomförandefasen eller utvärderingar av dålig kvalitet.

Vägledningen innehåller även exempel på utvärderingar och metoder för att hjälpa läsaren. De utvärderingar som har samlats in inom ramen för projektet för hjälpcentralen för utvärdering utgör den huvudsakliga informationskällan för de exempel och metoder som presenteras i vägledningen, men även andra källor beaktades, särskilt de utvärderingar som har genomförts av Gemensamma forskningscentrumet (*Centre for Research on Impact Evaluation (CRIE)*) och, i mer begränsad utsträckning, akademiska publikationer. Exempelen och erfarenheterna i vägledningen avser om möjligt ESF-insatser.

Vägledningen är indelad i fyra kapitel.

I **kapitel 1** diskuteras kontrafaktiska konsekvensbedömningars kännetecken och varför de är viktiga. Här ges en inledande översikt av olika strategier för kontrafaktiska konsekvensbedömningar, med betoning på skillnaden mellan experimentella och kvasiexperimentella strategier. Kapitlet beskriver även tekniska och praktiska problem som kan uppstå i samband med kontrafaktiska konsekvensbedömningar och hur de kan lösas. Olika typer av politiska frågor som kan hanteras med hjälp av dessa bedömningar och sambandet mellan metoder för kontrafaktiska konsekvensbedömningar och andra utvärderingsmetoder (t.ex. teoretiska metoder, processutvärdering och effektivitetsanalys) diskuteras allmänt.

En struktur som avspeglar stegen i en utvärdering

I **kapitel 2** diskuteras ett antal frågor som de förvaltande myndigheterna bör tänka på när de utformar och genomför kontrafaktiska konsekvensbedömningar. Denna vägledning belyser några av de viktigaste utmaningar som kan uppstå under utarbetandet av kontrafaktiska konsekvensbedömningar och innehåller rekommendationer om hur dessa problem kan lösas. Frågorna fungerar som vägledning för dig som planerar att utföra kontrafaktiska konsekvensbedömningar av ESF-finansierade insatser.

Kontrafaktiska konsekvensbedömningar ger robusta belegg för fonders ändamålsenlighet. Detta kräver dock att de planeras och utförs väl. De förvaltande myndigheterna måste därför överväga vissa viktiga frågor i samband med utvärderingen. Hur de förvaltande myndigheterna tar sig an dessa frågor och i vilken ordning de behandlas styrs av de praktiska och institutionella arrangemangen i respektive medlemsland. I denna vägledning uppmärksammar vi de förvaltande myndigheterna på några av de dessa viktiga frågor. Frågorna behandlas enligt de sex stegen i utvärderingsprocessen: 1) Välja den eller de insatser som ska bedömas. 2) Fastställa utvärderingsfrågor och utfallsvariabler. 3) Analysera data för att fastställa kontrollgruppen och mäta utfallsvariablerna. 4) Välja metod

för den kontrafaktiska konsekvensbedömningen. 5) Fastställa tidtabell och budget. 6) Genomföra utvärderingen. Särskild uppmärksamhet ägnas åt datatillgång och dataskydd, eftersom dessa aspekter i stor utsträckning kan begränsa möjligheterna att utföra en kontrafaktisk konsekvensbedömning.

I **kapitel 3** undersöks steg 4 ovan (välja metod för den kontrafaktiska konsekvensbedömningen), och kapitlet inriktas på kännetecknen hos de vanligaste metoderna för kontrafaktisk analys. Fem metoder presenteras: Experimentmetoden eller randomisering, benägenhetspoäng, skillnaden-mellan-skillnader, regressionsdiskontinuitetsdesign och instrumentvariabler. Detta avsnitt av vägledningen är inte avsett att utbilda dig som läsare i hur du använder metoder för kontrafaktiska konsekvensbedömningar, utan syftar till att göra dig medveten om varje metods starka och svaga sidor så att du kan avgöra när en metod är lämpligare än en annan.

I **kapitel 4** behandlas mer allmänna frågor om perspektivutveckling. Här diskuteras bland annat behovet av att utveckla kapacitet för att genomföra kontrafaktiska konsekvensbedömningar på ett effektivt sätt, både inom de förvaltande myndigheterna (beslutsfattare och tjänstemän) och bland forskare, akademiker och konsultföretag ute i medlemsländerna. I detta avsnitt diskuteras även rättsliga hinder i fråga om tillgång till uppgifter och hur kontrafaktiska konsekvensbedömningar av ESF+-program kan uppdateras med en mer avancerad utformning för att öka täckningen av politiska utbildnings- och inkluderingsåtgärder, uppskattningar av effekterna av "mjuka resultat" och genomförandet av metautvärderingsstrategier.

Sammanfattningsvis har denna vägledning följande syften: 1) Förespråka användningen av kontrafaktiska konsekvensbedömningar, 2) Beskriva viktiga steg för ett effektivt utförande av kontrafaktiska konsekvensbedömningar, och 3) etablera kontrafaktiska konsekvensbedömningar som ett centralt inslag i ESF+. Det slutliga målet är att kontrafaktiska konsekvensbedömningar ska bidra på ett bättre sätt till ESF+, och därmed till EU-invånarnas välbefinnande.

Författarna vill tacka för det värdefulla stöd och hjälp som personalen vid enhet G5 vid GD Sysselsättning (Bättre lagstiftning) har bistått med under utarbetandet av denna rapport, särskilt Linda Adamaite, Jeannette Monier och Maria José Cueto Faus. Författarna vill även tacka Andrea Pisano och Ernesto Belisario för deras bidrag till avsnittet om dataskydd och Paweł Hess för hans stöd och bidrag.

Kapitel 1. Begrepp och strategier

I detta kapitel behandlas grundläggande frågor om hur kontrafaktiska konsekvensbedömningar bidrar till att bedöma insatsers kausala verkan. Syftet är närmare bestämt att skapa förståelse kring de väsentliga aspekterna av kontrafaktiska konsekvensbedömningar, särskilt i samband med olika typer av insatser som samfinansieras via ESF+. Här undersöker vi även sambandet mellan kontrafaktiska metoder och andra utvärderingsmetoder och diskuterar varför kontrafaktiska konsekvensbedömningar är viktiga, särskilt för närvarande. Kapitlet innehåller en kort introduktion till de politiska frågor som kan hanteras med hjälp av kontrafaktiska konsekvensbedömningar och ger en kortfattad översikt av några av de viktigaste kontrafaktiska metoderna som är relevanta för att utvärdera insatser som samfinansieras via ESF+. En detaljerad presentation ges sedan i kapitel 3.

1.1. Väsentliga aspekter av kontrafaktiska analyser

Syftet med kontrafaktiska konsekvensbedömningar är att kartlägga nettoeffekter eller verkningar av insatser. Det som utmärker kontrafaktiska konsekvensbedömningar är att de ger stöd för påståenden om att en viss insats ger ett visst resultat, det vill säga att ett visst resultat endast kan tillskrivas den insatsen. Detta uppnås genom att insatsen isoleras och alternativa förklaringar till det observerade resultatet utesluts.

Grunden till kontrafaktiska konsekvensbedömningars förmåga att utesluta alternativa förklaringar är begreppet "kontrafaktisk", det vill säga svaret på frågan "vad skulle ha hänt – när det gäller resultaten av intresse – om insatsen inte hade genomförts"? För att skapa en tydlig förståelse av begreppet kontrafaktisk och mycket enkelt uttryckt för att klargöra begreppet, kan vi ta exemplet med en arbetslös person som deltar i ett utbildningsprogram som har ett sysselsättningsfrämjande mål. För att fastställa vilken effekt utbildningen har haft på personen föreställer man sig två möjliga resultat⁸ i den kontrafaktiska analysen. Det första är deltagarens sysselsättningsstatus efter att ha deltagit i utbildningen. Detta är det observerade resultatet för deltagaren. Det andra möjliga resultatet är deltagarens sysselsättningsstatus om han eller hon inte hade deltagit i utbildningsprogrammet när alla andra faktorer är lika. I detta fall motsvarar det andra resultatet det kontrafaktiska resultatet. Utbildningens effekt på den enskilda deltagaren fastställs genom skillnaden mellan observerade och kontrafaktiska resultat. Detta är den kausala nettoeffekten eller verkan av utbildningen på personen. Den enda skillnaden i de omständigheter eller förhållanden som gav upphov till de observerade och de kontrafaktiska resultaten är personens deltagande i utbildningen. Eventuella skillnader mellan de två resultaten måste därför vara utbildningens effekt på personens sysselsättningsstatus.

Kontrafaktisk situation

⁸ Du hittar en mer detaljerad diskussion om kausalitetsmodellen med "möjliga resultat" i [Holland P., 1986](#).

I själva verket observerar vi inte och kan inte heller observera kontrafaktiska resultat för personer som omfattas av en insats. Det huvudsakliga syftet med kontrafaktiska konsekvensbedömningar är dock att ge övertygande uppskattningar av kontrafaktiska resultat för grupper av personer eller företag som påverkas av insatser som samfinansieras via ESF+. Effekterna uttrycks därför till exempel i form av skillnader i medelvärden eller andelar mellan det observerade genomsnittet och "uppskattade" kontrafaktiska värden. I de flesta fall är syftet med kontrafaktiska konsekvensbedömningar att jämföra resultaten av en insats (en åtgärd eller ett instrument) för de enheter (personer, små och medelstora företag osv.) som gynnades av insatsen med en grupp som inte omfattas av insatsen. Enligt terminologin för kontrafaktiska konsekvensbedömningar skiljer sig den "behandlade" gruppen eller behandlingsgruppen från "kontrollgruppen", som bör vara så lik behandlingsgruppen som möjligt i alla andra avseenden (förutom den behandling som mottas). Uppskattningar av de kontrafaktiska resultaten fås från kontrollgruppen, med särskild uppmärksamhet på skillnader i kännetecken – observerade och ej observerade – mellan de två grupperna. Man kan också jämföra ett antal olika behandlingar genom att exponera stödberättigande personer (enheter) för ett antal olika behandlingsvarianter (t.ex. andra ESF-finansierade behandlingar eller insatser som finansieras via andra källor), som bildar ett antal behandlingsgrupper, och sedan jämföra resultaten mellan grupperna och/eller resultaten för den ej behandlade kontrollgruppen.

*Fastställa
behandlingsgrupper*

När en kontrollgrupp inte är exponerad för någon behandling ställs följande utvärderingsfråga: "Vad blir effekten för dem som omfattas av insatsen jämfört med dem som inte får hjälp eller stöd"? Omvänt gäller samma sak: när resultaten av den berörda behandlingen jämförs med en annan behandling är utvärderingsfrågan "vad blir effekten av att omfattas av den berörda insatsen jämfört med att omfattas av ett annat väldefinierat alternativ"? Kontrafaktiska konsekvensbedömningar utformas ofta för att undersöka dessa grundläggande frågor. Valet av vilken fråga som ska behandlas avgörs av beslutsfattarnas prioriteringar och eventuella praktiska begränsningar i utformningen.

*Fastställa
kontrollgrupper*

Om två olika behandlingar jämförs bör det finnas en tydligt definierad kontrast mellan dem som är meningsfull ur ett politiskt perspektiv.

1.2. Varför är kontrafaktiska bedömningar viktiga?

Kontrafaktiska konsekvensbedömningar ger viktig information om nettoeffekter eller verkningar av insatser. De ger uppskattningar av verkningarnas omfattning, deras tecken (positivt eller negativt) och statistiska osäkerhetsmått. De bidrar till att bekräfta eller motsäga det förmodade orsakssamband mellan insatsen och resultaten som ligger till grund för insatsens utformning. Dessa mätta effekter kan användas för en rad olika ändamål: visa att det råder öppenhet och ansvarsutkrävande när det gäller användningen av skattebetalarnas pengar, fatta politiska beslut som grundas på tydliga belägg, lära från resultat från regioner i samma land eller mellan medlemsländerna.

Kontrafaktiska konsekvensbedömningar kan vara ett mycket nyttigt verktyg för dig som ansvarar för insatser och fortsatt finansiering av program, eftersom de visar att programmen ger valuta för pengarna och

*Öppenhet och
ansvarsutkrävande*

ger mätbara resultat för deltagarna och samhället i stort. Belägg från kontrafaktiska konsekvensbedömningar är särskilt intressanta för dig som ansvarar för resursfördelning. De förvaltande myndigheterna vill ju visa att deras program faktiskt fungerar. För att göra detta på ett övertygande sätt behöver de kontrafaktiska konsekvensbedömningar av hög kvalitet.

Dessa aspekter av kontrafaktiska konsekvensbedömningar ger viktig information för beslutsfattare som har till uppgift att fördela medel till olika insatser. Beslut om finansiering av möjliga insatser fattas med hänsyn till resursbegränsningar. I detta sammanhang behöver du som beslutsfattare faktaunderlag för programmets verkningar och kostnadseffektivitet så att tillgängliga medel kan användas på bästa sätt. När insatsernas relativa effektivitet bedöms framgår de nettoeffekter som uppskattas i kontrafaktiska konsekvensbedömningar tydligare om man undersöker insatsens kostnadseffektivitet eller utför en fullständig kostnads-nyttoanalys.

*Underlag för
resursfördelningsbeslut*

En jämförelse mellan nettoeffekterna av liknande insatser som har genomförts i olika områden kan också ge viktig kompletterande kunskap. Om väsentliga skillnader i verkningarna upptäcks i den kontrafaktiska konsekvensbedömningen är det mycket sannolikt att insatsernas kvalitet varierar betydligt eller att vissa insatser passar in i sitt socioekonomiska sammanhang bättre än andra. Dessa jämförelser, som brukar kallas "metautvärderingar" eller "metaanalyser", gör det möjligt att generalisera resultaten av den kontrafaktiska konsekvensbedömningen i högre grad och fungerar som viktiga tester av de politiska åtgärderna (utbildning och integrerade tjänster osv.)⁹. För att utföra dessa jämförelser enligt vetenskapliga kriterier krävs ett lämpligt antal insatser och systematiska kontrafaktiska konsekvensbedömningar. God samordning på nationell nivå eller EU-nivå är således en grundläggande förutsättning för mer frekventa jämförelser och för att de ska kunna användas som underlag för politiska beslut.

*Jämföra insatser i olika
regioner eller
medlemsländer*

1.3. Varför är kontrafaktiska konsekvensbedömningar tekniskt komplicerade?

Det finns ett antal strategier som kan anses vara "opålitliga" när det gäller att uppskatta insatsernas verkningar. Dessa strategier diskuteras här för att åskådliggöra kontrafaktiska konsekvensbedömningars inneboende komplexitet, men inga hänvisningar görs till den faktiska utvärderingsmetoden.

*Se längre än förenklade
metoder*

Ett första steg för dig som beslutsfattare kan vara att utvärdera hur ett utbildningsprogram exempelvis har påverkat arbetslösa genom att jämföra deltagarnas inkomst före utbildningen med inkomsten för samtliga arbetslösa personer som inte deltog i programmet. Därefter tilldelas utbildningsprogrammet den observerade skillnaden i inkomst mellan de personer som har deltagit och de som inte deltagit.

Detta är inte en giltig strategi för att fastställa hur utbildningen har påverkat inkomsten eftersom de personer som inte har deltagit kan skilja

⁹ Se t.ex. [Card D., Kluve J. och Weber A., 2017.](#)

sig i viktiga aspekter från deltagarna, och dessa skillnader i sin tur kan påverka resultaten. En sådan metod utesluter därför inte alternativa förklaringar till eventuella skillnader i den observerade inkomsten. Deltagarna kan till exempel ha större inneboende förmågor än de som inte deltagit. Arbetslösa personer med större förmågor erbjuder sig med andra ord frivilligt att delta i utbildningsprogrammet. Dessa förmågor påverkar således inte bara beslutet att delta, utan även resultaten – det är mer sannolikt att arbetslösa med större inneboende förmågor begär en större inkomst än personer med mindre förmågor. Detta innebär att eventuella observerade inkomstskillnader mellan behandlings- och kontrollgrupper både kan bero på utbildningsprogrammet och/eller på inneboende skillnader i förmåga, vilket innebär att insatsens verkan överskattas.

Om förmågor inte kan mätas och skillnaden i inneboende förmågor mellan de två grupperna inte kan beaktas i uppskattningen av verkningarna genom att jämföra resultaten mellan de två grupper anses utbildningsprogrammets uppskattade verkan påverkas av **urvalsfel**. För att lösa detta problem försöker utvärderarna inhämta så mycket information som möjligt om viktiga faktorer som påverkar beslutet att delta och de resultat som detta leder till. Dessa uppgifter används för att välja en giltig kontrollgrupp från personer som inte deltagit – dvs. en grupp som är så lik behandlingsgruppen som möjligt när det gäller dessa faktorer – och utföra statistiska analyser med hänsyn till dessa faktorer. I detta sammanhang utgår utvärderarna ofta från antagandet att urvalet av programmet avgörs av dessa observerbara faktorer. Detta "identifierande" antagande är oftast svårt att bekräfta och utvärderaren måste argumentera på ett övertygande sätt för att antagandet är rimligt i sammanhanget på grundval av kunskap om institutionella faktorer och beteendeteori.

Motverka urvalsfel

En annan "opålitlig" strategi kan vara att beslutsfattaren tittar på deltagarnas inkomst före och efter utbildningen och tillämpar skillnaden före/efter på utbildningsinsatsen. I grunden förutsätter en sådan strategi att den genomsnittliga inkomsten förblir oförändrad i avsaknad av insatsen.

Före och efter förändringen

Denna strategi är nästan aldrig meningsfull för att mäta hur utbildning påverkar inkomsten om inte antagandet om temporal stabilitet är tillämpligt. Detta beror på att deltagarnas inkomst oundvikligen kommer att förändras med tiden av orsaker som inte har något att göra med utbildningen. Deltagarnas inkomster minskar till exempel ofta innan utbildningen, delvis beroende på övergående faktorer. I många fall skulle denna tillbakagång ha inträffat oavsett utbildningsinsatsen¹⁰. Metoden att mäta inverkan av utbildning utifrån skillnad i inkomst omedelbart innan deltagaren påbörjar utbildningen och inkomsten efter utbildningen är således osäker och bortser från det faktum att inkomsten ofta skulle ha ökat i alla fall.

För att justera sådana metoder krävs ett mått av det kontrafaktiska scenariot, det vill säga att mått av hur inkomsten skulle ha förändrats för deltagarna utan utbildningsinsatsen. Ett sådant kontrafaktiskt resultat kan till exempel åstadkommas med hjälp av en noggrant matchad

¹⁰ Detta mönster kallas ofta "Ashenfelter Dip" efter den ekonom som först undersökte detta fenomen. Se [Ashenfelter O., 1978](#).

kontrollgrupp som inte omfattas av utbildningsinsatsen och vars inkomster observeras vid samma tidpunkter som deltagarnas inkomster. I detta sammanhang tillämpas ofta parallella trendantaganden, det vill säga att inkomsttrenden bland deltagarna och kontrollgruppen skulle ha varit densamma i avsaknad av insatsen.

Begränsningarna hos dessa "opålitliga" strategier innebär att man bör söka efter mer övertygande utvärderingsmetoder. Som nämns ovan är mer övertygande metoder dock tekniskt svårare att genomföra. Nästa avsnitt i detta kapitel innehåller en kort översikt av några av de strategier för kontrafaktiska konsekvensbedömningar som sannolikt är mest relevanta för ESF.

1.4. Översikt av olika metoder för att utforma och genomföra kontrafaktiska konsekvensbedömningar

Vid mätningar av hur deltagare påverkas av en insats uppskattas kontrafaktiska resultat vanligen med hjälp av uppgifter från grupper av personer som inte har deltagit i den insats som ska utvärderas, men i övrigt liknar de personer som deltar i insatsen. I tabell 1 i slutet av det här kapitlet ges en kort översikt av metoder, några av deras fördelar och begränsningar och de viktigaste typer av uppgifter som krävs.

Den viktigaste åtskillnad som görs i kontrafaktiska konsekvensbedömningar är mellan experimentella och kvasiexperimentella tillvägagångssätt. Det experimentella tillvägagångssättet kallas ofta "randomiserad kontrollerad studie" eller *RCT (randomised control trial)*, och ibland även för "sociala experiment".

Den experimentella metoden anses utgöra den gyllene normen bland metoderna för kontrafaktiska konsekvensbedömningar för att utvärdera effekter av insatser som kan testas och manipuleras under förhållandevis korta tidsspann, och är i de flesta fall den idealiska lösningen. Uppskattningar av kontrafaktiska resultat ska vara opartiska, dvs. neutrala. Den experimentella metoden anses ge sådana neutrala uppskattningar. När man ska välja strategi för den kontrafaktiska konsekvensbedömningen är det därför bra att börja med att utforma den experimentella metoden. Syftet med kvasiexperimentella metoder är snarare att imitera det experimentella idealet.

Randomiserad utformning – den gyllene standarden

I diskussionen om hur den kontrafaktiska konsekvensbedömningen bör utformas beskrivs de viktigaste kännetecknen hos varje metod så enkelt som möjligt för att klargöra de underliggande principerna. Metoderna är betydligt mer komplicerade att tillämpa i verkligheten och de problem som kan uppstå, t.ex. att personer som har godkänts för deltagande i insatsen till slut inte gör det, kan förvärra problemen avsevärt.

Det finns många metoder som syftar till att imitera randomisering. Dessa metoder anses vara **kvasiexperimentella**. Vi har inte utrymme för att ta upp alla dessa metoder i denna vägledning eller redogöra för de tekniska aspekterna av varje metod. Generellt sett kan dock följande kvasiexperimentella metoder bli aktuella inom ramen för ESF: 1. Benägenhetspoäng. 2. Skillnaden-mellan-skillnader. 3. Regressionsdiskontinuitet. 4. Instrumentvariabler. En översikt av de viktigaste metoderna och deras fördelar ges i Tabell 5 i kapitel 3. Metoderna presenteras på ett mycket förenklat sätt för att belysa de grundläggande principerna för varje metod. Om du vill läsa mer om kvasiexperimentella metoder hittar du litteratur i bilaga 1.

Icke-randomiserade eller kvasiexperimentella metoder

1.5. Hur kontrafaktiska konsekvensbedömningar kan integreras i en bredare utvärderingsram

Kontrafaktiska konsekvensbedömningar behandlar vissa typer av frågor om insatsers kausala verkan. Dessa metoder har sina begränsningar när det gäller att undersöka andra frågor kring insatsen. Åtskillnad bör göras mellan utvärderingsfrågor som rör **kausala förklaringar** och utvärderingsfrågor som avser **kausala beskrivningar**. Kontrafaktiska konsekvensbedömningar syftar till att **beskriva** konsekvenserna av en insats. Sådana metoder är mindre lämpliga för att **förklara** de mekanismer och sammanhang som ger upphov till orsakssambanden. Denna åtskillnad är viktig eftersom den bidrar till att klargöra kontrafaktiska konsekvensbedömningars särskiljande funktion¹¹.

Kausala förklaringar och kausala beskrivningar

En väl utformad kontrafaktisk konsekvensbedömning informerar dig som beslutsfattare om huruvida en insats har lett till förändringar i de resultat som den utformades för att påverka. Den ger belägg för hur stora verkningarna eller resultaten har varit och om insatsen har haft en positiv eller negativ inverkan, men innehåller även ett mått av osäkerhet. Kontrafaktiska konsekvensbedömningar är däremot sämre på att redovisa varför och hur de effekter som har mätts i bedömningen har uppstått. Omvänt är det ofta svårt att utifrån en kontrafaktisk konsekvensbedömning fastställa varför en insats inte har haft någon påverkan om så visar sig vara fallet.

Vad kontrafaktiska konsekvensbedömningar kan berätta och inte berätta

Aktörer inom de flesta beslutsfattande organ brukar ha olika intressen och perspektiv när de utformar deskriptiva och kausala förklarande frågor. Programförvaltare och yrkesverksamma i stort brukar fokusera på kausala förklarande frågor. Personer som arbetar med resursfördelning och överordnade beslutsfattare med budgetansvar brukar däremot vara mer intresserade av kausala beskrivande frågor. I praktiken kan skillnaden mellan kausala förklaringar och kausala beskrivningar vara otydlig. Under vissa omständigheter kan kontrafaktiska konsekvensbedömningar förklara varför vissa effekter har påträffats, till exempel genom att undersöka hur insatsen har påverkat viktiga delgrupper. Det är dock viktigt att noggrant överväga de frågor som ska besvaras i utvärderingen av insatsen och välja lämpliga metoder för att besvara dem. Om den primära frågan är om en insats fungerar eller ej är en kontrafaktisk konsekvensbedömning ofta ett lämpligt alternativ. Om

¹¹ Se Shadish W.R., Cook T.D. och Campbell D.T., 2002, och [Stern E. m.fl., 2012](#).

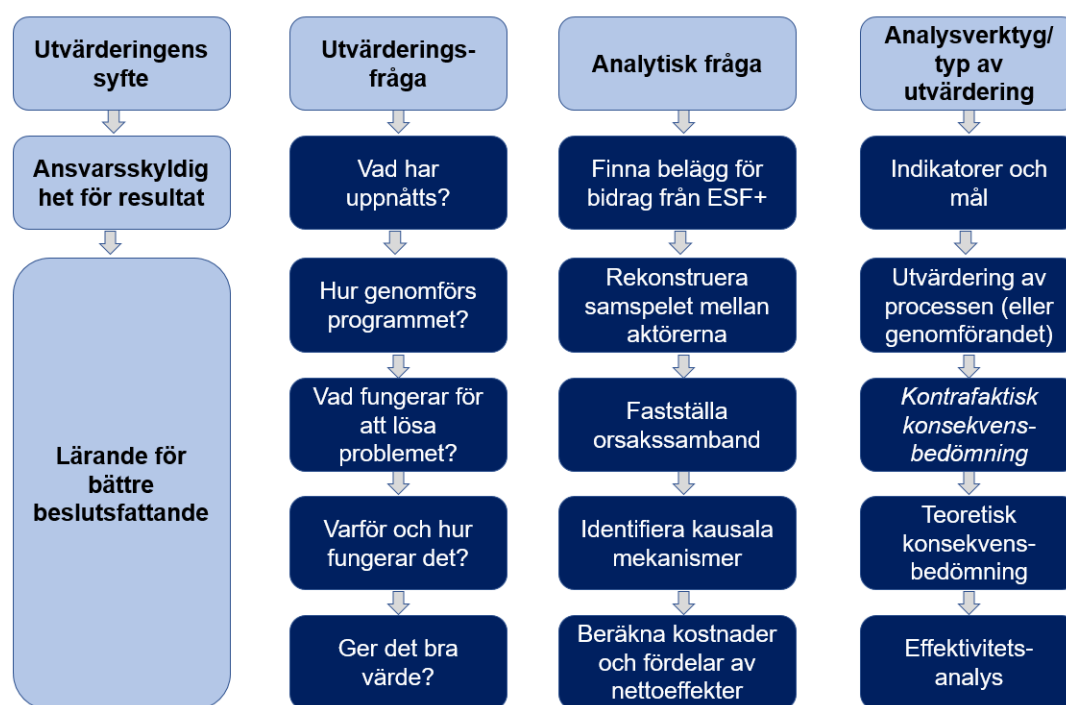
den primära frågan är hur en insats fungerar är däremot teoretiska metoder och processutvärderingsmetoder lämpligare.

Dessa olika nivåer och syften med frågorna sammanfattas i diagrammet nedan.

Slutsatsen av denna diskussion är att kontrafaktiska konsekvensbedömningar måste utformas inom ramen för utvärderingsplanen. Utvärderingsplanen ska bestå av olika typer av utvärderingar som är avsedda att besvara olika frågor för olika politiska aktörer. Om en kontrafaktisk konsekvensbedömning ingår i utvärderingsplanen krävs i stort sett alltid en processutvärdering.

Det finns många olika utvärderingsmetoder, som var och en tjänar sitt eget syfte. Den kritiska frågan är hur dessa metoder kan kombineras på användbara sätt till stöd för erfarenhetsbaserat lärande. Det svåra är att kombinera olika typer av utvärderingar – med olika syften inom ramen för programplaneringen – på ett lämpligt sätt. Som vi har diskuterat, och vilket även framgår av nästa diagram, kompletterar kontrafaktiska konsekvensbedömningar, processutvärderingar och teoretiska metoder varandra.

Figur 1 Olika ändamål med och olika typer av utvärderingar



Källa: anpassad från [Martini A., 2009.](#)

En robust utvärderingsstrategi bör bestå av följande delar:

- Teoretisk konsekvensbedömning.
- Utvärdering av processen (eller genomförandet).
- Kontrafaktisk konsekvensbedömning.
- Effektivitetsanalys.

Kombinera olika typer av utvärderingar

Denna vägledning inriktas på kontrafaktiska konsekvensbedömningar. Inom ramen för kontrafaktiska konsekvensbedömningar är syftet med teoretiska metoder att skapa förståelse för varför en insats har utformats på ett visst sätt.

Teoretiska utvärderingar används ibland för att beskriva hur insatsen är avsedd att fungera och för att testa om de förändringar i resultaten som insatsen är avsedd att ge i teorin kan observeras i verkligheten. Teoretiska metoder kan alltså användas för att bedöma effekterna av en insats genom att besvara "hur" effekten har uppstått när det inte är möjligt att genomföra en kontrafaktisk konsekvensbedömning. I nästa diagram visas ett stiliserat exempel på förändringsteori för en insats som syftade till att förbättra en arbetsförmedlings tjänster. I en teoretisk bedömning undersöker man om beläggen bekräftar sekvensen av kausala steg enligt förändringsteori, från insatser till inverkan, och vilka omständigheter och sociala mekanismer som möjliggjorde detta. Teoretiska metoder för att fastställa inverkan omfattas inte av denna vägledning och redovisas därför inte i detalj.

En teoretisk bedömning är en förändringsteori

Inom ramen för kontrafaktiska konsekvensbedömningar undersöker teoretiska utvärderingar hur en insats planeras och utformas och hur den är avsedd att genomföras. Metoden går i huvudsak ut på att skapa en gemensam förståelse av insatsens underliggande "förändringsteori" bland de aktörer som genomför insatsen, vilket visas i förenklad form i nästa diagram. Alla insatser har en programlogik som kopplar input och verksamheter till output på medellång sikt och därefter till resultat på längre sikt. Förändringsteorier är därför ett viktigt underlag även för kontrafaktiska konsekvensbedömningar: med förändringskategorier är det enklare att rekonstruera genomförandeprocessen och se hur den eventuellt har påverkat output och utfall, och det är även enklare att identifiera de mest representativa utfallsvariablerna som ska kontrolleras i den kontrafaktiska analysen.

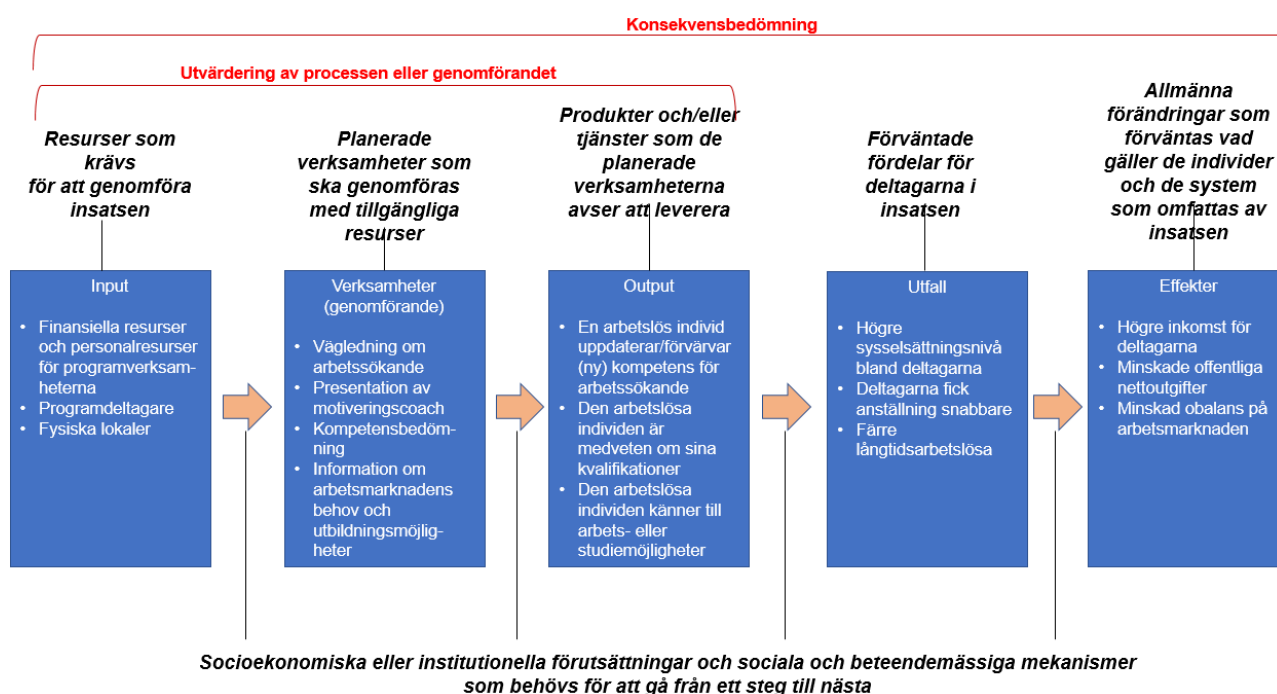
Teoretiska bedömningar kan kopplas till kontrafaktiska konsekvensbedömningar på en rad olika användbara sätt. En tydligt utformad förändringsteori (eller interventionslogik) kan tjäna som underlag för utformningen av en kontrafaktisk konsekvensbedömning. En väl definierad förändringsteori kan bland annat ge dig som tar fram konsekvensbedömningar följande information.

som bidrar till kontrafaktiska konsekvensbedömningar

- Vilka resultat är viktiga och behöver mätas?
- Vilket tecken och storlek kan insatsens effekter sannolikt ha?
- Vilken är den avsedda målgruppen och hur kan en kontrollgrupp väljas?
- Hur långt kan det dröja innan programmets effekter och resultat blir märkbara?
- Vilka data kan behövas för att mäta deltagandet i insatsen?

Hur rimlig är kontrollgruppen som ett mått på det kontrafaktiska scenariot?

Figur 2 Illustration av logikmodellen eller ”förändringsteorin”



Källa: Anpassad från [W.K. Kellogg Foundation, 2004](#) och [Bredgaard T., 2015](#).

Förändringsteorier kan även bidra till att identifiera eventuella oväntade effekter som bör beaktas i utformningen av en kontrafaktisk konsekvensbedömning. En tydligt utformad förändringsteori kan också i viss utsträckning hjälpa utvärderaren att tolka resultaten av en kontrafaktisk studie. När det gäller tolkning av resultat kan en **processutvärdering** också vara mycket upplysande.

Processutvärdering

Processutvärderingar har två syften när det gäller kontrafaktiska konsekvensbedömningar. Det första är att bedöma "tillförlitligheten" och det andra är att bedöma skillnaden mellan de erfarenheter som deltagarna i behandlingsgruppen och kontrollgruppen har haft.

Bedömningen av tillförlitlighet visar i hur stor utsträckning en insats har genomförts som avsett, dvs. på det sätt den utformats. I en processutvärdering undersöks vilka tjänster som egentligen tillhandahölls till deltagarna i insatsen. Motsvarar tjänsterna det som avses i insatsens teori? Om tjänsterna levererats på olika sätt på olika platser, vad beror dessa variationer på i så fall? De flesta insatserna har både en förvaltningslogik och en effektlogik.

Tillförlitlighetsbedömning

- Förvaltningslogiken handlar om hur genomförandeorganen förväntas svara på programmets regler och incitament.
- Effektlogiken handlar å sin sida om hur de personer som insatsen riktar sig mot förväntas reagera mot bakgrund av de resultat som faktiskt levereras.

Processanalysens tillförlitlighetsaspekt ger information om vad som faktiskt har uppnåtts med hjälp av insatsen, det vill säga vilka faktorer som faktiskt bidrar till de observerade effekterna. Den ger även viktig återkoppling för projektledningen.

Bedömningen av skillnader är särskilt viktig i kontrafaktiska bedömningar. Fokus ligger ofta på insatsers effekter, precis som i den här vägledningen. Innan man tittar på hur resultaten påverkas, vilket i sin tur visar hur input påverkar resultaten, måste man dock beräkna skillnaden mellan behandlings- och kontrollgrupperna när det gäller de möjligheter som en insats faktiskt ger. I princip kan alla kontrafaktiska konsekvensbedömningar "kastas om", vilket även gäller den behandlingsgrupp som användes som kontroll för att bedöma resultaten för personer i den grupp som före omvändningen kallades kontrollgruppen. Förutsättningen är att man vet lika mycket om vilka faktorer som styr erfarenheterna som man vet om behandlingen, eftersom orsakssambanden när det gäller skillnader i resultat i en kontrafaktisk konsekvensbedömning styrs av skillnaden mellan behandling och kontroll av inputfaktorerna.

*Skillnad mellan
behandlingsgrupper och
kontrollgrupper*

Om vi går tillbaka till exemplet med utbildningsinsatsen kan man tänka sig två olika ingångsscenarier. I det ena scenariot tillhandahålls utbildningsprogrammet i allmänt sammanhang och inga utbildningar av den aktuella typen finns tillgängliga. Kontrollindividerna klarar sig helt enkelt utan utbildning. En annan möjlighet är att det finns vissa substitut. Utbildning kan till exempel finnas tillgänglig från företag som specialiserar sig på yrkesutbildning. I detta fall måste processanalysen i möjligaste mån omfatta en bedömning av skillnaden mellan utbildningsviljan i behandlings- och kontrollgruppen, eftersom det inte bara kan förutsättas att alla dimensioner av behandlingen finns inom räckhåll för individerna i kontrollgruppen.

Processutvärderingar kan göras helt oberoende av andra utvärderingar. De är dock viktiga för både utvärderingen av projektledningen och den kontrafaktiska konsekvensbedömningen, vilket i sin tur innebär att det är viktigt att process- och konsekvensbedömningarna planeras tillsammans.

*Processutvärderingar är
en förutsättning för
kontrafaktiska
konsekvensbedömningar*

En väl utförd processanalys kan skapa tillförlitlighet och processutvärderingarna redovisar insatsens orsakssamband. Utan en väl utformad processutvärdering är det ofta svårt att fullständigt tolka resultaten av en kontrafaktisk konsekvensbedömning eller att beräkna kostnaden för kostnads-nyttanalyser om effektuppskattningar finns tillgängliga.

Som nämns ovan kan processutvärderingar även bidra till att tolka resultaten av konsekvensbedömningar genom att visa i vilket sammanhang insatsen genomfördes. Det är viktigt att man förstår sammanhanget eftersom det är sammanhanget som avgör hur effektiv insatsen blir. Det ger också en indikation på om insatsen kan ge liknande effekter om den genomförs på annat håll, inom olika geografiska områden eller vid olika tidpunkter. Detta är särskilt viktigt om man överväger att överföra politiska strategier och för att lyfta fram god praxis i gränsöverskridande lärande och utbyten. Processanalyser bidrar till att skapa förtroende för vad som benämns utvärderingsresultatens **externa giltighet**.

Effektivitetsanalyser omfattar i de flesta fall antingen en kostnadseffektivitetsanalys eller en kostnads-nyttanalyser.

Fastställa kostnadseffektivitetskvoter

Kostnadseffektivitetsanalyser handlar om att jämföra kostnaderna för insatsen med dess effekter enligt den kontrafaktiska konsekvensbedömningen. Enkelt uttryckt får man fram en kostnadseffektivitetskvot genom att dividera insatsens effekt – antingen uttryckt i måtenheter eller standardiserade enheter – med nettokostnaden för genomförandet av insatsen per behandlad enhet.

En kostnadseffektivitetskvot för ett utbildningsprogram som syftar till att hjälpa arbetslösa att hitta arbete kan visa hur mycket det kostar att hjälpa varje arbetslös deltagare att få jobb.

Kostnads-nyttanalyser som ett verktyg för att jämföra fördelen med nettokostnaden

Kostnadseffektivitetskvoten är ett viktigt mått för dig som ansvarar för att fördela medel mellan olika program. Kvoter från olika insatser utgör således ett underlag för dig som beslutsfattare som hjälper dig att bedöma vilka insatser som ger mer valuta för pengarna.

I stället för att uttrycka programmets effekter genom måtenheter eller standardiserade enheter syftar en **kostnads-nyttanalyser** till att kostnadsberäkna de uppskattade effekterna enligt den kontrafaktiska konsekvensbedömningen och jämföra dem med insatsens nettokostnader. Syftet med kostnads-nyttanalyser är således att avgöra om penningvärdet för ett program överskrider dess nettokostnader. I en kostnads-nyttanalyser av ett typiskt ESF-utbildningsprogram jämför man vanligen insatsens fördelar för deltagarna, staten och samhället i stort med insatsens nettokostnader. Programmets fördelar för deltagarna (vanligen förbättrad anställbarhet och ökad inkomst) fås fram med hjälp av en kontrafaktisk konsekvensbedömning. Därefter drar man av värdet på skatter och andra anställningsrelaterade kostnader som deltagarna betalar för att få fram nettofördelen. Ur statens perspektiv ger insatsen fördelar i form av ytterligare skatteintäkter och minskade socialbidragskostnader, samtidigt som staten bär den största delen av kostnaderna för insatsen. Kostnaden för samhället i stort fås fram genom att fördelarna för deltagarna och staten läggs ihop och summan av deltagarnas och statens kostnader dras av från detta belopp.

Det är fortfarande inte särskilt vanligt att kostnadseffektivitetsanalyser och kostnads-nyttanalyser används i samband med ESF-utvärderingar. Sådana analyser är dock mycket användbara som underlag för beslut om huruvida en insats bör finansieras igen i framtiden eller för att identifiera den mest effektiva insatsen bland en uppsättning liknande insatser (se exempel i följande ruta).

Effektuppskattningar från kontrafaktiska konsekvensbedömningar är en viktig faktor i både kostnadseffektivitets- och kostnads-nyttanalyser. I den förstnämnda analysen är de ett mått på effektiviteten, medan de i den sistnämnda analysen är en viktig källa för att uppskatta fördelarna mätt i penningvärde. Både kostnadseffektivitetsstudier och kostnads-nyttanalyser kräver korrekta kostnadsuppgifter varifrån nettokostnaderna kan härledas. Insamlingen av dessa uppgifter kallas vanligen kostnadsundersökning. I vissa komplexa utvärderingar där flera metoder kombineras integreras kostnadsundersökningar vanligen

i processutvärderingen, och forskningsmetoderna anpassas för att samla in viktiga kostnadsuppgifter.

Ruta 1 Exempel på en kontrafaktisk konsekvensbedömning för jämförelse av fördelar och kostnader

Ett exempel på användningen av en kontrafaktisk metod för att uppskatta fördelar och kostnader ges i [Bazzoli M. m.fl., 2018](#). Studien inriktas på yrkesutbildningsprogram som genomfördes i den italienska autonoma provinsen Trento 2010–2011, som omfattade över 300 timmar utbildningsaktiviteter. Två grupper av kurser utvärderades: dels kurser som finansierades med provinsiella medel (PVR) och dels kurser som finansierades via ESF, med 954 respektive 205 deltagare.

Kostnads-nyttoanalysen genomfördes i följande huvudsteg: 1) Författarna utvärderade hur utbildningskurserna påverkade sannolikheten för att deltagarna skulle hitta ett jobb under tre år efter det att kursen inletts, med tillämpning av benägenhetspoäng¹². 2) Kursernas inverkan på bruttointkomsten fram till utgången av 2013 uppskattades. 3) Författarna uppskattade de ytterligare finanspolitiska vinsterna (härledda från effekterna på inkomsterna) och besparingarna i statliga medel till följd av det minskade antalet mottagare av arbetslöshetsförmåner. 4) Kursernas fördelar och kostnader jämfördes.

Flera administrativa datauppsättningar användes: a) övervakning av data rörande deltagarna och deras förutsättningar, b) data från PES-registret (Centri per l'impiego, dvs. arbetsförmedlingar) över arbetslösa för att fastställa kontrollgruppen, c) uppgifter från COB-databasen, dvs. arkivet över arbetsgivares obligatoriska anmälan av anställningsavtal som skickas till den offentliga arbetsförmedlingen, för att fastställa anställningsstatus för behandlings- och kontrollgrupperna före och efter deltagandet i utbildningskursen, d) data från skatteinkomstregister, för att beräkna de berörda personernas inkomst, och e) data från det italienska socialförsäkringsinstitutet (INPS), för information om de arbetslöshetsförmåner som de berörda personerna mottagit.

Efter 36 månader var sannolikheten för att få jobb för personerna i behandlingsgruppen i PVR-kurserna omkring 5 procentenheter högre än i kontrollgruppen, medan ESF-kurserna hade en mycket större effekt, cirka 28 procentenheter. Under de tre undersökta åren efter insatsen tjänade personer som hade deltagit i PVR-kurserna i genomsnitt totalt 2 250 euro p.a. mer än personerna i kontrollgruppen, medan de som deltagit i ESF-kurserna tjänade 4 106 euro p.a. mer än kontrollgruppen. På grundval av dessa uppgifter kunde författarna även uppskatta fördelarna för den offentliga förvaltningen i form av ökade skatteintäkter och minskade utgifter för sociala förmåner. De personer som deltagit i PVR-kursen betalade 126 euro mer i inkomstskatt än kontrollgruppen för varje år som omfattades av undersökningen, medan uppskattningen för personerna i ESF-kursgruppen uppgick till 318 euro p.a. När det gäller den offentliga förvaltningens minskade utgifter för arbetslöshetsersättningar hade kurserna en försumbar effekt, förmodligen för att penningvärdet för arbetslöshetsersättningar varierade beroende på hur länge personerna varit anställda innan de blev arbetslösa, och att många av deltagarna var ungdomar med begränsad yrkeserfarenhet.

Kostnaderna för insatserna uppgick till 4 800 euro per deltagare i PVR-kurserna och 14 500 euro per deltagare i ESF-kurserna. Författarna jämförde även kostnader och fördelar på individuell nivå för båda typerna av kurser. De fann att kostnaderna i genomsnitt var större än fördelarna för perioden 2010–2013¹³.

¹² Författarna tillämpade närmare bestämt blockering med regressionsjusteringar.

¹³ En annan liknande metod beskrivs i [Lammers M. och Kok L. \(2021\)](#).

Kapitel 2. Praktiska överväganden vid utformningen av kontrafaktiska konsekvensbedömningar

I detta kapitel tittar vi närmare på vilka praktiska frågor som bör övervägas i förberedelserna för utvärderingen. Detta kapitel bör utgöra underlag för planeringen av utvärderingen, beslutet om vilka insatser som bör ingå i den kontrafaktiska konsekvensbedömningen och för att fastställa de viktigaste frågor som bör ingå i bedömningen.

Förbereda en kontrafaktisk konsekvensbedömning

Utgångspunkten är en situation där du som programförvaltare inom en förvaltningsmyndighet (eller som förvaltare vid ett förmedlande organ) med ansvar för att genomföra en ESF-insats överväger vilka insatser som ska utvärderas och hur en kontrafaktisk konsekvensbedömning kan integreras på bästa sätt i utvärderingen. Vi förutsätter också att ni vid den förvaltande myndigheten inte kommer att utföra utvärderingarna själva, utan i stället lägga ut utvärderingen eller anlita externa experter. Även om den kontrafaktiska konsekvensbedömningen utförs av en extern uppdragstagare måste den förvaltande myndigheten (eller det förmedlande organet) ändå planera och förbereda för en konsekvensbedömning innan uppdraget läggs ut.

Utvärderingsstrategin, även de olika typer av utvärderingar som beskrivs i det föregående kapitlet, måste fastställas i utvärderingsplanen.

Utvärderingsplaner är obligatoriska för samtliga program och ska godkännas av övervakningskommittén senast ett år efter godkännandet av programmet¹⁴. Planerna ska utarbetas i början av programperioden och ska omfatta arrangemang för utvärderingsprocessen (hur utvärderingarna ska ledas och kopplingen mellan utvärdering och övervakning), faktiska utvärderingsverksamheter (t.ex. en preliminär förteckning över de utvärderingar som ska genomföras, omfattningen av varje utvärdering, huvudsakliga frågor, nödvändiga data, potentiella användningsområden, preliminär tidsplan, ledningsstruktur), tidpunkten för utvärderingarna, total budget och kapacitetsbehov.

Utvärderingsplaner brukar vara allmänt hållna, medan det krävs mer detaljerade förberedelser för att planera en kontrafaktisk konsekvensbedömning. Den kontrafaktiska konsekvensbedömningen bör planeras samtidigt som utvärderingsplanen, men vissa detaljer kan kompletteras i ett senare skede. Den förvaltande myndigheten eller det förmedlande organet måste dock vara medvetna om att det inte är någon lätt uppgift att fastställa kopplingar till berörda aktörer och andra nödvändiga arrangemang för att samla in data om insatsen, utan att detta bör planeras i god tid.

¹⁴ Utvärderingsplaner är ett krav enligt artikel 44.5 och 44.6 i förordning (EU) 2021/1060. Flera planerade utvärderingar av fler än ett program kan ingå i planen, men samtliga program måste omfattas av en utvärderingsplan.

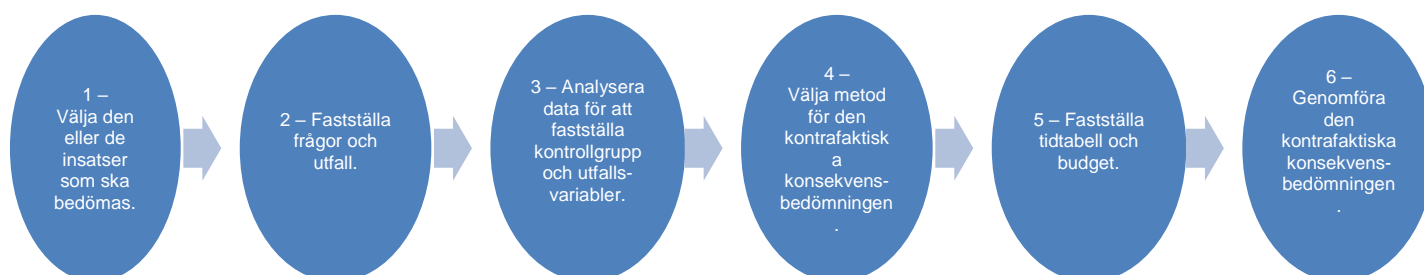
Denna vägledning inriktas på olika sätt att utforma ett utvärderingsschema för specifika insatser som är lämpade för kontrafaktiska konsekvensbedömningar. Schemat bör ingå i utvärderingsplanen eller alternativt ingå som ett operativt steg inom ramen för en uppföljning av en utvärderingsplan.

Utforma ett utvärderingsschema för specifika insatser

Kontrafaktiska konsekvensbedömningar passar inte alla typer av ESF-finansierade insatser. Som beslutsfattare måste du välja vilka insatser du vill koncentrera dig på. Det behövs en process för att välja de insatser som ska konsekvensbedömas. I denna vägledning lyfter vi fram några aspekter som de förvaltande myndigheterna behöver ta hänsyn till i valet av lämpliga insatser. Det centrala syftet med denna vägledning är att hjälpa dig som har ansvar för utförandet av kontrafaktiska konsekvensbedömningar genom att uppmärksamma dig på problem som kan uppstå under arbetets gång och ta fram utvärderingsplaner för de kontrafaktiska konsekvensbedömningar som övervägs.

Vi förutsätter här att de förvaltande myndigheterna när de har valt ut de insatser som ska omfattas av kontrafaktiska konsekvensbedömningar utarbetar ett utvärderingsschema för varje vald insats. Begreppet "schema" används för att skilja denna aktivitet från de formella "utvärderingsplaner" som krävs enligt förordningen om gemensamma bestämmelser (förordning (EU) 2021/1060) för programperioden 2021–2027. Begreppet "schema" avser den uppsättning standardiserade steg som krävs för att fastställa och genomföra en kontrafaktisk konsekvensbedömning. Detta ska ske i förväg, enligt följande bild.

Bild 3 Ordningsföljd för de viktigaste stegen i en kontrafaktisk konsekvensbedömning



Dessa scheman kommer att utgöra underlaget för utförandet av kontrafaktiska konsekvensbedömningar och lägger grunden för en noggrann studie som ger användbara resultat. I återstoden av kapitlet tittar vi på de frågor som behöver lösas i planeringen av utvärderingarna. Utvärderingsscheman måste skräddarsys efter de särskilda omständigheterna för varje insats. Det är omöjligt att veta hur omständigheterna kommer att se ut. Här diskuterar vi därför frågor som a) bör hanteras i schemat, eller b) frågor som kan tjäna som inspiration för att lösa de problem som schemat är avsett att hantera

Frågor som behöver besvaras i planeringen av utvärderingen

Efter att ha undersökt vilka insatser som är lämpade för kontrafaktiska konsekvensbedömningar och om det är möjligt att genomföra sådana bedömningar med de uppgifter som finns tillgängliga går vi nu över till vissa centrala frågor som är viktiga att tänka på när man tar fram ett utvärderingsschema. Utvärderingsschemat måste vara klart innan den kontrafaktiska konsekvensbedömningen (eller en mer omfattande utvärderingsstudie) beställs för att ta fram specifikationerna och utse en uppdragstagare. Det huvudsakliga innehållet i ett utvärderingsschema anges i tabellen nedan.

Tabell 1 Rekommenderat innehåll i ett utvärderingsschema

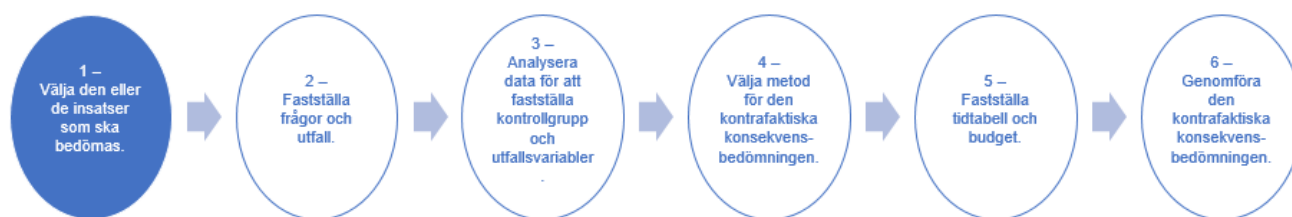
Huvudsteg i förberedandet och genomförandet av en kontrafaktisk konsekvensbedömning	Innehåll
1. Välja den eller de insatser som ska bedömas	<ul style="list-style-type: none"> – De ESF-insatser, eller typer av insatser (typen av insatser ska om möjligt fastställas på förhand), som ska utvärderas med hjälp av en kontrafaktisk konsekvensbedömning. – En sammanfattande beskrivning av huvuddragen i de utvalda insatserna (mål, huvudsakliga kriterier för stödberättigande, målgrupp, typer av stöd, genomförandemodell, preliminära start- och slutdatum, preliminär budget).
2. Fastställa utvärderingsfrågor och utfallsvariabler	<ul style="list-style-type: none"> – Utvärderingsfrågor. – Hur insatserna förväntas fungera ("förändringsteorin") och utfallsvariabler (sysselsättningsstatus, förändrad inkomst, fattigdomsstatus, genomsnittliga betyg i prov osv.).
3. Analysera data för att fastställa kontrollgruppen och mäta utfallsvariablerna	<ul style="list-style-type: none"> – De administrativa uppgifter som ska användas (t.ex. arbetslöshetsregister, skatteregister, försäkringsadministrativa uppgifter, skolors register eller databaser över elever) eller andra data (t.ex. undersökningar, stordata). - De variabler för databasen eller undersökningen som ska användas i den kontrafaktiska konsekvensbedömningen. - Tidsserier för de variabler som behövs. - Centrala regler för och aspekter av datatillgång (direkt åtkomst för förvaltande myndigheter, nödvändiga överenskommelser med andra förvaltningar, integritetsregler och begränsningar).
4. Välja metod för den kontrafaktiska konsekvensbedömningen	<ul style="list-style-type: none"> – Vilken metod som ska användas för den kontrafaktiska konsekvensbedömningen (detta kan beslutas senare, men användning av randomiserade kontrollerade studier eller andra metoder bör anges i förväg så att utvärderingsprocessen blir enhetlig).
5. Fastställa tidtabell och budget	<ul style="list-style-type: none"> – Tidtabell för den kontrafaktiska konsekvensbedömningen, viktiga delmål: a) Bestämna vilken metod som ska användas, experimentell eller kvasiexperimentell, b) detaljerade utvärderingsfrågor, c) utarbeta uppdragsbeskrivningen, d) utse utvärderare, e) förbereda data, f) dataanalys, g) halvtidsrapporter och slutrapporter, h) validera resultaten, i) sprida resultat och erfarenheter. – Tillgänglig budget för den kontrafaktiska konsekvensbedömningen.
6. Genomföra utvärderingen	<ul style="list-style-type: none"> – Val av utvärderare. – Övervakning av genomförandet av den kontrafaktiska konsekvensbedömningen. – Rapportering om den kontrafaktiska konsekvensbedömningen. – Spridning av resultaten (viktigaste aktörer som bör delta, huvudsakliga verktyg).

Ruta 2 Utvärdering av kontrafaktiska konsekvensbedömningar inom en bredare ram

Många kontrafaktiska konsekvensbedömningar av ESF-finansierade insatser som görs i medlemsländerna är integrerade i bredare utvärderingsramar.

- I det tyska operativa programmet Bund ESF 2014–2020 ingår den kontrafaktiska utvärderingen av programmet för integrering av långtidsarbetslösa på arbetsmarknaden som finansieras inom det integrerade projektet 9.i i en bredare utvärderingsstrategi med delårsrapporter där även frågor kring genomförandet av insatserna undersöks. Fyra rapporter togs fram under åren 2017–2021¹⁵. Detsamma gäller utvärderingen av ESF-åtgärder till stöd för integrering av långtidsarbetslösa i Baden-Württemberg, där den kontrafaktiska analysen utgjordes av en mer kvalitativ analys som inriktades på deltagarnas bedömning av insatserna¹⁶.
- I Piedmont (Italien) var en kontrafaktisk konsekvensbedömning av effekterna på sysselsättningen av ”sysselsättningscheckar” för utsatta personer som finansierades inom de integrerade projekten 8.i och 9.i ett steg i en mer omfattande pågående utvärdering. Genomförandefrågor och deltagarnas subjektiva uppfattningar behandlades i två inledande rapporter (april 2018 och februari 2019), medan två ytterligare rapporter i juli 2019 och i slutet på 2020 inriktades på sysselsättningseffekter med tillämpning av en kontrafaktisk metod¹⁷.
- Den kontrafaktiska konsekvensbedömning av insatsernas effekter för långtidsarbetslösa 2020 som genomfördes i Marche (Italien) har en tematisk inriktning. Den följde efter en mer allmän konsekvensanalys 2019 av ESF-insatser för arbetslösa (rapport om sysselsättningsresultat)¹⁸.
- I Polen beställdes ett antal utvärderingar (8 rapporter¹⁹) inom ramen för det operativa programmet Kunskap – utbildning – tillväxt 2014–2020 mellan årsslutet 2015 och maj 2020. Syftet var att analysera stödet till ungdomar från ESF och sysselsättningsinitiativet för unga ur olika perspektiv och med användning av flera analysmetoder (t.ex. kvalitativa analyser via undersökningar och intervjuer, makromodeller, fältverksamhet och en kontrafaktisk metod).

2.1. Välja insatser som ska konsekvensbedömas



Urvalet av insatser som ska konsekvensbedömas sker i tre huvudsakliga *Kriterier för att välja insatser* steg:

1. Första steget är att fastställa de strategiska prioriteringarna.

¹⁵ Se [Boockmann B. m.fl., 2017](#) – [Boockmann B. m.fl., 2018](#) – [Boockmann B. m.fl., 2019](#) och [Boockmann B. m.fl., 2021](#). Kontrafaktiska konsekvensbedömningar genomfördes åren 2019 och 2021.

¹⁶ Se [Hunger K. och Sattler K., 2017](#) och [Scheller F. och Seidel K., 2020](#). Kontrafaktiska konsekvensbedömningar tillämpades 2020.

¹⁷ Se [Pomatto G., 2017](#) – [Pomatto G., 2019](#) – [Poy S., 2019](#); [Poy S., 2020](#). Kontrafaktiska konsekvensbedömningar genomfördes åren 2019 och 2020.

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

¹⁹ Se [Instytut Badań Strukturalnych – Imapp – IQS, 2015](#) – [Baran J. m.fl., 2016](#) – [Baran J. m.fl., 2017](#) – [Baran J. m.fl., 2018](#) – [Baran J. m.fl., 2018a](#) – [Palczyńska M. m.fl., 2019](#) – [Kalinowski H., 2020](#) – [Kalinowski H. m.fl., 2020](#). Kontrafaktiska konsekvensbedömningar tillämpades 2017 och 2020.

2. När de strategiska prioriteringarna är klara bedöms insatserna var för sig för att se om de uppfyller de grundläggande kraven för en kontrafaktisk metod, i hur stor utsträckning de är innovativa och/eller bidrar på ett värdefullt sätt till kunskapsbasen.
3. Man måste också veta om det finns tillgång till eller möjligheter att få tillgång till de data som krävs för att genomföra en kontrafaktisk konsekvensbedömning. Denna tredje punkt har hittills visat sig vara ett stort hinder för genomförandet av kontrafaktiska konsekvensbedömningar av ESF-insatser och förtjänar därför särskild uppmärksamhet.

Ruta 3 Urvalsfrågor för att välja insatser som ska genomgå en kontrafaktisk konsekvensbedömning

Kontrafaktiska konsekvensbedömningar är inte lämpliga för alla insatser och det är oftast inte kostnadseffektivt att låta alla insatser genomgå sådana bedömningar. De förvaltande myndigheterna bör fördela sina medel för att uppnå största möjliga fördel. Dessa val ska anges i utvärderingsplanen. Tre huvudsakliga aspekter bör övervägas i planeringen av kontrafaktiska konsekvensbedömningar: a) strategiska prioriteringar, b) genomförbarhet, och c) tillgång till nödvändiga data.

Utvärderingsstrategin påverkas av faktorer som omfattning, politisk utveckling, oklarheter och behovet av kunskap. De förvaltande myndigheterna bör ställa följande frågor:

- Ett stort belopp har anslagits för insatsen. Innebär detta att det är särskilt viktigt att motivera utgiften? Sådana insatser är relativt enkla att identifiera eftersom de tilldelas merparten av medlen för varje specifikt mål (enligt artikel 2 i ESF+-förordningen (förordning (EU) 2021/1057)).
- Inriktas åtgärden i fråga på en reformprocess och kan utvärderingsresultaten bidra till en kritisk granskning av insatsen? Sådana insatser är kopplade till nyligen genomförda reformer av arbetsmarknaden, utbildningssystemet eller socialpolitiken. De tilldelas inga större belopp, men är viktiga för att reformen ska bli effektiv.
- Är insatsen innovativ och har den testats i pilot- eller försöksskala innan den utvidgades? Sådana insatser kanske inte heller tilldelas några större resurser, men det krävs en bedömning redan i ett tidigt skede för att besluta om de ska fortsättas och utökas eller om försöket ska avslutas.
- Inriktas insatsen på områden där det krävs ytterligare belägg för deras ändamålsenlighet? Denna grupp omfattar insatser som inte har utvärderats tidigare eller har utvärderats för så länge sedan att det krävs en uppdatering. Det finns många sådana insatser inom ESF+-programmen eftersom kontrafaktiska konsekvensbedömningar har börjat tillämpas i bred omfattning först nyligen och "nettoeffekterna" är okända.

Genomförbarhet avser både insatsernas kännetecken och de omständigheter de genomförs under. Som planerare bör du kunna svara ja på följande frågor:

- Är insatsens behandling diskret, särskiljande och tillräckligt enhetlig?
- Är jämförelsen mellan behandlings- och kontrollgrupperna tillräckligt meningsfull för att det ska vara möjligt att mäta effekten?
- Är målgruppen tillräckligt stor och väl definierad?
- Är teorin som kopplar insatsen till de avsedda målen logiskt konsekvent?
- Kan behandlingsgruppen identifieras tydligt inom målgruppen?
- Är behandlingsgruppen tillräckligt stor?
- Är det möjligt att identifiera en trovärdig kontrollgrupp?
- Kan skillnaden mellan behandlings- och kontrollgruppens erfarenheter upprätthållas tillräckligt länge för att det ska vara möjligt att mäta effekten?

Data är en avgörande faktor. Mätning är en grundläggande aspekt av kontrafaktiska konsekvensbedömningar och mätning kräver kvantitativ information, både om behandlings- och kontrollgrupperna och om förutsättningarna för bedömningen. Exakt vilka data som behövs avgörs vanligen av insatsteorin och den strategi som används för att fastställa det kontrafaktiska scenariot. Förvaltande myndigheter som planerar för kontrafaktiska konsekvensbedömningar bör ställa följande frågor när det är dags att välja insatser:

- Vad är viktigast att veta om personerna i mål- och kontrollgrupperna?
- Vad är viktigast att veta om insatsen i den form som den faktiskt levereras till behandlingsgruppen?
- Mottar kontrollgruppen någon annan behandling? Finns det i så fall uppgifter om detta?
- Vilka uppgifter finns tillgängliga från administrativa källor och andra källor?
- Finns det uppgifter om individuella karriärer?
- Kan individualiserade uppgifter från olika källor kopplas till varandra?

Dessa frågor diskuteras närmare längre fram i detta kapitel.

2.1.1. Prioritera insatser för konsekvensbedömning

Innan specifika insatser prioriteras för kontrafaktiska konsekvensbedömningar bör ett antal mer allmänna strategiska aspekter övervägas. Målet bör vara att välja de insatser där konsekvensbedömningar kan ge mest information om vad som fungerar. De fördelar som väl utformade och noggranna utvärderingar ger är inte bara viktiga för den förvaltande myndighet och det medlemsland som utför eller beställer dem, utan är även värdefulla för andra aktörer och för kommissionen.

Kontrafaktiska konsekvensbedömningar bidrar till att motivera utgifter

Med tanke på kontrafaktiska konsekvensbedömningar behandlar frågor som är centrala för beslutsfattare, särskilt de som ansvarar för resursfördelning, bör konsekvensbedömningarna inriktas på särskilt resursintensiva program och insatser. Ju mer tid och andra resurser som läggs på ett program eller en insats, desto viktigare är det att ta reda på om programmet eller insatsen fungerar och om fördelarna är större än kostnaderna. Kostsamma insatser som inte ger samhällsligt eller ekonomiskt värde kanske behöver omprövas, medan andra insatser med bevisat mervärde kan förtjäna ökad finansiering och uppmärksamhet.

Fokus på resursintensiva insatser

Resultat från färskas utvärderingar av ESF-insatser som finansierades under programperioden 2014–2020 visar att det krävs strategiska justeringar och ökat fokus på centrala politiska mål. Insatser för sysselsättning och arbetskraftens rörlighet visade sig vara mindre ändamålsenliga för äldre och personer som står längre från arbetsmarknaden. Dessa grupper kräver nya och mer effektiva instrument²⁰. Socialpolitiska insatser varierar stort i kostnad per deltagare och typ av insats, men trots detta är det ovanligt med systematiska kostnads-nyttoanalyser av sådana insatser²¹. Utbildningsinsatser visade positiva resultat, men det begränsade antalet konsekvensanalyser gör

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

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

det omöjligt att fullständigt bedöma deras långsiktiga verkningar²². Kontrafaktiska konsekvensbedömningar gör det möjligt att sälla bland insatserna och på så sätt identifiera de mest ändamålsenliga och effektiva strategierna för de aktuella målgrupperna. Därigenom maximeras valutan för pengarna för nya ESF+-program.

Insatser som ingår i reformprocesser

Insatser som är centrala inslag i bredare reformprogram finansieras ofta med stora belopp. ESF-insatser som är centrala för en strategi för social inkludering eller en plan för aktiv arbetsmarknadspolitik uppmärksammas naturligtvis i högre grad.

Insatser som bidrar till politisk innovation

Innovativa och undersökande insatser

Nya och innovativa pilotinsatser är särskilt lämpade för kontrafaktiska konsekvensbedömningar. Att testa effekterna av insatser genom ett pilotprojekt eller ett försök kräver dock en noggrann utvärdering. Utvärderingar genom väl utformade kontrafaktiska konsekvensbedömningar är ännu viktigare om det finns en tydlig avsikt att utöka insatsen eller genomföra den i bredare skala om den skulle visa sig vara effektiv.

Insatser som bidrar till lärande

Att ägna extra uppmärksamhet och resurser åt specifika programområden och särskilda insatser inom dessa områden är motiverat om det finns få eller inga belägg för vad som fungerar. Så är exempelvis fallet om det råder osäkerhet kring den framtida politiska inriktningen och det finns en risk för att man förlitar sig på belägg som inte är direkt relevanta (t.ex. från andra länder).

Ta fram belägg

Utvärderingar av hög kvalitet kommer hela samhället till del. Den information som sådana utvärderingar ger kan användas av aktörer utanför den förvaltande myndigheten. Därför är det viktigt att överväga vilka andra aktörer som kan ha nytta av den föreslagna konsekvensbedömningen. Sådana aktörer kan vara de förmedlande organen eller organ med ansvar för insatser inom samma program, andra förvaltande myndigheter eller förmedlande organ inom det berörda medlemslandet eller myndigheter eller institutioner som förvaltar nationella eller regionala medel. En annan uppenbar extern berörd aktör som bör övervägas är Europeiska kommissionen, och det finns även aktörer i andra medlemsländer som kan ha nytta av den aktuella utvärderingen. Genom att ta hänsyn till andra aktörer än de närmaste kan du som är beslutsfattare eller programförvaltare bidra på ett värdefullt sätt till ömsesidigt lärande.

Ett sista strategiskt övervägande i valet av vilka områden som ska prioriteras för kontrafaktiska konsekvensbedömningar är att fundera på vilka insatser som kan visa fördelarna med denna typ av bedömningar och fungera som modeller.

Främja kontrafaktiska konsekvensbedömningar

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

2.1.2. Välja insatser som är lämpade för kontrafaktiska konsekvensbedömningar

Efter att ha gått igenom de allmänna strategiska överväganden som kan motivera valet av insatser för kontrafaktiska konsekvensbedömningar tittar vi i detta avsnitt närmare på vilka särskilda kännetecken som gör insatser lämpade för denna typ av konsekvensbedömningar. Kännetecknen är många och varierande. Vissa aspekter av en insats kan vara lämpade för en kontrafaktisk konsekvensbedömning under vissa omständigheter, men inte passa för andra omständigheter. Det är därför inte möjligt att exakt beskriva vilka faktorer som bör övervägas. Vissa särskilda kännetecken som gör insatser lämpade för kontrafaktiska konsekvensbedömningar är dock värda att nämna.

Ruta 4 De vanligaste typerna av insatser och målgrupper som väljs för kontrafaktiska konsekvensbedömningar inom ramen för ESF

De flesta kontrafaktiska konsekvensbedömningar inom ESF fokuserar på aktiv arbetsmarknadspolitik som riktas mot arbetslösa, och i ett ytterligare steg arbetslösa personer som är missgynnade på ett eller annat sätt. Det finns många konsekvensbedömningar av insatser för ungdomar, vilket beror på att sådana utvärderingar ska genomföras med regelbundna mellanrum enligt sysselsättningsinitiativet för unga.

Ungefär hälften av de kontrafaktiska konsekvensbedömningar som identifierats i projektet för hjälpcentralen för utvärdering²³ sedan 2015 avser tematiskt mål 8, "främja hållbar och kvalitativ sysselsättning och arbetskraftens rörlighet". De flesta kontrafaktiska konsekvensbedömningar av insatser som finansieras inom tematiskt mål 9 inriktas likaså på ändamålsenligheten av insatser för att integrera utsatta arbetslösa personer på arbetsmarknaden (t.ex. den tyska insatsen för långtidsarbetslösa, som finansieras inom det operativa programmet Bund ESF 2014–2020).

De typer av stöd till arbetslösa som oftast analyseras är åtgärder för utbildning, praktikplatser eller andra typer av yrkeserfarenhet och subventionerade arbeten. Detta framgår tydligt av de exempel som analyseras i denna vägledning (se insatsen i Marche i Italien, utvärderingen av utbildningskurser för migranter i Tyskland och yrkesutbildningsprojektet för unga som varken arbetar eller studerar i Lettland). Analyser av yrkesrådgivnings- och jobbmatchningstjänster är mindre vanliga (ett exempel i vägledningen kommer från Sverige, där en pilotåtgärd för att öka arbetsförmedlingarnas stöd till arbetslösa utvärderades med en randomiserad metod). ESF-insatser till stöd för egenföretagande eller skapande av företag utvärderas sällan enligt uppgifter från hjälpcentralen²⁴.

Ibland analyseras olika typer av insatser tillsammans i kontrafaktiska konsekvensbedömningar, vilket innebär en risk för att olika typer av interventionslogik blandas och gör jämförelsen mellan behandlings- och kontrollgrupper mindre trovärdig.

Försöken att utvärdera insatsers effekter för utsatta människor är värda att nämnas, inte bara när det gäller sysselsättningsresultat, utan även "mjuka" resultat: en kontrafaktisk konsekvensbedömning som genomfördes i Tyskland inriktades på denna aspekt – man bedömde hur sysselsättningsåtgärder påverkade långtidsarbetslösa genom att mäta indikatorer för uppfattat hälsotillstånd, tillfredsställelse med tillvaron, känsla av tillhörighet och social status.

ESF-insatser på utbildningsområdet bedöms mer sällan med hjälp av kontrafaktiska konsekvensbedömningar. Färre kontrafaktiska konsekvensbedömningar avser tematiskt mål 10 än de tematiska målen 8 eller 9. Bedömningarna inriktas ofta på insatser som finansieras inom ramen för det integrerade programmet 10.IV om yrkesutbildning och dess effekter när det gäller integrering på arbetsmarknaden. Detta beror bland annat på mer begränsad tillgång till datauppsättningar och strängare integritetsregler²⁵. Försök har dock gjorts i den riktningen: i Spanien (Asturien) bedömdes en insats vid gymnasieskolor för att förebygga skolavhopp, i Polen (Podlasien) undersöktes effekterna av ett projekt för

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

²⁴ Nedan följer några exempel: [Borik V. m.fl., 2015](#) – [Ires Piemonte, 2019](#) – [Openfield, 2019](#).

²⁵ Se t.ex. [Ismeri Europa – Ecores – Institute for Employment Studies, 2019](#).

främjande av yrkesinriktad utbildning bland elever genom en kontrafaktisk konsekvensbedömning, i Portugal analyserades studiebidrag till högskolestudenter.

Kontrafaktiska konsekvensbedömningar kan genomföras på olika politiska nivåer (dvs. inom ett eller flera prioriterade områden, specifika mål eller insatser²⁶ inom ett program), och kan omfatta nationella eller regionala ESF+-program. De kan inriktas på enhetliga målgrupper (män eller kvinnor, ungdomar eller utsatta grupper, långtidsarbetslösa osv.) eller typer av insatser (t.ex. utbildning, tjänster för social inkludering eller förebyggande av skolavhopp (se föregående ruta)).

Exempel från medlemsländerna visar att en rad olika instrument som används inom ESF+ är lämpade för kontrafaktiska konsekvensbedömningar, bland annat yrkesinriktad utbildning, sysselsättningsincitament och arbetsmarknadstjänster (t.ex. yrkesrådgivning, handledning). Det är däremot svårare att genomföra kontrafaktiska konsekvensbedömningar av insatser för arbetsrotation och arbetsdelning, stimulansåtgärder för nyetablering eller stöd till system och strukturer och insatser för utbildning och social inkludering.

Följande frågor kan vara till hjälp i bedömningen av vilka insatser som är lämpade för kontrafaktiska konsekvensbedömningar:

Är insatsen särskiljande och förhållandevis enhetlig?

Det måste vara möjligt att den eller de behandlingar som tillhandahålls genom insatsen från behandlingar i andra insatser. Det måste dessutom finnas en meningsfull skillnad mellan vad insatsens deltagare får och vad andra liknande grupper av individer får. Om behandlingarna är så suddiga att det inte är möjligt att urskilja en diskret grupp av mottagare är en kontrafaktisk konsekvensbedömning varken möjlig eller önskvärd.

*Tydligt urskiljbar
behandling*

Kontrafaktiska metoder blir mycket komplicerade och svåra om en viss enhets behandlingsstatus (ett företag eller en individ) påverkar det potentiella resultatet för andra enheter (genom så kallade allmänna jämviktseffekter). I utbildningsprogram kan denna situation uppstå om utexaminerade från programmet gör det svårt för andra personer utan utbildning att hitta arbete på kort sikt. Om detta anses vara ett stort problem (t.ex. när det gäller storskaliga insatser) kan makroekonomiska analyser krävas för att bedöma omfattningen av substitutions- och undanträngningseffekter. De förvaltande myndigheterna bör inhämta expertrådgivning om sådana effekter kan förekomma.

Insatsen i sig bör vara förhållandevis enhetlig. Detta innebär att alla deltagare i en insats bör motta i stort sett samma åtgärder. Om stödåtgärderna till deltagarna är alltför olika inom en enskild insats kan detta få en rad följder för den kontrafaktiska konsekvensbedömningen. För det första kan insatsen inte sägas vara enhetlig, utan det handlar snarare om insatser med separata kausala processer. För det andra blir det svårt att toka de effekter som rapporteras som genomsnittliga nettoeffekter för en grupp med olikartade insatser. För det tredje kan det vara berättigat att analysera delgrupper, men om det finns för många delgrupper inom

Enhetliga insatser

²⁶ Enligt artikel 2.4 i förordning (EU) 2021/1060 avser "insats" a) ett projekt, ett kontrakt, en åtgärd eller en grupp projekt som valts ut inom de berörda programmen, b) när det gäller finansieringsinstrument, ett programbidrag till ett finansieringsinstrument och det ekonomiska stöd som sedan ges till slutmottagare genom det finansieringsinstrumentet.

behandlingsgruppen kan begränsningar i urvalsstorleken innebära att en sådan analys inte ger användbara slutsatser.

Jämförs behandlingen med ingen behandling eller finns det andra relevanta behandlingsformer?

ESF samfinansierar nationella och regionala arbetsmarknadspolitiska åtgärder och åtgärder för social inkludering. När man tar fram ett system för kontrafaktiska konsekvensbedömningar bör man således noggrant beakta om insatserna är tydligt identifierbara och individerna har möjlighet att tillgodogöra sig tjänster från andra (nationella eller regionala) program och finansieringskällor. Det viktiga är att de utvärderade behandlingarna faktiskt förändrar de möjligheter eller resurser som finns tillgängliga för deltagarna jämfört med vad som finns tillgängligt för kontrollerna och att skillnaderna kan mätas och övervakas.

Sådana komplexa behandlingsfrågor brukar vara kontextspecifika och komplicerar utformningen och genomförandet av kontrafaktiska konsekvensbedömningar. Det är därför viktigt att utvärderingarna planeras noggrant i förväg.

*Komplexa
behandlingar*

Finns det en stor och väl definierad målgrupp?

Kontrafaktiska konsekvensbedömningar kräver stora urvalsstorlekar jämfört med andra utvärderingsformer. Det är viktigt att målgrupperna består av ett lämpligt antal individer och det måste vara möjligt att urskilja tillräckligt stora kontrollgrupper inom dessa målgrupper. Denna fråga diskuteras närmare nedan.

Stort urval

Det är viktigt att insatser som övervägs för kontrafaktiska konsekvensbedömningar riktas mot en väl definierad grupp. Om det inte finns en tydlig förståelse av målgrupperna är det svårt att definiera en meningsfull kontrollgrupp. Vissa insatser har en uttrycklig avsikt att "rekrytera" individer till behandling via informella mekanismer, dvs. processer för att uppmuntra till deltagande som inte är fastställda på förhand eller är alltför normativa (sociala grupper som är svåra att nå ut till, t.ex. unga som varken arbetar eller studerar eller missgynnade grupper, kan involveras genom tillfälliga procedurer eller från fall till fall). I sådana fall kan det vara svårt att exakt identifiera de behandlade individerna och den relaterade kontrollgruppen.

*Fastställa
målgruppens
identitet*

Finns det en tydlig kausal mekanism?

Som redan har nämnts när de viktigaste utvärderingsmetoderna och "förändringsteorin" presenterades är det ofta användbart att utföra en teoretisk utvärdering innan eller i kombination med en kontrafaktisk konsekvensbedömning. Att ta fram en förändringsteori eller en detaljerad logik för insatsen kan vara till hjälp för dig som utformar kontrafaktiska konsekvensbedömningar på en rad olika sätt, särskilt när det gäller att fastställa om insatsen underbyggs av en sammanhängande kausal mekanism. Insatser som inte har en tydlig och övertygande kausal mekanism kommer sannolikt inte att ge effekter som är tillräckligt stora för att de ska kunna fastställas statistiskt med hjälp av en kontrafaktisk konsekvensbedömning.

*Distinkt politisk
mekanism*

Kan utfallen definieras kvantitativt?*Resultatmätning*

Kvantifierbara mått av utfall (eller resultat) är ett måste. Sådana uppgifter och indikatorer kan inhämtas från administrativa källor eller särskilt riktade undersökningar.

I vissa fall kan insatser ha avsedda resultat som kräver särskilda mått och steg för att kunna mätas kvantitativt. Insatsen kan till exempel handla om föränderliga attityder, övertygelser eller åsikter. I sådana fall måste undersökningarna utformas för att mäta dessa förändringar. Andra insatser kan ha mycket vaga eller otillräckligt definierade resultat. Även här kan det vara bra att ta fram en interventionslogik, eftersom detta kan vara till hjälp för att förstå målet för insatsen och hur det är meningen att den ska bidra till att åstadkomma förändringar i de berörda resultaten.

Införs insatsen på ett sätt som gör det möjligt att hitta en meningsfull kontrollgrupp?

För att hitta en meningsfull kontrollgrupp är det viktigt att överväga hur de behandlade enheterna (personer eller företag) väljs ut för insatsen eller varför de bestämmer sig för att delta, om samma datakälla – t.ex. samma undersökningsinstrument – kan tillämpas på kontrollurvalet och på behandlingsgruppen, och slutligen om det är nödvändigt att välja ett kontrollurval som är föremål för samma arbetsmarknadsförhållanden som behandlingsgruppen. Några exempel ges i rutan nedan.

Urvalsmekanismer för behandling

Om en insats är obligatorisk och genomförs mer eller mindre samtidigt för hela målgruppen kan det vara svårt att hitta en obehandlad del av målgruppen som kan användas som kontroll.

Ruta 5 Fastställa kontrollgrupper

I de exempel på kontrafaktiska konsekvensbedömningar som undersöks i denna vägledning valdes kontrollgrupperna utifrån insatsernas kännetecken (t.ex. kriterier för stödberättigande) och även tillgång till lämpliga uppgifter.

Om vi tittar på erfarenheten av kontrafaktiska konsekvensbedömningar från den föregående programperioden ser vi att kontrollgruppen oftare identifieras och väljs baserat på administrativa data, särskilt arbetslöshetsregister. Den vanligaste strategin är att identifiera potentiella kontrollindivider med liknande kännetecken som är anmälda som arbetslösa hos den offentliga arbetsförmedlingen under en viss tid, vilket krävs enligt kriterierna för stödberättigande. Denna strategi tillämpades i utvärderingar i Italien (Marche, provinsen Trento och Piedmont), Polen (Lubelskie och Podlasien) samt i de tyska utvärderingarna av utbildningskurser för migranter samt sysselsättnings- och integrationsåtgärder för arbetslösa i Baden-Württemberg.

I andra utvärderingar som inte ingår i de exempel som presenteras i detalj i denna vägledning valdes personer som hade anmält sig men inte valdes ut för behandling som kontrollgrupp. Så var fallet i den italienska utvärderingen av ungdomsgarantin, den tyska utvärderingen av pilotprojektet "Medborgararbete som finansierats 2007–2013" och utvärderingen av Pipol-projektet som genomfördes i Friuli Venezia Giulia²⁷. Denna strategi är dock ovanlig, eftersom informationssystemen i de flesta länder och regioner inte innehåller uppgifter om personer som ansökt men inte deltagit.

Den tyska strategin för utvärderingar av insatser för långtidsarbetslösa²⁸ var annorlunda utvärderingarna inriktades på "avsikten att delta i behandling" och inte "genomsnittliga behandlingseffekter på behandlade". I detta fall bestod behandlingsgruppen av personer som potentiellt var stödberättigade under programmets

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

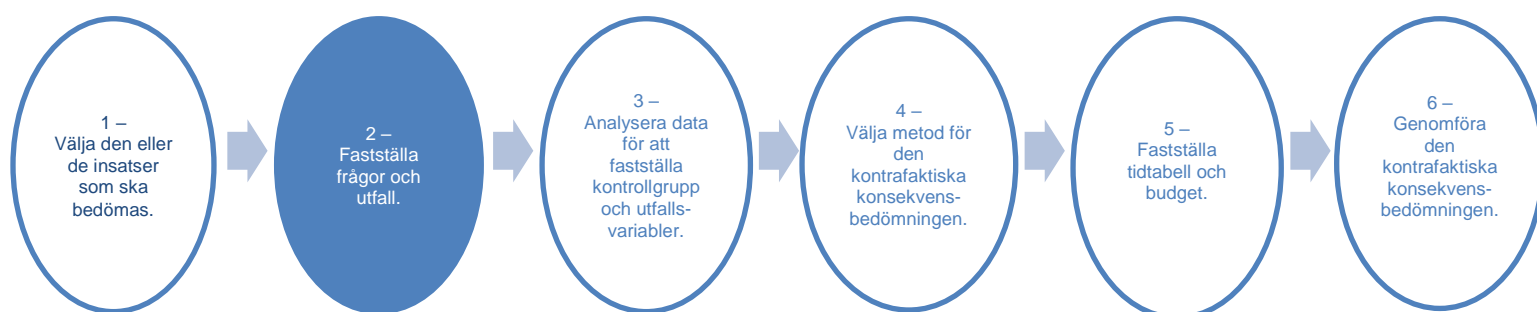
²⁸ Se [Boockmann B. m.fl., 2019](#).

genomförandeperiod, oavsett om de faktiskt behandlades eller ej (från 2015), medan kontrollgruppen bestod av personer med samma kriterier för stödberättigande men som var långtidsarbetslösa före programmets genomförande (2010–2012).

Endast i det svenska exemplet, där en randomiserad metod tillämpades, fastställdes kontrollgruppen slumpmässigt: som standard tillhandahölls behandlingen (stärkt stöd från offentliga arbetsförmedlingar) slumpmässigt till en grupp av ungdomar, medan kontrollgruppen mottog det ordinarie stöd som arbetsförmedlingarna tillhandahåller.

I de få utvärderingar som inriktas på ESF-åtgärder för företag drogs gränsen mellan behandlings- och kontrollgrupperna mellan sökande som fick finansiering och sökande som inte fick det, t.ex. i vissa danska utvärderingar som jämför prestationer för företag som finansierats via ESF med företag i ett urval med liknande kännetecken men som inte mottog stödet²⁹.

2.2. Fastställa utvärderingsfrågor och utfallsvariabler



2.2.1. Vilka syften och mål har insatsen?

Utvärderingsschemat bör först och främst innehålla en beskrivning av insatsens syften, mål och viktigaste kännetecken.

Ofta finns det redan dokument som beskriver insatsens syften och mål. När det gäller kontrafaktiska konsekvensbedömningar är det dock viktigt att exakt ange de resultat och förändringar, dvs. de förväntade effekterna, som insatsen syftar till att åstadkomma.

Det är ofta bra att formulera en förändringsteori för insatsen med en beskrivning av hur insatsens input och aktiviteter är tänkta att kopplas till output, utfall (eller resultat) och därmed effekter.

2.2.2. Vad är syftet med utvärderingen?

När man tar fram ett utvärderingsschema för en kontrafaktisk konsekvensbedömning är det viktigt att man noggrant överväger syftet med utvärderingen. Om det inte står klart varför utvärderingen behövs kommer den sannolikt inte heller att ge de belägg som är avsikten. När det gäller utvärderingar av ESF-finansierade insatser finns det en rad frågor att överväga.

- Vad är syftet med utvärderingen och hur är den utformad inom ramen för kommissionens regler och riktlinjer?
- Vilka aktörer berörs främst av eller deltar i utvärderingen?
- Hur kommer utvärderingens resultat att användas?
- Vilka specifika frågor är utvärderingen avsedd att besvara?

Använda insikter från interventionslogiken i utformningen av kontrafaktiska konsekvensbedömningar

²⁹ Bland annat [Danmarks Statistik m.fl., 2017](#) och [Danmarks Statistik m.fl., 2018](#).

Vilka syften har insatsen och hur är den utformad?

För det första är det viktigt att definiera **skälen** till utvärderingen. Följande anges i förordning (EU) 2021/1060: "Medlemsstaten eller den förvaltande myndigheten ska utvärdera programmen mot ett eller flera av följande kriterier: ändamålsenlighet, effektivitet, relevans, samstämmighet och mervärde för unionen, i syfte att förbättra kvaliteten när det gäller utformning och genomförande av program. Utvärderingarna får också omfatta andra relevanta kriterier, såsom inkludering, icke-diskriminering och synlighet, och får omfatta fler än ett program³⁰." Såsom visas ovan avser resultaten från kontrafaktiska konsekvensbedömningar generellt ändamålsenligheten (i vilken utsträckning de förväntade utfallen har uppnåtts) och effektivitet (insatsens kostnadseffektivitet eller kostnadsnyttoförhållande).

Insatsens syfte och art

Mer allmänt uppmanar kommissionen medlemsländerna att tillämpa resultatorienterade strategier i sina beslutsprocesser och genomföra utvärderingar som uppfyller deras respektive krav vad gäller omfattning, utformning och tidsram. I detta avseende kan kontrafaktiska konsekvensbedömningar av ESF+-program även användas för att besvara specifika utvärderingsfrågor enligt den nationella utvärderingspolicyn.

För det andra måste **utvärderingens art** fastställas³¹:

- Utvärderingar av **effekter** undersöker just effekterna av ett program eller en grupp av program i förhållande till prioriteringar på EU-nivå och nationell nivå (det kan handla om den makroekonomiska effekten av ESF, särskilda politiska åtgärder och teman eller övergripande prioriteringar som barnfrågor eller lika möjligheter).
- Utvärderingar av **processer (eller genomförande)** stöder genomförandet av ett program, analyserar framsteg och genomförandemetoder och ger rekommendationer om hur programmet kan förbättras.

I princip kan den kontrafaktiska metoden tillämpas på konsekvensbedömningar, medan processutvärderingar kräver andra metoder (se även

Figur 1 ovan).

Förordningen om gemensamma bestämmelser föreskriver inte exakt hur många konsekvensbedömningar som ska genomföras under programperioden 2021–2027, till skillnad från programperioden 2014–2020. Förordningen kräver i stället en utvärderingsstrategi med vilken det är möjligt att bedöma hur stödet från EU-fonderna har bidragit till att uppnå programmets mål med avseende på samtliga strategiska profiler³². Förordningen föreskriver inte heller vilka prioriteringar eller insatser som ska utvärderas. Det beslutet lämnas till den förvaltande myndigheten som fastställer detta i utvärderingsplanen. Detta innebär att den blandning av effekt- och processutvärderingar som ska tillämpas på varje program och de prioriteringar och insatser som utvärderingarna koncentreras på ska anges i utvärderingsstrategin. En sista och allmän

³⁰ Artikel 44.1 i förordningen om gemensamma bestämmelser (förordning (EU) 2021/1060).

³¹ [Se Europeiska kommissionen, 2007.](#)

³² Artikel 44.1 i förordningen om gemensamma bestämmelser (förordning (EU) 2021/1060).

konsekvensbedömning ska göras senast i juni 2029³³, men det finns inga tidsgränser för de andra utvärderingar som anges i utvärderingsplanen.

Vilken är den huvudsakliga målgruppen?

Utvärderingens målgrupp ska definieras. Beroende på utvärderingens art kan målgruppen omfatta beslutsfattare, de förvaltande myndigheterna och programförvaltare, andra förvaltande myndigheter eller förmedlande organ i medlemslandet eller nationella eller regionala myndigheter som genomför liknande program. Om institutioner som inte ingår i programförvaltningen tillhandahåller data bör även dessa organ ("dataägare") anses vara berörda aktörer. Det är viktigt att samla alla viktiga aktörer i en styrgrupp för utvärderingen för att skapa ett gemensamt ansvar för arbetet med att utforma och utföra utvärderingen. Utvärderingsexperter från den akademiska världen eller offentliga institutioner bör också ingå i styrgruppen eftersom de bistår den förvaltande myndigheten med teknisk rådgivning.

Identifiera berörda aktörer

Hur kommer utvärderingens resultat att användas?

När utvärderingens målgrupp har identifierats är nästa steg att besluta hur resultaten ska användas. Detta kan göras genom att styrgruppen involveras i arbetet med att ta fram utvärderingsfrågorna och i diskussionerna om specifikationerna.

Två viktiga beslut som resultat av kontrafaktiska konsekvensbedömningar kan bidra till är

- om en befintlig insats bör fortsätta eller
- om en ny typ av insats bör genomföras i bredare skala (dvs. utökas).

I det första fallet kan en kontrafaktisk konsekvensbedömning användas för att utvärdera ändamålsenligheten av ett befintligt eller pågående program om budgeten är pressad och de medel som går till programmet kan användas bättre på annat håll. I denna situation har insatsen sannolikt inte utvärderats tidigare med hjälp av kontrafaktiska scenarier.

I det andra fallet kan det finnas begränsningar i genomförandet av insatsen. Insatsen kanske genomförs i en viss region eller ett visst område i medlemslandet, eller endast under begränsad tid. I sådana situationer kan resultaten från en kontrafaktisk konsekvensbedömning användas för att ta reda på om insatsen är ändamålsenlig och därför kan genomföras på andra håll med lyckat resultat. Denna typ av insatser brukar kallas pilotprojekt eller testas innan de genomförs i bredare skala.

Vilka frågor behöver besvaras?

När insatsens mål och utvärderingens syfte och användningsområden har fastställts och målgruppen har identifierats är det dags att mer detaljerat överväga de frågor som den kontrafaktiska konsekvensbedömningen är avsedd att besvara. Målgrupperna och de berörda aktörerna har ofta frågor av kausal natur som de vill undersöka med hjälp av den kontrafaktiska konsekvensbedömningen. Den förvaltande myndigheten eller utvärderaren bör samla in dessa frågor genom omfattande samråd och ta hänsyn till synpunkter och förslag från berörda personer i olika stadier av insatsen. Därefter beslutar den

³³ Artikel 44.2 i förordningen om gemensamma bestämmelser (förordning (EU) 2021/1060).

förvaltande myndigheten och/eller utvärderaren om vilka frågor som ska prioriteras så att den kontrafaktiska konsekvensbedömningen kan inriktas på de viktigaste och mest relevanta frågorna.

Följande frågor kan vara till nytta för att fastställa en slutlig förteckning över viktiga forskningsfrågor för en kontrafaktisk konsekvensbedömning:

*Centrala
forskningsfrågor*

- Ledde eller bidrog insatsen till de avsedda utfallen på kort, medellång och lång sikt? Fanns det stora skillnader mellan de kortsiktiga och de långsiktiga effekterna? Frågor av denna typ bör prioriteras.
- Är det möjligt att kvantitativt mäta utfallen? Kontrafaktiska konsekvensbedömningar måste bygga på en lämplig datauppsättning (administrativa data eller direkta enkäter till deltagarna), oberoende av vilka källor som föredras.
- I vilken utsträckning kan förändringar i deltagarnas omständigheter eller det socioekonomiska sammanhanget tillskrivas insatserna? Denna typ av frågor kräver ett mått av insatsens nettoeffekter för en jämförelse med deltagare i kontrollgruppen och kontextindikatorer.
- Hade insatsen samma effekter på alla personer i målgruppen? Inriktades insatsen till exempel på såväl långtidsarbetslösa män som kvinnor på samma sätt? Var målgruppen långtidsarbetslösa yngre än 25 år eller äldre än 50? Kontrafaktiska konsekvensbedömningar är ett mycket effektivt verktyg för att undersöka effekter i olika delgrupper om antalet individer i behandlings- och kontrollgruppen är tillräckligt stort.
- Har insatsen varit kostnadseffektiv (jämfört med andra alternativ)? Vilket kostnads-nyttoförhållande har insatsen? Denna effektivitetsanalys förutsätter att den kontrafaktiska konsekvensbedömningen ger ett exakt mått på effekterna och noggranna uppgifter om direkta och indirekta kostnader och fördelar.
- Finns det mycket information och kunskap om effekterna av liknande insatser? I vilken utsträckning kan denna kunskap tillämpas på den undersökta insatsen? Om en politisk åtgärd är välkänd och dess effekter har undersökts noggrant kan det vara bättre att inrikta utvärderingsfrågorna på specifika aspekter av politiken för att undvika att upprepa andra analyser. En detaljerad genomgång av befintlig litteratur kan användas som underlag för beslut i detta avseende och är i alla händelser till nytta vid utformningen av den kontrafaktiska konsekvensbedömningen.

Det är viktigt att redan på förhand ha en tydlig förståelse av de olika forskningsfrågor som den kontrafaktiska konsekvensbedömningen är avsedd att besvara. Diskussioner om vilka frågor som ska behandlas i utvärderingen är ett viktigt inslag i utvärderingsschemat.

Det är viktigt att prioritera frågor och inte falla för frestelsen att överbelasta utvärderingen med för många frågor. Det är en svår balansgång att utforma utvärderingen så att den är relevant för olika aktörer med olika intressen och samtidigt är lätthanterlig. Om utvärderingen omfattar för många forskningsfrågor kan den förlora fokus och bli ytlig. Planeringen av utvärderingen bör vägledas av principen "mindre är mer".

Prioritera frågor

I valet av vilka frågor som ska prioriteras i utvärderingen är det viktigt att först sälla bort alltför likartade frågor. Därefter kan varje fråga poängsättas enligt relevanta principer. Detta kan innefatta följande: Hur pass viktig och engagerad den aktör som formulerade frågan är, huruvida frågan överensstämmer med programmets förändringsteori, frågans relevans för det övergripande syftet med utvärderingen och om frågan är lämplig i förhållande till de uppgifter, den tid och de resurser som finns tillgängliga³⁴. På så sätt rangordnas frågorna efter vikt så att de lämpligaste frågorna kan väljas.

Ibland är det möjligt att samla en grupp av utvärderingsfrågor i en mer allmän fråga, till exempel under frågan "vilken nettoeffekt har insatsen gett"? Andra frågor, till exempel effekter på olika grupper av deltagare, kanske även under olika tidpunkter, kan också samlas i en övergripande fråga. Målet ska dock alltid vara att skapa en hanterlig uppsättning frågor som faktiskt kan besvaras.

Vilka utvärderingskriterier kan tillämpas på utvärderingen?

Förhållandet mellan utvärderingskriterier och utvärderingsfrågor diskuteras ovan, men är värt att klargöras ytterligare. Utvärderingskriterier (effektivitet, ändamålsenlighet, EU-mervärde, samstämmighet osv.) är nödvändiga för att tilldela insamlade belegg ett värde och för att kunna bedöma en politisk åtgärd (effektivitet osv.). Utvärderingsfrågor är alltså nödvändiga för att tydliggöra syftet med utvärderingen och fokusera på de viktigaste politiska frågorna. Utvärderingskriterier och utvärderingsfrågor är dock sammankopplade. Varje fråga kan oftast hänföras till ett visst kriterium. Detta inbördes förhållande är viktigt för såväl förvaltande myndigheter som utvärderare eftersom det kopplar utvärderingens utformning, som nödvändigtvis baseras på de frågor som ska besvaras, till kraven i förordningen om gemensamma bestämmelser för programperioden 2021–2027, som i sin tur är kopplade till utvärderingsfrågorna på ett annat sätt.

Utvärderingskriterier och utvärderingsfrågor

Nedan ges några exempel på vanliga utvärderingsfrågor som grupperats med respektive utvärderingskriterier. Frågorna har valts från Europeiska kommissionens verktygslåda för bättre lagstiftning³⁵ (verktyg nr 47 om utvärderingskriterier och frågor) och anpassats så att exemplen motsvarar frågor som kräver en konsekvensanalys och när så är möjligt en kontrafaktisk konsekvensbedömning:

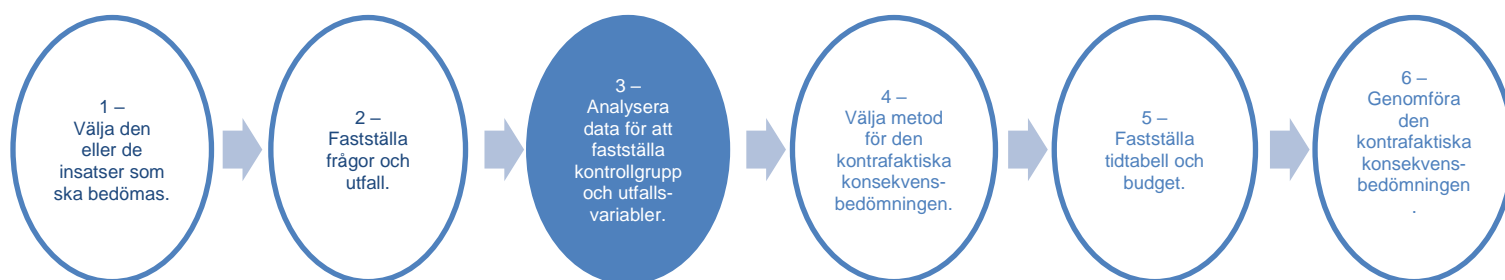
- *Exempel på vanliga frågor om ändamålsenlighet*
 - Vilka kvantitativa effekter har insatsen gett?
 - I vilken utsträckning kan dessa förändringar/effekter tillskrivas insatsen?
 - I vilken utsträckning kan faktorer som påverkar de observerade resultaten kopplas till EU-insatsen?
- *Exempel på vanliga frågor om effektivitet*
 - Har insatsen varit kostnadseffektiv?
 - I vilken utsträckning är kostnaderna för insatsen motiverade i förhållande till de förändringar/effekter som har uppnåtts?

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

³⁵ Se [European Commission, 2017](#).

- Om det finns stora skillnader i kostnader (eller fördelar) mellan olika områden, vad orsakas dessa skillnader i så fall av? Vilken koppling har dessa skillnader till insatsen?
- *Vanliga frågor om EU-mervärde*
 - Vilket mervärde ger ESF+-insatsen jämfört med det mervärde som liknande nationella och/eller regionala insatser ger?
 - Vilka skulle vara de mest troliga konsekvenserna av att stoppa eller upphäva den nuvarande ESF+-insatsen?

2.3. Data för att fastställa kontrollgruppen och mäta utfallsvariablerna



2.3.1. Finns de uppgifter som behövs tillgängliga eller är det möjligt att få tillgång till dem?

Under diskussioner med förvaltande myndigheter och utvärderingsexperter i EU har det framkommit att tillgång till lämpliga uppgifter är ett av de största problemen i samband med genomförandet av kontrafaktiska konsekvensanalyser. Ett viktigt praktiskt övervägande är om de typer av uppgifter som krävs finns tillgängliga. I detta avsnitt presenteras en förenklad kategorisering av de typer av uppgifter som behövs, tillsammans med en diskussion om de källor som sådana uppgifter kan hämtas från och huvudsakliga datainsamlingsmetoder. Den centrala frågan om dataskydd diskuteras också.

Noggrann planering är en absolut förutsättning. Det finns fall där kontrafaktiska konsekvensbedömningar inte har kunnat genomföras som avsett på grund av att det saknats uppgifter. Orsaken till detta är att konsekvensbedömningen inte har planerats i tid. När det gäller befintliga insatser är det viktigt att identifiera de personer som ska ingå i den grupp som mottar behandling och den grupp som inte gör det, eftersom de kommer att stå i centrum för utvärderingen. För nya insatser bör åtgärder vidtas redan i ett tidigt skede för att säkerställa att rätt uppgifter samlas in i god tid.

Vilka typer av uppgifter behövs?

I stort sett krävs tre typer av uppgifter för att utföra en kontrafaktisk konsekvensbedömning. I vissa fall kan en enda datakälla innehålla en eller flera uppgiftstyper. Dessa är uppgifter om behandlings- och kontrollgruppen, utfallsdata och kontextdata.

- **Uppgifter om behandlings- och kontrollgruppen** Det måste finnas datakällor som gör det möjligt för utvärderarna att identifiera de individuella enheterna i behandlings- och kontrollgruppen (företag, personer eller eventuellt geografiska områden).

*Planering av
datainsamlingen*

*Huvudsakliga
datakällor*

- **Utfallsdata:** Såsom framgår av diagrammen 6 och 7 i kapitel 3 av denna vägledning kräver kontrafaktiska konsekvensbedömningar att utfallen mäts både för behandlingsgruppen och kontrollgruppen. Utfallsdata för båda grupperna bör helst samlas in med hjälp av samma insamlingsmetoder och resultatmätningarna bör göras samtidigt.
- **Kontextdata:** Denna typ av data behövs för att välja väl matchade kontroll- och behandlingsgrupper och för att kontrollera eventuella återstående skillnader mellan grupperna genom en analys. Det är viktigt att samla in så mycket uppgifter som möjligt om enheternas kännetecken och faktorer som kan ha samband med både valet att delta i insatsen och med de potentiella resultaten, särskilt om resultatindikatorer som mäts före behandlingen. Kontextdata kan även omfatta data som beskriver de lokala arbetsmarknaderna (t.ex. lokala arbetslöshetstal eller mått på arbetskraftsbrist) samt data som möjliggör analyser per delgrupp.

I tabell 2 nedan beskrivs dessa tre datatyper, med förslag på källor som de kan hämtas från. Exempel på data som används för kontrafaktiska konsekvensbedömningar av ESF-insatser ges i den första rutan nedan, medan den andra rutan innehåller exempel på integrerade datauppsättningar som används i analyser av arbetsmarknadsfrågor och i utvärderingar av arbetsmarknadspolitiska åtgärder.

Tabell 2 Datatyper och källor

Datatyper	Källor
Data om behandlingsgruppen	<ul style="list-style-type: none"> - Register över deltagare i insatsen (innehas vanligen av stödmottagarna). - Övervakningsdata för ESF+ (insatsens kännetecken, start- och slutdatum, remissregister, ansökningsregister).
Data om kontrollgruppen	<ul style="list-style-type: none"> - Administrativa data, t.ex. om socialförsäkrings-, utbildnings- och arbetslöshetsersättningsregister. - Deltaganderegister (personer som var berättigade att delta men inte gjorde det av andra skäl än stödberättigande³⁶). - Befintliga nationella undersökningar såsom arbetskraftsundersökningen.
Utfallsdata (behövs både för behandlings- och kontrollgrupper)	<ul style="list-style-type: none"> - Administrativa data, t.ex. socialförsäkrings- och arbetslöshetsregister, kan också användas för att ta fram utfallsmått (förmåner/socialbidragsutbetalningar), liksom nationella försäkrings- och skatteregister (inkomster och sysselsättningsresultat). - Administrativa utbildningsregister (standardiserade kompetenstester, antal utexaminerade, anmälan och deltagande). - Officiella företagsregister eller skatteregister (produktivitet eller omsättning före och efter intern utbildning eller nyanställningar). - Sysselsättnings- eller outputregister från officiell statistik (vid territoriella kontrafaktiska analyser för att mäta sysselsättnings- och BNP-nivåer). - Anpassade undersökningar av behandlings- och kontrollgrupper.
Kontextdata/kontrollvariabler (behövs både för behandlings- och kontrollgrupper)	<ul style="list-style-type: none"> - Administrativa system – t.ex. register över förmåner som visar ansökningshistorik före behandlingen. Nationella försäkrings- och skatteregister, historiska inkomster och sysselsättningsuppgifter. - Officiell arbetsmarknads- eller utbildningsstatistik (t.ex. arbetskraftsundersökningen, grundläggande data på regional eller nationell nivå som även ger mikrodata på individuell nivå för specialbehandling). - Undersökningar av kontroll- och behandlingsgrupperna. Om reglerna för behandlingen är tydliga kan kontrollgrupper identifieras i förväg och referensdata samlas in. - Övervakningssystem – under vissa omständigheter kan övervakningssystem användas för att sammanställa mått för både behandlings- och kontrollgrupperna, t.ex. ansökningsystem, där sökande som inte godkänts kan användas som kontroller.

Ruta 6 Exempel på data som används för kontrafaktiska konsekvensbedömningar

Bland de exempel som undersöks i detalj i denna vägledning har de tyska utvärderingarna de mest detaljerade administrativa datauppsättningarna. I de ”integrerade anställningsbiografierna” (IEB) samlas en rad olika administrativa informationskällor som förvaltas av den federala arbetsförmedlingen och innehåller uppgifter om anställning (med undantag för egenföretagare) och arbetslöshetsperioder, sociodemografiska kännetecken, betalningsöverföringar (arbetslöshetsersättning) och deltagande i aktiva arbetsmarknadspolitiska åtgärder. IEB har en lång historia och det har krävts kontinuerliga ansträngningar och investeringar för att ta fram och förvalta datauppsättningen. IEB beskrivs närmare i ruta 7³⁷. I den tyska

³⁶ Dessa deltagare var berättigade att delta men godkändes inte för deltagande. Nekat deltagande beror vanligen på att medlen har tagit slut eller på andra externa orsaker (att läraren har blivit sjuk, trafikstörningar osv.). Dessa individer kan bilda en prioriterad kontrollgrupp eftersom de är stödberättigade och är lika intresserade av att delta som de individer som faktiskt deltar. Ibland finns det dock inte tillräckligt många sådana individer för att skapa en kontrollgrupp och i andra fall saknas uppgifter.

³⁷ Ett liknande exempel är den irländska databasen ”Jobseekers Longitudinal Dataset” (JLD), som samlar information om

utvärderingen av sysselsättningsåtgärder kompletterades dessa uppgifter med en undersökning som strukturerades i tre omgångar för att samla in information om ”mjuka utfall” (deltagarnas uppfattningar) i olika skeden av insatserna.

Liknande data användes i de italienska utvärderingar som undersöks i detalj i denna vägledning. Arkiv över arbetslösa personer som var anmälda hos den offentliga arbetsförmedlingen användes för att identifiera potentiella kontrollgrupper, medan administrativa uppgifter om byte av anställning (*Comunicazioni obbligatorie* – COB), ett register över alla anställningsavtal som privata företag och offentliga arbetsgivare ingår (förutom egenföretagare) användes för att mäta utfallvariablerna. Uppgiftskvaliteten varierar eftersom COB-registren förvaltas på regional nivå. Landsomfattande normer har dock införts under de senaste åren och kvaliteten har förbättrats. I en utvärdering (för provinsen Trento) kunde utvärderaren även slå ihop dessa uppgifter med deklaraionsuppgifter från INPS (det nationella socialförsäkringsinstitutet) och kunde på så sätt mäta genomslaget utifrån de berörda personernas inkomster.

I Polen användes administrativa data från den offentliga arbetsförmedlingens arbetslöshetsregister för att identifiera kontrollgrupper i utvärderingen. Dessa data innehåller dock inte information om sysselsättningsstatus eller historik för de berörda personerna. Utvärderingarna baserades i stället på ett närmevärde (genom att uppgifter om antalet personer som är anmälda som arbetslösa hos den offentliga arbetsförmedlingen stryks) eller genom insamling av information via undersökningar av ett urval av individer, både behandlade och obehandlade.

I exemplet från Lettland kunde utvärderaren slå ihop två huvudsakliga administrativa datauppsättningar. Data från Lettlands statliga arbetsförmedling gav information om både deltagare och icke-deltagare som var anmälda som arbetslösa på vissa datum, och data från det statliga skattekontoret gav information om sysselsättningsstatus på olika datum samt inkomstnivåer. Tack vare denna information kunde utvärderaren bedöma effekter både när det gällde sannolikheten för att de berörda personerna skulle hitta arbete vid olika tidpunkter och deras inkomstnivåer.

I det svenska exemplet, som är det enda fallet av randomisering, användes både administrativa data för att mäta utfallsvariabler och en enkät till arbetsförmedlingar och förmedlande organ för att mäta stödintensiteten och de typer av stöd som tillhandahölls.

I utvärderingarna på utbildningsområdet användes administrativa data för att identifiera kontrollgruppen och bedöma utfallvariablerna, med undantag för den polska utvärderingen, där man använde en undersökning för att mäta de variabler som användes som utfall.

Exemplen visar överlag att det är viktigt att det finns tillgång till lämpliga administrativa data för kontrafaktiska konsekvensbedömningar. Uppgifterna ska helst kunna användas både för att identifiera kontrollgrupper och mäta utfallvariablerna. Förvaltande myndigheter som planerar att genomföra en kontrafaktisk konsekvensbedömning bör därför redan på förhand försäkra sig om att administrativa data för att utföra utvärderingarna finns tillgängliga och vidta lämpliga åtgärder för att lösa eventuella problem med tillgänglighet, integrering av data eller andra problematiska aspekter.

Ruta 7 Exempel på integrerade databaser för kontrafaktiska konsekvensbedömningar

Jobseekers longitudinal dataset (JLD) i Irland

Jobseekers Longitudinal Dataset (JLD) är en administrativ databas som förvaltas av det irländska socialskyddsdepartementet (DSP).

JLD är ett ambitiöst initiativ för att anpassa administrativa data till forskningsändamål. Arbetet med databasen inleddes för ungefär tio år sedan, efter det att socialskyddsdepartementet gett University College i Dublin i uppdrag att genomföra en preliminär studie av förvaltningen av Live-registret³⁸ och mer allmänt av arbetsmarknadsdata. I studien gavs flera förslag om hur datainsamlingen kunde förbättras. Flera problem framkom också (dubbel datainsamling i olika datasystem, saknad information osv.).

JLD omfattar flera informationskällor: betalningsuppgifter och administrativa data från

utbetalningar för socialt stöd och socialförsäkringar, arbetsmarknadsprogram och anställningshistorik i en enda databas. JLD användes bland annat i utvärderingen av aktiveringsprogrammet JobBridge, som finansierades under programperioden 2007–2013. Se [Indecon, 2016](#).

³⁸ Live-registret innehåller uppgifter om personer som ansöker om arbetslöshetsbidrag, arbetslöshetsersättning eller andra lagstadgade förmåner vid socialskyddsdepartementets lokalkontor.

socialskyddsdepartementet, t.ex. utbetalningar av socialt stöd eller socialförsäkringar till de personer i arbetsför ålder som är registrerade i Live-registret och uppgifter om program för en aktiv arbetsmarknadspolitik som förvaltas av socialskyddsdepartementet, data om arbetsmarknadsprogram som förvaltas av det nationella utbildningsorganet Solas, data som samlas in av skattemyndigheterna (*Revenue Commissioners*). När det gäller ESF-insatser täcker JLD delvis de insatser som finansieras av fonden, men i synnerhet de insatser som finansieras av DSP och Solas³⁹.

Data från de ovannämnda källorna ordnas om i en serie händelser. Händelsen inleds när individen blir arbetslös och avslutas när individen hittar ett arbete eller deltar i en annan aktivitet i utbildningsprogrammet. När individens sysselsättningsstatus förändras igen inleds en ny händelse. Inledningen av en arbetslöshetsperiod sammanfaller med det datum då individen ansöker om arbetslöshetsbidrag eller arbetslöshetsersättning⁴⁰. En fördel med att JLD är strukturerad i händelser är att sammanhängande perioder av arbetslöshetsbidrag eller arbetslöshetsersättning kan kopplas ihop. De utgör då en arbetslöshetsperiod. Såsom visas ovan kan händelserna överlappa varandra, vilket innebär att forskaren måste bestämma sig för vilken händelse som ska prioriteras.

JLD har spårat socialförsäkringsansökningar, aktivering, utbildning och anställningshistorik sedan 2004 och innehåller omkring 13 miljoner händelser och 2 miljoner personer. Varje händelse har ett start- och ett slutdatum samt en operativ kod⁴¹, som gör det möjligt att se vilken situation individen befinner sig i under varje omgång. JLD täcker en stor uppsättning variabler: kön, ålder, civilstånd, medborgarskap, utbildningsnivå, tidigare yrke, anställnings- och arbetslöshetshistorik (varaktighet och antal händelser) samt uppgifter om anställningen (t.ex. bransch), historik över utbildning som erbjudits under arbetslösheten (typ, varaktighet och antal händelser) förmånstyp (arbetslöshetsbidrag eller arbetslöshetsersättning), antal barn, personer i beroendeställning, typ av familjebidrag (dvs. bidrag till den vuxna personen och barnbidrag eller endast bidrag till den vuxna personen), inkomst, skatt och ort⁴².

JLD kopplas dessutom till andra administrativa data via individuella identifierare. För utvärderingen av programmet JobsPlus matchades JLD till exempel med en separat övervakningsdatabas med detaljerad information om start- och slutdatum för programmet, antal dagar i Live-registret i början av programmet och behandlingstyp.

Det finns två huvudsakliga kanaler för åtkomst till JLD-data: forskningsprojekt som socialskyddsdepartementet beställer via ett anbudsförfarande, eller på begäran från forskare med lämpliga meriter. I det senare fallet behandlas begäran om åtkomst från fall till fall, och om begäran godkänns ingås ett rättsligt bindande datadelningsavtal mellan forskarna eller institutionerna och departementet. Alla data pseudonymiseras och endast de datafält som är relevanta för forskningsämnet överförs. Forskarnas eller forskningsinstitutens åtkomst till uppgifterna är begränsad och uppgifterna måste raderas i slutet av forsknings- eller utvärderingsprojektet. Forskarna förses med en kodbok tillsammans med JLD-uppgifterna, men kodboken behöver förbättras betydligt.

JLD har använts för en rad olika forsknings- och utvärderingsverksamheter under årens lopp: Bland de program som har valts ut för utvärdering finns *Back to Education Allowance* (2015), *JobBridge Activation* (2016), *Back to Work Enterprise Allowance* (2017), *JobPath* (2019) och *JobPlus* (2020).

JLD:s struktur och innehåll har förbättrats sedan inledningskedet. Datauppsättningen kräver dock kontinuerliga åtgärder för att hålla den aktuell, ta fram förbättringar och fylla luckor, t.ex. genom att ta med mer detaljerad information om utbildning, integrera mer programspecifika uppgifter (innehåll, slutförande osv.) och förbättra inkomstuppgifterna. Ett kommande projekt syftar till att förbättra JLD genom regelbundna uppdateringar (t.ex. månadsvis) med automatisering och testning av datakanalen och tillägg eller utbyte av datakällor för att säkerställa att sysselsättnings- och utbildningsstöd och andra socialskyddsprogram täcks helt samt inkomstuppgifter i realtid⁴³.

³⁹ Program inom områden som rättvisa (projekten *Youth Diversion* och *Young Persons Probation* samt *Integration and Employment of Migrants*), utbildning (*Third Level Access* och *Adult Literacy*), samhälle (programmet *Community Training Centres*) och försvar (programmet *Defence Forces Employment Support Scheme*) ingår inte i JLD.

⁴⁰ Arbetslöshetsbidrag är en veckovis betalning från DSP för personer som är heltidsarbetslösa eller arbetar deltid på grund av att arbetsgivaren har begränsat arbetstiden. Bidraget betalas ut till arbetssökande som har betalat in minst 104 socialförsäkringsavgifter (PRSI) under det berörda beskattningsåret. Arbetslöshetsersättning är en behovsprövad betalning till arbetslösa personer som söker arbete och inte har rätt till arbetslöshetsbidrag eller är utförsäkrade.

⁴¹ Representerar en kombination av aktiverings-/utbildningsaktiviteter, bidragsansökningar och anställningstid.

⁴² Vissa data avser olika perioder. Inkomster registreras t.ex. årligen, medan socialförsäkringsbetalningar uppdateras varje vecka.

⁴³ Vi vill tacka Frank Humphreys, Ciaran Judge, Saidhbhín Hardiman och Krzysztof Gigon från DSP för den information de lämnat

Institut für Arbeitsmarkt- und Berufsforschung (IAB) i Tyskland och "integrerade anställningsbiografier" (IEB)

I Tyskland förbättrades åtkomsten till data för vetenskapliga ändamål efter arbetsmarknadsreformerna 2003–2005, med betoning på utvärdering av politiska åtgärder och enligt rekommendationerna från kommissionen för förbättring av informationsinfrastrukturen mellan vetenskap och statistik, med målet att inrätta en datacentral hos varje offentligt organ som producerar mikrodata. Till följd av detta inrättade den federala arbetsförmedlingen en forskningsdatacentral inom institutet för arbetsmarknad- och yrkesforskning (*Institut für Arbeitsmarkt- und Berufsforschung* (IAB)) 2004. IAB ansvarar för att extrahera data från administrativa processer i syfte att förbereda datauppsättningar som är lämpade för empirisk forskning.

IAB förvaltar över femton datauppsättningar som finns tillgängliga för forskare. Uppgifterna härrör från administrativa data från socialförsäkringssystemet, den federala arbetsförmedlingens interna processer och från undersökningar som utförs av IAB⁴⁴.

När det gäller socialförsäkringssystemen är alla arbetsgivare skyldiga att rapportera ett antal poster och kännetecken för sina anställda. Dessa uppgifter bildar en rik informationskälla om människors sysselsättningshistorik. De administrativa uppgifterna om den federala arbetsförmedlingens interna förfaranden omfattar uppgifter om arbetslöshetsförsäkring, arbetslöshetsersättning och motsvarande ersättningsperioder, rådgivande möten med arbetslösa och handläggare vid arbetsförmedlingar samt aktiva arbetsmarknadsåtgärder. IAB kombinerar uppgifterna till en enda heltäckande datauppsättning, som kallas "integrerade anställningsbiografier" EIB. Dessa administrativa uppgifter började samlas in 1975, men alla variabler finns inte tillgängliga för hela observationsperioden. IEB kan integreras med ett antal olika uppgifter från undersökningar som utförs av IAB, t.ex. IAB:s etableringspanel, IAB:s undersökning av lediga arbetstillfällen, tyska förvaltnings- och organisationsmodeller (GMOP), panelstudien om arbetsmarknaden och socialförsäkringssystemet (PASS), projektet "Arbeta och lära i en föränderlig värld", undersökningen om bonusbetalningar till anställda, löneökningar och rättvisa (BLoG), IAB:s program *Linked Employer-Employee Data* (LIAB), Panelen "WeLL" – undersökning bland anställda för projektet "fortbildning som en del av livslångt lärande".

IAB uppdaterar sina dataprodukter regelbundet och erbjuder olika urval av dessa omfattande administrativa datakällor för forskningsändamål. För varje dataprodukt tillhandahåller IAB detaljerad dokumentation på tyska och engelska. Den rättsliga grunden för åtkomst till data är den tyska lagen om social trygghet (flera versioner). Mer konkret kan forskare få åtkomst till data på följande fyra sätt: "Campusfiler" (fullständigt anonymiserade uppgifter som endast är användbara för undervisningssyften). Filer för vetenskaplig användning – anonymiserad mikrodata som lämnas ut till vetenskapliga institutioner i Tyskland och EU-medlemsländerna för arbetsmarknadsforskningsprojekt, men inte för undervisning eller kommersiella forskningsintressen. Den vetenskapliga institution som ansöker om åtkomst till data måste garantera att uppgifterna behandlas på ett säkert sätt, medan veckovisa anonymiserade uppgifter med mer detaljerad information endast finns tillgängliga på plats. IAB tillhandahåller separata arbetsplatser i en säker datormiljö i Nürnberg och på flera platser i Tyskland, USA och Förenade kungariket. Forskarna har direkt åtkomst till uppgifterna, men får endast tillgång till programmets output efter det att IAB:s personal har kontrollerat att uppgifterna kan lämnas ut. Programmen fjärrstyrs, vilket innebär att forskarna förbereder sina program med artificiella data och laddar upp dem i applikationen JoSuA (*Job Submission Application*). Forskarna ser aldrig de ursprungliga uppgifterna under processen, utan får endast ut sina resultat. Standardiserade ansökningsformulär finns tillgängliga för alla typer av dataåtkomst. När ansökan har godkänts ingår forskningsinstitutionen och IAB ett avtal om användning av de berörda uppgifterna för ett specifikt projekt och en fastställd tidsperiod. Dataskyddsreglerna anges i avtalet och stränga påföljder tillämpas i händelse av överträdelser. En del av de datauppsättningar som förvaltas av IAB får endast användas på plats (t.ex. de kopplade datauppsättningarna).

Baserat på de uppgifter som tas fram av IAB utvärderas aktiva arbetsmarknadspolitiska åtgärder med de senaste empiriska metoderna och i vissa fall har resultaten lett till att politiken ändrats⁴⁵. IAB ger inte bara

om JLD.

⁴⁴ År 2011 grundades ett centrum för sammanlänkning av register inom IAB, ett gemensamt projekt med Duisburg-Essens universitet som finansierades av den tyska forskningsstiftelsen. Syftet med centrumet är att förenkla sammanlänkning av datauppsättningar utan specifika identifierare.

⁴⁵ Ett exempel är en utvärdering av obligatoriska integrationsavtal mellan arbetssökande och handläggare. Med hjälp av ett

forskare i Tyskland tillgång till sina datauppsättningar, utan även forskare i andra länder. Antalet användare ökar stadigt (under 2016 ingicks t.ex. nästan en tredjedel av alla dataanvändningsavtal med institutioner utanför Tyskland).

Du hittar mer information om användning av administrativa data för kontrafaktiska konsekvensbedömningar och praktiska frågor i samband med detta i den europeiska handledning som tagits fram av forskare från JRC, se [Europeiska kommissionen, 2020](#).

Vilka dataskyddsfrågor kan uppstå i samband med kontrafaktiska konsekvensbedömningar?

Det kan vara svårt att få fram uppgifter som identifierar individer eller företag som har deltagit i insatser som finansieras via ESF+⁴⁶. Kontrafaktiska konsekvensbedömningar kräver mikrodata som innehåller observationer av individuella enheter i både behandlings- och kontrollgrupperna.

Enligt bilaga I till ESF+-förordningen (förordning (EU) 2021/1057) ska uppgifterna om deltagare delas upp efter kön, arbetsmarknadsstatus, åldersgrupp, utbildningsresultat och utsatta grupper (migranter, minoriteter, personer med funktionsnedsättning, andra missgynnade personer). Enligt förordningen om gemensamma bestämmelser och ESF+-förordningen är förvaltande myndigheter under programperioden 2021–2027 skyldiga att samla in och behandla personuppgifter i form av deltagarregister. När det gäller övervakning och utvärdering anges dessutom följande i artikel 17.6 i ESF+-förordningen: "När uppgifter finns tillgängliga i register eller likvärdiga källor får medlemsstaterna göra det möjligt för de förvaltande myndigheterna och andra organ som ansvarar för insamlingen av de uppgifter som behövs för övervakning och utvärdering av det allmänna stödet från ESF+-delen inom ramen för delad förvaltning att inhämta dessa uppgifter från dessa register eller likvärdiga källor i enlighet med artikel 6.1 c och e i förordning (EU) 2016/679."

*ESF+-
förordningens
bestämmelser om
datahantering*

Dessa regler i förordningen om allmänna bestämmelser och ESF+-förordningen underlättar åtkomst till och användning av personuppgifter som behövs för övervakningen av ESF+, och när det gäller kontrafaktiska konsekvensbedömningar, för att definiera behandlingsgruppen. Åtkomsten till de personuppgifter som behövs för att bilda kontrollgrupper och behandla uppgifter om behandlings- och kontrollgrupper måste dock överensstämja med förordning (EU) 2016/679 (allmänna dataskyddsförordningen), som reglerar överföring och användning av personuppgifter i allmänhet inom EU, inklusive särskilda kategorier av personuppgifter⁴⁷. I rutan nedan ges en översikt av de viktigaste bestämmelserna och skyldigheterna i dataskyddsförordningen.

randomiserat fältexperiment och "arbetsmarknadsbiografierna" för de personer som ingick i experimentet kunde IAB visa att för vissa grupper av arbetslösa är det obligatoriska integrationsavtalet kontraproduktivt och att instrumentet bör hanteras mer flexibelt för dessa personer (van den Berg m.fl. 2016).

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

⁴⁷ Begreppet "känsliga uppgifter" används inte längre i dataskyddsförordningen, utan den hänvisar nu till "särskilda kategorier av personuppgifter". Enligt artikel 9 i dataskyddsförordningen utgörs dessa av följande: "personuppgifter som avslöjar ras eller etniskt ursprung, politiska åsikter, religiös eller filosofisk övertygelse eller medlemskap i fackförening och behandling av genetiska uppgifter, biometriska uppgifter för att entydigt identifiera en fysisk person, uppgifter om hälsa eller uppgifter om en fysisk persons sexualliv eller sexuella läggning".

Ruta 8 EU:s regelverk om behandling av personuppgifter

Den relevanta lagstiftningen om behandling av personuppgifter i EU utgörs huvudsakligen av förordning (EU) 2016/679 och riktlinjer och åtgärder som antagits av behöriga myndigheter som Europeiska dataskyddsstyrelsen (EDPB). Lagstiftningen föreskriver ett stort antal villkor och begränsningar för behandling av personuppgifter i syfte att skydda de registrerades rättigheter och friheter. Det krävs dock en rimlig avvägning mellan behovet av skydd enligt lagstiftningen och behovet av att inte begränsa vetenskaplig forskning så att personuppgifter blir en tillgång i forskningen. EU:s regelverk föreskriver därför att undantag får göras på vissa villkor för att möjliggöra forskningsverksamhet och spridning av output, så länge de berörda personernas grundläggande rättigheter skyddas, dvs. rätten till integritet⁴⁸.

DEN ALLMÄNNA DATASKYDDSFÖRORDNINGEN (FÖRORDNING (EU) 2016/679)

Den allmänna dataskyddsförordningen trädde i kraft den 24 maj 2016 och blev fullständigt tillämplig i medlemsländerna den 25 maj. 2018. Förordningen tillämpas på "behandling av personuppgifter som helt eller delvis företas på automatisk väg samt på annan behandling än automatisk av personuppgifter som ingår i eller kommer att ingå i ett register"⁴⁹.

När det gäller det territoriella tillämpningsområdet omfattar förordningen både behandling av personuppgifter som utförs av personer inom EU:s territorium och när behandlingen omfattar registrerade som finns inom EU, även om personuppgiftsansvarige eller personuppgiftsbiträdet finns utanför EU.

I materiellt hänseende och jämfört med den tidigare lagstiftningen stärks de registrerades rättigheter ytterligare genom förordningen och att en rad nya skyldigheter införs för registeransvariga enligt en logik som baseras på riskanalys och principen om ansvarsskyldighet. Förordningen föreskriver dessutom en rad nya krav som registeransvariga ska uppfylla som inte fanns med i den tidigare lagstiftningen. De nya kraven rör bland annat konsekvensbedömning avseende integritet (artiklarna 35–36), minimering av behandling enligt kriterierna för inbyggt dataskydd och dataskydd som standard (artikel 25), införande av register över behandling (artikel 30) och utnämning av dataskyddsombud (artiklarna 37–39).

Det är viktigt att framhålla de huvudsakliga principer som anges i artikel 5 i förordningen och som du som behandlar personuppgifter ska uppfylla. Dessa principer är följande:

- a) Laglighet, korrekthet och öppenhet: personuppgifter ska behandlas på ett lagligt, korrekt och öppet sätt.
- b) Ändamålsbegränsning: personuppgifter ska samlas in för särskilda, uttryckligt angivna och berättigade ändamål och inte senare behandlas på ett sätt som är oförenligt med dessa ändamål. Vid behandling för statistiska ändamål eller vetenskaplig forskning ska den registeransvarige vidta nödvändiga skyddsåtgärder.
- c) Uppgiftsminimering endast uppgifter som är strikt nödvändiga för att uppnå det avsedda ändamålet får samlas in.
- d) Korrekthet: de insamlade uppgifterna ska vara korrekta och, när så är nödvändigt, hållas aktuella.
- e) Lagringsminimering: personuppgifter får inte förvaras i en form som möjliggör identifiering av den registrerade under en längre tid än vad som är nödvändigt för de ändamål för vilka personuppgifterna behandlas. I detta syfte får personuppgifter lagras under längre perioder i den mån personuppgifterna enbart behandlas för arkivändamål i allmänhetens intresse, vetenskapliga eller historiska forskningsändamål eller statistiska ändamål i enlighet med artikel 89. I sådana fall ska personuppgifterna omfattas av lämpliga tekniska och organisatoriska åtgärder för att säkerställa den registrerades rättigheter och friheter.
- f) Integritet och konfidentialitet: Personuppgifter ska behandlas på ett sätt som säkerställer lämplig säkerhet för personuppgifterna, inbegripet skydd mot obehörig eller otillåten behandling och mot förlust, förstörelse eller skada genom olyckshändelse, med användning av lämpliga tekniska eller organisatoriska åtgärder.
- g) Ansvarighet är en av förordningens viktigaste principer: den registeransvarige ansvarar för behandlingen av personuppgifter och ska kunna visa att eventuella nödvändiga åtgärder har vidtagits.

Dessa principer utgör grunden för dataskyddsförordningen och för de skyldigheter som registeransvariga och personuppgiftsbiträden ska uppfylla.

⁴⁸ Annan relevant EU-lagstiftning om behandling av personuppgifter är förordning (EU) nr 557/2013 om europeisk statistik vad gäller tillgång till konfidentiella uppgifter för vetenskapliga ändamål, förordning (EU) 2018/1725 om skydd för fysiska personer med avseende på behandling av personuppgifter som utförs av EU:s institutioner, organ och byråer och om det fria flödet av sådana uppgifter.

⁴⁹ Artikel 2 i Europaparlamentets och rådets förordning (EU) 2016/679.

BEHANDLING AV PERSONUPPGIFTER INOM RAMEN FÖR KONTRAFAKTISKA KONSEKVENSBEDÖMNINGAR: VILLKOR, BEGRÄNSNINGAR OCH VIKTIGA FRÅGOR

Kontrafaktiska konsekvensbedömningar omfattar behandling av stora datamängder, även personuppgifter. När medlemsländerna och deras förvaltande myndigheter genomför kontrafaktiska konsekvensbedömningar är det således viktigt att de känner till det ovannämnda regelverket och de villkor, begränsningar och viktiga frågor som det innehåller.

Rättslig grund för och ändamål med behandlingen

För att behandling av personuppgifter ska vara tillåten måste minst ett av villkoren för laglighet i artikel 6 i dataskyddsförordningen vara uppfyllt⁵⁰. Om de uppgifter som ska behandlas utgör "särskilda kategorier av personuppgifter" krävs även en hänvisning till artikel 9 i dataskyddsförordningen.

När det gäller kontrafaktiska konsekvensbedömningar förefaller **allmänt intresse** enligt artikel 6.1 e i dataskyddsförordningen ("utföra en uppgift av allmänt intresse eller som ett led i den personuppgiftsansvariges myndighetsutövning") vara den lämpligaste rättsliga grunden för behandling av personuppgifter. Denna rättsliga grund ska fastställas i enlighet med unionsrätten eller nationell lagstiftning enligt artikel 6.3 i dataskyddsförordningen. Ett "allmänt intresse" motsvarar tydligt de förvaltande myndigheternas skyldigheter enligt artikel 17.6 i ESF+-förordningen vid användning av uppgifter som "finns tillgängliga i register eller likvärdiga källor". Följande anges dessutom i artikel 4 i förordningen om gemensamma bestämmelser: "Medlemsstaterna och kommissionen får endast behandla personuppgifter om det är nödvändigt för att de ska kunna fullgöra sina respektive skyldigheter enligt denna förordning, särskilt när det gäller övervakning, rapportering, kommunikation, offentliggörande, utvärdering [...]." Registeransvariga kan även tilldelas liknande befogenheter enligt nationell lagstiftning för hantering och behandling av personuppgifter i allmänt intresse.

Andra rättsliga grunder för kontrafaktiska konsekvensbedömningar kan hänvisa till de andra villkor som anges i artikel 6.1, särskilt

- den registrerades **samtycke** (artikel 6.1 a i dataskyddsförordningen). Samtycke kan till exempel vara en lämplig rättslig grund när uppgifterna för den kontrafaktiska konsekvensbedömningen samlas in genom en undersökning och de registrerade enkelt kan ge sitt samtycke till behandlingen (se artikel 7 om samtycke i dataskyddsförordningen). Det kan vara svårt att inhämta de registrerades samtycke utan god förhandsplanering. Denna rättsliga grund kan anses motsvara en alternativ rättslig grund för kontrafaktiska konsekvensbedömningar när andra rättsliga grunder inte kan åberopas.
- Behandling är nödvändig för att **fullgöra en rättslig förpliktelse** som åvilar den personuppgiftsansvarige (artikel 6.1 c i dataskyddsförordningen med hänvisning till artikel 17 i ESF+-förordningen). Denna rättsliga grund måste fastställas i lag och kan omfatta privata eller offentliga enheter. Ett privat eller offentligt organ kan till exempel på grund av särskilda rättsliga föreskrifter vara ansvarigt för en datauppsättning som krävs för att samarbeta med den förvaltande myndigheten i utförandet av en kontrafaktisk konsekvensbedömning.

Den personuppgiftsansvarige får **dessutom använda de insamlade uppgifterna för andra ändamål om de är förenliga med de ursprungliga ändamålen**. I detta avseende ska den registeransvarige beakta villkoren i artiklarna 6.4 och 5.1 b i dataskyddsförordningen. Dessa bestämmelser är särskilt relevanta när data behandlas för vetenskaplig forskning eller statistiska ändamål. Bestämmelserna gör det nämligen möjligt att använda administrativa data för andra ändamål än de ursprungliga. Det krävs inget särskilt samtycke för den nya användningen, men ändamålet ska uppfylla skyddsreglerna i artikel 89.1, främst rörande pseudonymisering (se nedan). Dessa bestämmelser kan till exempel vara relevanta om uppgifter från arbetslöshetsregister ska användas i en kontrafaktisk konsekvensbedömning.

Till följd av de särskilda förutsättningarna för kontrafaktiska konsekvensbedömningar kan den registeransvarige som samlar in uppgifterna och den enhet som utför forskningsarbetet vara olika. I sådana

⁵⁰ Villkoren är följande: a) Den registrerade har lämnat sitt samtycke till att dennes personuppgifter behandlas för ett eller flera specifika ändamål. b) Behandlingen är nödvändig för att fullgöra ett avtal i vilket den registrerade är part eller för att vidta åtgärder på begäran av den registrerade innan ett sådant avtal ingås. c) Behandlingen är nödvändig för att fullgöra en rättslig förpliktelse som åvilar den personuppgiftsansvarige. d) Behandlingen är nödvändig för att skydda intressen som är av grundläggande betydelse för den registrerade eller för en annan fysisk person. e) Behandlingen är nödvändig för att utföra en uppgift av allmänt intresse eller som ett led i den personuppgiftsansvariges myndighetsutövning. f) Behandlingen är nödvändig för ändamål som rör den personuppgiftsansvariges eller en tredje parts berättigade intressen, om inte den registrerades intressen eller grundläggande rättigheter och friheter väger tyngre och kräver skydd av personuppgifter, särskilt när den registrerade är ett barn.

fall måste det finnas ett villkor som gör dataöverföringen laglig, så att mottagaren kan gå vidare med bedömningen.

Avidentifiering och pseudonymisering behandling för statistiska ändamål

Enligt artikel 89 i dataskyddsförordningen ska behandling av uppgifter som utförs för allmänintressen, forskningsändamål eller statistiska ändamål omfattas av lämpliga skyddsåtgärder för de registrerades rättigheter och friheter, och ska särskilt vara förenlig med principen om ”minimering”. Detta innebär att pseudonymiseringstekniker ska användas⁵¹.

När ändamålen kan uppfyllas genom ytterligare behandling av uppgifter som inte medger eller inte längre medger identifiering av de registrerade ska ändamålen uppfyllas genom att uppgifterna avidentifieras och sedan behandlas i aggregerad form. I kontrafaktiska konsekvensbedömningar är resultaten aggregerade per definition och denna risk föreligger inte, såvida inte de ursprungliga datauppsättningarna offentliggörs av statistiska skäl. I detta fall måste datauppsättningarna anonymiseras (se exempel i rutan nedan).

Datalagring och säker databehandling

Datalagring är en av de viktigaste aspekterna av databehandling. Lagstiftningen föreskriver inte specifikt hur data ska lagras, men de principer som nämns ovan innebär att lagring och behandling alltid måste vara kopplade till forskningsändamålet. När ändamålet med behandlingen har uppnåtts och det inte längre är nödvändigt att behålla uppgifterna ska de raderas. Denna allmänna regel ska anges i det meddelande om behandling av personuppgifter som de registrerade får när de registreras i administrativa datauppsättningar. Med tanke på att det sannolikt är omöjligt att informera alla registrerade personligen finns det dock alternativa sätt att tillhandahålla denna information (t.ex. via en informationssida där integritetspolicyn för forskningsverksamheterna anges eller på den förvaltande myndighetens webbplats).

När den kontrafaktiska konsekvensbedömningen har utförts och resultaten har gjorts tillgängliga och aggregerats för statistiska ändamål anses ändamålet vara uppfyllt. Vid denna tidpunkt bör uppgifterna antingen avidentifieras (om de ska användas för andra ändamål i ett senare skede) eller raderas. Lagar eller andra författningar som tillåter lagring kan gälla under en obegränsad tid, men endast enligt uttryckliga rättsliga hänvisningar.

Det andra villkor som ställs i dataskyddsförordningen är att behandlingen ska omfattas av lämpliga skyddsåtgärder. Enligt artikel 32 i förordningen ska den som utför behandlingen (den personuppgiftsansvarige eller personuppgiftsbiträdet) vidta lämpliga tekniska och organisatoriska åtgärder för att säkerställa en säkerhetsnivå som är lämplig i förhållande till risken. Förordningen föreskriver alltså inga standardåtgärder, utan åtgärderna beslutas från fall till fall med särskild hänsyn till de risker som den berörda behandlingen medför.

RÄTTSLIGA SKYLDIGHETER

Information till registrerade

Den personuppgiftsansvariges främsta skyldighet är att informera de registrerade. Denna skyldighet anges i artikel 13 i dataskyddsförordningen och är tillämplig när uppgifterna samlas in från den registrerade. Om så inte är fallet gäller artikel 14. Den registrerade ska informeras om ändamålet med och metoderna för behandlingen, den rättsliga grunden, lagringsperioder och de rättigheter som den registrerade kan utöva. När det gäller kontrafaktiska konsekvensbedömningar ges denna information vid tidpunkten för insamlingen av de uppgifter som ska användas för statistisk forskning och utvärdering, t.ex. övervakning inom ramen för ESF eller datauppsättningar som innehåller uppgifter från arbetslöshetsregister. Om det är praktiskt taget omöjligt att meddela de registrerade kan som sagt webbplatser eller andra allmänna informationsverktyg användas för att informera personer som finns med i register som används för kontrafaktiska konsekvensbedömningar.

Styrning av förhållandet mellan olika enheter som är involverade i behandlingen

Forskningsverksamheter omfattar ofta flera organ och i sådana fall måste förhållandet mellan dem regleras

⁵¹ Pseudonymisering avser behandling av personuppgifter på ett sätt som innebär att personuppgifterna inte längre kan tillskrivas en specifik person utan att kompletterande uppgifter används. Avidentifiering avser behandling av personuppgifter på ett sådant sätt att det är omöjligt att identifiera individer utifrån uppgifterna. En översikt av pseudonymiseringstekniker finns här. <https://www.enisa.europa.eu/publications/pseudonymisation-techniques-and-best-practices>.

genom särskilda arrangemang, som kallas "avtal om behandling av personuppgifter"⁵². Avtalen ingås från fall till fall enligt varje aktörs roll i behandlingen. Sådana förhållanden kan t.ex. omfatta följande aktörer:

- *Personuppgiftsansvarig–personuppgiftsansvarig* föreligger när enheterna samarbetar i genomförandet av ett projekt, dock i olika roller och med olika uppgifter, och där enheternas ändamål med behandlingen är olika.
- *Personuppgiftsansvarig–personuppgiftsbiträde* föreligger när en enhet (den personuppgiftsansvarige) fastställer ändamålen och medlen för behandlingen och använder en annan enhet (personuppgiftsbiträdet) för att utföra viss behandling. I detta fall är den rättsliga grunden artikel 28 i dataskyddsförordningen och de skyldigheter som anges där gäller.
- *Gemensamt personuppgiftsansvariga*: Detta förhållande definieras i artikel 26 i dataskyddsförordningen, som föreskriver följande: "Om två eller fler personuppgiftsansvariga gemensamt fastställer ändamålen med och medlen för behandlingen ska de vara gemensamt personuppgiftsansvariga." I sådana fall ska parterna fastställa sitt respektive ansvar för att följa bestämmelserna på ett öppet sätt, särskilt när det gäller de registrerades rättigheter.

Konsekvensbedömning avseende dataskydd

Om en typ av behandling av personuppgifter sannolikt leder till en hög risk för de registrerades rättigheter och friheter föreskriver artikel 35 i dataskyddsförordningen att den personuppgiftsansvarige före behandlingen ska utföra en bedömning av den planerade behandlingens konsekvenser⁵³. När det gäller kontrafaktiska konsekvensbedömningar är en sådan bedömning endast relevant om särskilda kategorier av registrerade behandlas i stor skala. I sådana fall kan den personuppgiftsansvarige i förvaltningen utföra bedömningen enligt den nationella dataskyddsmyndighetens regler och verktyg.

FRAMTIDSUTSIKTER: DATAFÖRVALTNINGSAKTEN

Det regelverk som beskrivs ovan är det som gäller för närvarande i EU. Syftet med den nya dataskyddsförordningen var ett nå ett genombrott i skyddet av personuppgifter och skapa ett gemensamt system i EU enligt den rådande tekniska utvecklingen och dagens datoriserade samhälle. Andra reformer är på gång som kan få en stor inverkan på uppgiftsinsamlingen under de kommande åren.

Dataförvaltningsakten befinner sig för närvarande i förslagsfasen hos Europeiska kommissionen⁵⁴. Det uttryckliga syftet med dataförvaltningsakten är att främja tillgång till användbara data genom att förbättra förtroendet för dataförmedlare och stärka mekanismerna för datautbyte inom EU. Personuppgifter kommer troligen att omfattas det nya regelverket och personuppgifter för statistiska ändamål kan därför komma att användas i bredare skala, tack vare en ny aktör som benämns "förmedlare av delning av personuppgifter". En sådan åtgärd skulle kunna bidra till att skapa mer gynnsamma förutsättningar för kontrafaktiska konsekvensbedömningar och underlätta tillgång till mycket information i stor skala.

Dataskyddsförordningens regler utesluter inte användning av kontrafaktiska konsekvensbedömningar. För att kontrafaktiska konsekvensbedömningar ska utföras på ett säkert sätt krävs dock vissa förfarandemässiga och operativa steg som alla medlemsländer måste följa. I normalfallet, där övervakningsdata (behandlade personer) kombineras med data från ett offentligt register (kontrollgruppen) och en extern utvärderar utför den kontrafaktiska konsekvensbedömningen, ska följande viktiga steg följas:

1. Den förvaltande myndigheten når en överenskommelse med de förvaltningar som ansvarar för uppgifterna (arbetslöshetsregister, skatteregister osv.) om vilka uppgifter som behövs för att identifiera kontrollgruppen och analysera behandlings- och kontrollgrupperna. Den förvaltande myndigheten kontrollerar att användningen av

Viktiga steg för att förbättra efterlevnaden av den allmänna dataskyddsförordningen

⁵² Så är t.ex. fallet om den personuppgiftsansvarige och de som faktiskt utför forskningen är olika personer. I sådana situationer ska avtalet reglera förhållandena mellan parterna i fråga om tillgången till uppgifterna och utförandet av forskningen.

⁵³ Du hittar mer information i de riktlinjer som har offentliggjorts av Europeiska dataskyddsstyrelsen: <https://ec.europa.eu/newsroom/article29/items/611236>.

⁵⁴ Europeiska kommissionen antog förslaget till dataförvaltningsakt i slutet av 2020. Information om det aktuella läget för dataförvaltningsakten hittar du här: <https://eur-lex.europa.eu/legal-content/SV/HIS/?uri=CELEX:52020PC0767>.

uppgifterna överensstämmer med artikel 6.1 e eller 6.4 och artikel 5.1 b i dataskyddsförordningen, om inte uttryckligt samtycke har inhämtats.

2. Den förvaltande myndigheten ingår överenskommelser med andra enheter (dataägarna och utvärderaren) för att reglera informationsflödena och det ömsesidiga ansvaret enligt dataskyddsförordningens regler. Tjänsteavtalet mellan utvärderaren och den förvaltande myndigheten ska innehålla en särskild bestämmelse om skydd av personuppgifter. När det gäller andra förvaltningar kan ett samförståndsavtal eller särskilda nationella förfaranden reglera dataskyddet i samband med kontrafaktiska konsekvensbedömningar.
3. Samförståndsavtal innebär att den förvaltande myndigheten får data i pseudonymiserad form från dataägaren och överför data för behandling av behandlings- och kontrollgrupperna till utvärderaren i enlighet med dataskyddsförordningen⁵⁵.
4. Personuppgifterna för behandlings- och kontrollgrupperna får endast lagras under den tid som forskningsarbetet pågår och för forskningens ändamål. De säkerhetsregler som föreskrivs i dataskyddsförordningen ska följas. Dessa regler ska följas av alla enheter som deltar i den kontrafaktiska konsekvensbedömningen.

Nationell praxis överensstämmer vanligen med dataskyddsförordningen men varierar mellan medlemsstaterna. Utvärderare rapporterar att nationella dataskyddsregler fortfarande utgör allvarliga hinder för användningen av mikrodata. Detta kan bero på att det tar tid att införa nationella regler och att vänja sig vid den nya dataskyddsförordningen, att förordningen tolkas på olika sätt i de nationella förvaltningarna, överreglering som gör dataskyddsreglerna strängare i vissa länder eller andra feltolkningar. I några länder bör de förvaltande eller de nationella myndigheterna vidta åtgärder enligt dataskyddsförordningen och ESF+-förordningen för att komma till rätta med dessa problem. I rutan nedan beskrivs hur regionen Veneto i Italien löste några av de vanligaste svårigheterna. Ett exempel på god praxis är att kontakta den nationella dataskyddsmyndigheten för att diskutera de föreslagna arrangemangen innan de genomförs.

Ruta 9 Dataskydd och datautbyte

Ett bra exempel på hur man relativt snabbt kan få åtkomst till avidentifierade personuppgifter är den kontrafaktiska konsekvensbedömningen i "pilot- och genomförbarhetsstudien om hållbara och ändamålsenliga resultat för ESF-deltagare med användning av kontrafaktiska konsekvensbedömningar"⁵⁶, som utfördes för GD Syssestänings räkning 2019.

Utvärderingen, som var ett av de fyra fall som ingick i studien, inriktades på ESF-insatser (integrerat program IP 9.i) och utfördes i regionen Veneto 2015–2016 med långtidsarbetslösa som målgrupp. Övervakningsdata för deltagare och typer av åtgärder kombinerades med data om anställningsperioder före och efter ESF-insatserna (*Comunicazioni obbligatorie*). Utvärderingen tog ungefär två månader (från slutet av juli till slutet av september) från det första mötet om datakrav med den förvaltande myndigheten och den institution som hade uppgifterna, det regionala institutet i Veneto Lavoro. Veneto Lavoro ansvarade för

⁵⁵ Om behandlings- och kontrollgrupperna behandlas genom en undersökning måste utvärderaren inhämta de intervjuade personernas samtycke för att få använda uppgifterna för forskningsändamål.

⁵⁶ [Imeri Europa – Ecorys – Institute for Employment Studies, 2019.](#)

avidentifieringen av uppgifterna och alla datauppsättningar kunde enkelt kopplas till varandra⁵⁷ genom en unik gemensam identifierare som tillhandahölls av Veneto Lavoro.

Även om de berörda organisationerna inte direkt beställde utvärderingen beror den positiva erfarenheten på två huvudfaktorer: nära (om än informellt) samarbete mellan den förvaltande myndighetens lokalkontor och det regionala institutet i Veneto Lavoro, vilket var nödvändigt för att lösa dataskyddsfrågor, och en databas (Mercurio) som drivs av Veneto Lavoro och finns tillgänglig som en anonymiserad fil för forskningsändamål⁵⁸.

Mercurio är en statistisk databas som innehåller alla uppgifter som registreras av arbetsförmedlingarna i regionen Veneto avseende anställnings- och arbetslöshetsperioder för befolkningen. Mercurio innehåller dessutom detaljerad information om alla registrerade arbetstagare och företag. Vid tidpunkten för analysen innehöll Mercurio information om över 4 miljoner arbetstagare, 17,6 miljoner anställningshändelser och 4 miljoner arbetslöshetshändelser⁵⁹.

Datauppsättningen uppdateras och rensas med jämna mellanrum, vilket innebar att det var relativt enkelt att matcha den mot den förvaltande myndighetens övervakningsdata. En potentiell begränsning är att den anonymiserade filen inte uppdateras tillräckligt ofta (den senaste tillgängliga versionen uppdaterades i december 2020), eftersom det kräver stora ansträngningar för att säkerställa god datakvalitet⁶⁰. Denna erfarenhet visar att den förvaltande myndigheten när den tar fram utvärderingsplaner redan på förhand bör överväga möjliga former för samarbete med externa organ som förvaltar nödvändiga uppgifter för kontrafaktiska konsekvensbedömningar, lösningar för eventuella rättsliga hinder, sätt att få åtkomst till och hantera administrativa uppgifter och hur de kan tillhandahållas i en avidentifierad form.

2.3.2. Hur identifierar man "behandlingsgruppen"?

En förutsättning för utförandet av kontrafaktiska konsekvensbedömningar är en tydlig definition av vad det innebär att en person har behandlats eller deltagit i insatsen. När det råder en tydlig förståelse av när en individ eller ett företag kan anses ha behandlats är det naturligtvis viktigt att de kan identifieras. I det här kapitlet diskuteras de viktigaste frågorna kring definitionen av behandlings- och kontrollgrupper, medan mer detaljerade förklaringar av gruppernas metodologiska kännetecken ges i kapitel 3.

Att definiera "deltagande" kan verka enkelt vid första anblicken. Det finns dock ett antal frågor som kanske inte är uppenbara direkt, men som är viktiga och därför kräver noggrant övervägande. Anses till exempel praktikanter i ett yrkesutbildningsprogram som lämnar insatsen ha behandlats? Hur många lektioner i en utbildningskurs måste praktikanterna delta i innan de anses vara deltagare? Det finns även förväntade effekter att överväga. Personer som mottar socialförsäkringsförmåner kan till exempel skriva ut sig från det sociala omsorgssystemet innan de deltar i en insats för att undvika aktiveringsåtgärder. Behandlas dessa personer trots att de aldrig fysiskt finner sig till bokade möten vid den offentliga arbetsförmedlingen?

Definition av behandlingsgruppen

När "behandlingsgruppen" definieras måste det göras åtskillnad mellan "avsikt att behandla" och "behandling av de behandlade". Från ett politiskt perspektiv är den avgörande frågan vanligen om det är effekterna av att en person har erbjudits att delta i en insats eller effekterna av det faktiska

Avsikt att behandla eller behandling av de behandlade

⁵⁷ En ytterligare datauppsättning fanns tillgänglig. Den omfattade start- och slutdatum för varje förklaring om "omedelbar tillgång till arbete" som alla behandlade och icke-behandlade individer undertecknade.

⁵⁸ Se <https://www.venetolavoro.it/public-use-file> för en beskrivning av Mercurio och dess åtkomstregler.

⁵⁹ Följande variabler av intresse utvanns från de administrativa uppgifterna: Start- och slutdatum för varje anställnings- och arbetslöshetsperiod. Typ av avtal för varje anställningsperiod (fast anställning, tidsbegränsad anställning, praktikplats). Planerad varaktighet av tidsbegränsade anställningsavtal. Sociodemografisk information (kön, födelsedatum, utbildning, medborgarskap).

⁶⁰ Myndigheten behöver mer aktuell information än uppgifterna i den senaste tillgängliga versionen av Mercurio, vilket kräver extraarbete för att integrera och rensa den mest aktuella informationen om anställningsperioder.

deltagandet som är intressanta. I det förstnämnda fallet kan de personer som erbjuds insatsen välja att delta eller inte delta. I det sistnämnda fallet, där det är behandlingens effekt på de behandlade som är intressant, består behandlingsgruppen endast av de personer som väljer att delta⁶¹.

Vid en första anblick förutsätter beslutsfattare ofta att det intressanta är att fastställa behandlingens nettoeffekter på deltagarna. Vid närmare övervägande kanske dessa frågor inte är så klara. Om det är möjligt att identifiera de personer som erbjuds behandling kan det ur politisk synvinkel vara mer användbart att definiera dem som den ”behandlade” gruppen. Så är i synnerhet fallet om deltagandet i insatsen är frivilligt. Som beslutsfattare kan du alltså välja att koncentrera dig på de personer som erbjuds deltagande i insatsen, och i det fallet blir den relevanta frågan ”vilka blir effekterna när det gäller efterföljande anställning och inkomst för de personer som erbjuds att delta i utbildningsprogrammet”?

Erbjudande om behandling eller faktisk behandling

För att uppskatta effekterna av erbjudandet om behandling utifrån olika resultat måste det vara möjligt att identifiera de personer som får erbjudandet, vilket det ofta kan vara svårt att göra.

Var hittar jag användbara uppgifter?

När frågorna om vem som behandlas och vad som utgör en behandling har klargjorts är det viktigt att överväga hur de personer som behandlas ska identifieras i undersökningen. För att göra detta behöver du en datakälla utifrån vilken de behandlade enheterna, vare sig de är personer eller företag, antingen kan räknas eller bilda ett urval. Sådana uppgifter hämtas vanligen från ESF:s övervakningssystem och – i förekommande fall – från ytterligare dataregister som har upprättats för insatsen i fråga.

Hitta datakällor för behandlade personer

Till följd av övervaknings- och rapporteringskraven för ESF (och ESF+) måste stödmottagande organisationer registrera antal och vissa personliga kännetecken för de personer som mottar tjänster via en insats. När det gäller kontrafaktiska konsekvensbedömningar krävs fler uppgifter i form av mikrodata för dem som har deltagit i uppsatsen. Utvärderarna kommer att begära uppgifter om varje behandlad enhet (personer eller företag) med information om deras huvudsakliga kännetecken (kön, ålder, utbildningsnivå osv.). Dessa uppgifter kan avidentifieras eller pseudonymiseras av integritetsskäl, men när det krävs en undersökning för att genomföra den kontrafaktiska konsekvensbedömningen är det nödvändigt att fastställa huruvida det är tillåtet enligt dataskyddsförordningens krav att överföra identiteten för dessa enheter (namn, telefonnummer osv.), för att ta fram ett urval. Unika identifierare för varje enhet krävs också för att det ska bli enklare att koppla registren till de olika datakällorna.

2.3.3. Faktorer som ska övervägas vid fastställandet av kontrollgrupper

För att uppskatta det kontrafaktiska scenariot är det oftast nödvändigt att identifiera en kontrollgrupp. Valet av kontrollgrupper begränsas vanligen av olika faktorer, till exempel om insatsen är obligatorisk för deltagarna eller ej och om den genomförs över hela linjen i jurisdiktionen eller begränsas till ett visst område eller en viss tidsperiod. Valet av en lämplig kontrollgrupp har tre

Tre aspekter att överväga i definitionen av kontrollgruppen

⁶¹ Om deltagandet i en insats är obligatoriskt finns det i stort sett ingen skillnad mellan dessa två situationer – alla som erbjuds behandling måste delta. I de flesta fall är deltagandet i en insats dock frivilligt (vilket förutsätts i denna vägledning).

aspekter: 1) en analytisk aspekt, 2) en politisk aspekt, och 3) en praktisk aspekt.

Definiera kontrollgruppen från ett analytiskt perspektiv

Syftet med kontrafaktiska konsekvensbedömningar är att erhålla neutrala uppskattningar av hur en insats påverkar ett antal resultat. För att åstadkomma detta krävs uppskattningar av de kontrafaktiska resultaten. Uppskattningar av de kontrafaktiska resultaten fås från kontrollgruppen (se avsnitt 1.1.). Såsom framgår av diagrammen 6 och 7 uppskattas effekterna genom att en uppskattning av det kontrafaktiska resultatet dras av från det observerade resultatet för den "behandlade" gruppen. Den utsträckning i vilken en effekt är neutral beror på hur väl det kontrafaktiska resultat som har räknats fram för kontrollgruppen överensstämmer med resultatet för behandlingsgruppen om den inte hade behandlats, om alla andra faktorer förblir oförändrade.

Om inte randomisering tillämpas måste vara kontrollgruppen likvärdig med behandlingsgruppen i alla viktiga aspekter, både när det gäller observerbara och icke-observerbara faktorer. Så är fallet med alla kvasiexperimentella metoder, vilket i sin tur innebär att användningen av en sådan metod är en nödvändig förutsättning för alla kontrafaktiska konsekvensbedömningar som inte tillämpar randomisering.

Eftersom nästan alla ESF-insatser är a) frivilliga (personerna i målgruppen är inte tvungna att delta i insatsen), och/eller b) är begränsade på något sätt – det kan röra sig om pilotinsatser eller instrument som begränsas till en viss region eller jurisdiktion – har utvärderaren en rad olika enheter att välja mellan som kontroller. Som utvärderare behöver du sålla bland dessa möjliga enheter för att förfina det slutliga valet av kontrollindivider så att de väl motsvarar deltagarna (behandlingsgruppen). I de flesta fall finns det fyra alternativ att välja mellan⁶²:

Alternativ för valet av kontrollgrupper

- **Plats** – personerna i kontrollgruppen liknar de personer som deltar i insatsen, men finns i olika områden i medlemslandet där insatsen inte finns tillgänglig (i förekommande fall). Valet faller ofta på metoden skillnaden-mellan-skillnader om sådana kontrollgrupper och lämpliga uppgifter finns tillgängliga. Personer som bor på olika platser kan ofta ha stora likheter. Sådana grupper har inte haft möjlighet att delta i insatsen eller kanske avböjer erbjudandet om att delta, vilket innebär att en viktig källa till eventuella fel undanröjs. Arbetsmarknadsvillkoren skiljer sig dock mellan olika platser. Tekniker som tillämpar skillnaden-mellan-skillnader på sådana variationer brukar fungera bra, eftersom skillnader i lokala arbetsmarknadsvillkor tenderar att vara rimligt oföränderliga över tid. Om matchning används för att uppskatta effekterna är det dock mindre tillrådligt att hämta kontrollurval från olika lokala arbetsmarknader. Det har visat sig att felfrekvensen vid urvalet av kontrollindivider från olika arbetsmarknader kan vara större än urvalsfelet⁶³.
- **Tid** – kontrollindivider som liknar deltagarna men observeras vid olika tidpunkter, antingen före eller efter insatsen. Sådana kontrollgrupper behövs ofta när en insats är allmän och obligatorisk – med andra ord när alla personer i målgruppen är tvungna att delta och programmet

⁶² Detta avsnitt bygger på [Card, D., Ibarraran, P. och Villa, J.M., 2011.](#)

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

genomförs inom hela jurisdiktionen. Det finns en stor nackdel med kontrollgrupper som är utformade på detta sätt, nämligen att deras resultat kommer att mätas vid en annan tidpunkt än resultatmätningarna för behandlingsgruppen. Detta innebär i sin tur att resultaten är mottagliga för konjunktursvängningar, förändringar av sammansättningen och skiftande makroekonomiska trender, som kan begränsa möjligheterna att identifiera ett opartiskt kontrafaktiskt resultat. Sådana kontrollgrupper bör endast övervägas om variationen i resultaten över tid är begränsad och en samtidig kontrollgrupp inte finns tillgänglig.

- **Stödberättigande** – här väljs kontrollindivider från grupper på samma plats och vid samma tidpunkt, men från grupper av personer som av en eller annan anledning inte var berättigade att delta. Sådana kontrollindivider eftersträvas ofta om insatsen är allmän och deltagandegraden hög eller om deltagandet är obligatoriskt och det finns tydliga regler för stödberättigande. I detta fall kan personer som "nätt och jämnt" är stödberättigade utgöra en källa till kontrollindivider. Målet är att hitta grupper av personer som liknar de behandlade personerna, men för vilka det finns välkända och fasta skäl (som kan kvantifieras i uppgifterna) till att de inte var berättigade till behandling. Möjligheterna att delta i insatser som finansieras enligt ESF+ baseras ofta på andra regler för stödberättigande som är enkla att mäta och inte kan manipuleras (deltagarna har en viss ålder, de har varit arbetslösa under en viss tid osv.). Dessa trösklar kan därför användas som diskriminerande faktorer i valet av kontroller, enligt vilka behandlade individer och kontrollindivider fördelas.
- **Val/medvetenhet** – i grunden är både behandlings- och kontrollgruppen (inte bara behandlingsgruppen) föremål för urvalsförfaranden som baseras på val som motiveras av potentiellt icke-observerbara faktorer⁶⁴. Kontrollindividerna kan väljas bland de personer som var berättigade att delta men inte gjorde det. Fördelen med detta är att sådana personer vanligen hämtas från samma arbetsmarknad som de personer som behandlades. Sådana kontrollindivider bör övervägas om matchning har valts som metod för den kontrafaktiska konsekvensbedömningen och om det finns mycket information som underlag för urvalsbeslutet. I andra situationer, till exempel om metoden skillnaden-mellan-skillnader tillämpas, är kontroller som baseras på val eller medvetenhet mindre attraktiva.

En ytterligare aspekt är värd att framhålla. Om resultatmått före behandlingen finns tillgängliga både för behandlings- och kontrollgrupperna är det viktigt att undersöka trender i resultatmåtteten före insatsen för båda grupperna. Genom att kontrollera antagandet om "gemensamma trender" undviker man problemet med tillfälliga minskningar av sysselsättningsgrad och inkomster före insatsen för en del av de personer som är berättigade till aktiva arbetsmarknadspolitiska åtgärder (annars skulle de inte vara berättigade till stöd, det så kallade *Ashenfelter's Dip*-problemet). Utvärderaren letar efter liknande tidstrender i resultatmåtteten både för behandlings- och kontrollgruppen, så att återhämtningen från visstidsanställningar eller inkomstförluster inte förväxlas med de långsiktiga relativa vinster som kontrafaktiska konsekvensbedömningar syftar till att upptäcka.

*Analysera trender
före insatsen*

⁶⁴ Detta är vad [Card, D., Ibarrran, P. och Villa, J.M., 2011](#) kallar "tvåsidigt urvalsfel".

Att välja en lämplig kontrollgrupp är svårt både tekniskt och metodologiskt sett. När utvärderingsschemat tas fram rekommenderar vi därför att du bekantar dig med de viktigaste koncepten och vidtar åtgärder redan i ett tidigt skede för att identifiera potentiella kontroller. Det är viktigt att du som beställer en utvärdering anlitar experter tidigt under projekteringen för stöd och hjälp.

Vilka är de relevanta politiska övervägandena?

Valet av en lämplig kontrollgrupp är inte bara en teknisk eller analytisk process. De analytiska aspekterna är grundläggande när lämpliga kontrollindivider ska identifieras, men det är också viktigt att kontrollgruppen utgör ett relevant alternativ till insatsen från ett politiskt perspektiv.

Definiera ett alternativ till insatsen

Kontrafaktiska konsekvensbedömningar kan utformas på en rad olika sätt. De kan till jämföra resultaten för en eller flera behandlingsgrupper med en kontrollgrupp som inte mottar behandling, eller jämföra en behandlingsgrupp med en annan, utan en kontrollgrupp som inte har mottagit behandling. Valet av kontrollgrupp beror på vilken typ av jämförelse som är mest relevant i politiskt hänseende, och om det ens är möjligt att hitta en kontrollgrupp som inte har mottagit behandling. Ruta 10 nedan illustrerar ett exempel på en jämförelse mellan två behandlingar utan en "obehandlad" kontrollgrupp – målet är att bedöma om man hellre ska fortsätta med en insats än en annan. Det bör också påpekas att jämförelser mellan program utan fördelen med en kontrollgrupp av personer som inte har behandlats kan leda till tolkningsskillnader (detta diskuteras närmare i ruta 11).

Jämföra behandling och icke-behandling eller jämföra med ett alternativ

Ruta 10 Politiska frågor i samband med ett utbildningsprogram

Tänk dig ett exempel där beslutsfattaren har för avsikt att inleda en ny utbildningsinsats som ska finansieras via ESF – kalla denna insats "A". Anta att medlemslandet redan har ett utbildningsprogram ("B") som riktas mot samma målgrupp men finansieras via nationella medel. I ett sådant fall kan en lämplig politisk fråga vara om sysselsättnings- och inkomstnivåerna för deltagarna i insats A är större än för deltagarna i insats B efter deltagandet. I vidare betydelse, ger insats A bättre valuta för pengarna än insats B? Om inkomsterna är högre för deltagarna i insats A är det uppenbara politiska svaret att avsluta insats B till förmån för insats A om insats A även visar sig vara kostnadseffektiv.

Ruta 11 Tolka nettoeffekter

En studie kan ge som resultat att det inte finns några inkomstskillnader mellan deltagarna i insats A och deltagarna i insats B. Det politiska svaret på den informationen kanske inte är lika självklart om insats B till exempel var mycket effektiv när det gäller personer som inte fick behandling. Detta skulle innebära att båda insatserna var mycket effektiva. I vissa fall kanske det inte finns belägg för att insats B har varit effektiv för de personer som inte fick behandling. Alternativt kan både insats A och insats B vara ineffektiva, även om en insats kan verka mer effektiv än den andra. I situationer där en insats riktas mer mot vissa grupper i befolkningen än andra kan det fortfarande vara nyttigare att hitta en lämplig grupp av obehandlade enheter som jämförelse.

Tänk på att metoden skillnaden-mellan-skillnader inte kan användas för att jämföra flera behandlingar när det saknas en kontrollgrupp som inte mottagit behandling.

Vilka är de praktiska övervägandena i valet av kontrollgrupp?

Förutom analytiska och politiska överväganden finns det även praktiska aspekter i valet av kontrollgrupper som bör beaktas. Det krävs en

lämplig urvalsram för valet av de urvalsenheter (personer eller företag) som ska fungera som kontroller. Urvalsramarna bör dessutom omfatta individuella enheter som överensstämmer med de analytiska och politiska behoven. Det bästa sättet att göra detta varierar mellan utvärderingarna beroende på det särskilda sammanhanget för den insats som testas.

Det finns särskilt två datakällor som ofta används för att identifiera lämpliga kontrollgrupper. Behandlingsgruppen måste vara känd i båda fallen.

Folkbokföringsregister av olika slag kan användas för att hitta kontroller. För en insats för aktiv integrering på arbetsmarknaden som vänder sig till personer i åldern 18–24 år som mottar arbetslöshetsförmåner kan till exempel register över förmånsutbetalningar användas för att identifiera målgruppen. Om behandlingsgruppen är känd och kan matchas mot uppgifterna om förmånsutbetalningar är det således möjligt att hitta dessa 18–24-åringar som inte mottar behandling och därför kan utgöra kontroller. Tänk dig i stället att en insats riktas mot små och medelstora företag. I det här fallet kan nationella företagsregister (i förekommande fall) användas för att definiera målgruppen, medan det är möjligt att hitta potentiella kontrollgrupper genom information om vilka företag som behandlas.

Folkbokföringsregister och bolagsbeskattningsregister

Register över sökande kan användas om insatsen inte genomförs allmänt, t.ex. om inte alla personer som ansöker till ett yrkesutbildningsprogram godkänns (en kontrollgrupp som bygger på val/medvetenhet). Samma sak gäller för det andra exemplet. Alla företag som ansöker om finansiering kommer inte att godkännas, och i vissa fall kan dessa företag användas som kontrollindivider (se den tidigare diskussionen i detta avsnitt om att man bör vara försiktig i valet av kontrollgrupper under dessa omständigheter).

Ansökningsregister

2.3.4. Vilka uppgifter bör ingå i utvärderingsschemat?

Vilka typer av uppgifter behövs och hur kommer de att samlas in?

Som vi har sett kräver kontrafaktiska konsekvensbedömningar vanligen att det finns tillgång till stora mängder mikrodata (i vissa fall kan grupperade data användas, t.ex. regionala uppgifter). Uppgifterna måste samlas in, sammanställas och dokumenteras. Data från olika källor behöver kopplas ihop på grundval av gemensamma identifieringsfält. De måste också kunna lagras och överföras på ett säkert sätt mellan de personer som ansvarar för att förvalta och genomföra den kontrafaktiska konsekvensbedömningen. Eventuella analytiska datauppsättningar konstrueras från dessa datakällor för att underlätta uppskattningen av effekter.

Hantera datakällor

Följande datarelaterade frågor är viktiga att överväga i utformningen av ett utvärderingsschema:

Checklista för datahantering

Tillgång

- Vilka källor kan användas för att inhämta olika typer av grupperade data/mikrodata?

- Finns de nödvändiga individuella uppgifterna tillgängliga? Gäller detta även särskilda kategorier av uppgifter (i förekommande fall)?
- Om en undersökning är nödvändig, har mål- och kontrollgrupperna identifierats på ett sätt som gör det möjligt att följa upp dem genom intervjuer – finns aktuella kontaktuppgifter tillgängliga?

Konsekvens

- Finns det en enda datakälla eller är det nödvändigt att koppla ihop olika datakällor (statistik avseende arbetslöshet, sociala förmåner, socialförsäkringar, företags-/etableringsuppgifter osv.)?
- Är källorna enhetliga? Kan individer identifieras på ett konsekvent sätt mellan källorna?

Tillgänglighet

- Är det möjligt att få tillgång till nationella datakällor om individuella karriärer, inkomster eller sociala förmåner för att jämföra ESF-deltagare med en potentiell kontrollgrupp?

Avtal

- Finns det särskilda avtal om tillgång till uppgifter mellan den förvaltande myndigheten och andra dataägare? Vilka rättsliga eller organisatoriska hinder behöver diskuteras?
- Vem kommer att ansvara för att förhandla om åtkomst och tillstånd om dataanvändning?

Lagring

- Var kommer uppgifterna att lagras? Vilka it-system och it-infrastrukturer kommer att behövas?
- Vilka åtgärder kommer att vidtas för att säkerställa att uppgifterna lagras på ett säkert sätt och att åtkomst endast beviljas de personer som begär att få tillgång till uppgifter för utvärderingen?
- Hur avidentifieras uppgifterna? Är det möjligt att följa individer och relaterade datakällor över tid?

Efterlevnad av dataskydd

- Uppfyller tidigare lösningar de grundläggande bestämmelserna i den allmänna dataskyddsförordningen och ESF+-förordningen? Är eventuella hinder för tillgången till data motiverade på grund av de tillämpliga rättsliga grunderna enligt artikel 6 a, c och e i dataskyddsförordningen och artikel 17 i ESF-förordningen och eventuell användning av avidentifierade uppgifter?

Hur kommer uppgifterna att behandlas?

Kontrafaktiska konsekvensbedömningar kräver ofta mikrodata – dvs. data som innehåller observationer av individuella enheter (privatpersoner eller företag) såväl i behandlings- som kontrollgrupperna (ibland kan grupperade data användas, t.ex. regionala uppgifter eller uppgifter från arbetsförmedlingar). Vi skiljer mellan tre huvudsakliga datatyper: a) Register över personer i behandlings- och kontrollgruppen, b) resultatregister, och c) kontextuella data (data som används för att undersöka om det finns viktiga skillnader mellan behandlings- och kontrollgruppen). Uppgifterna kan komma från samma eller separata källor. Källorna måste

Koppla mikrodata mellan källor

struktureras för att bilda analytiska datauppsättningar (eller analytiska urval), som används för att uppskatta effekterna. Detta innebär ofta att register över privatpersoner eller företag kopplas ihop mellan källorna. För att göra detta krävs antingen identifierare på individuell nivå (t.ex. individuella socialförsäkringsnummer som gör det möjligt att koppla en persons uppgifter (t.ex. skatteuppgifter) till ett deltaganderegister, eller tillräckligt med data för att koppla ihop register mellan olika källor (uppgifter om namn och födelsedatum måste t.ex. finnas tillgängliga i de olika källorna). Det är viktigt att överväga vilka datakällor som kommer att användas för den planerade kontrafaktiska konsekvensbedömningen, men även om det är möjligt att koppla ihop register mellan källorna.

2.3.5. Vilka är de viktigaste hindren i data- och resultatanalysen?

Såsom diskuteras ovan fastställs effekter i kontrafaktiska konsekvensbedömningar ofta genom en jämförelse av resultaten i behandlingsgruppen och kontrollgruppen. Sådana jämförelser är ett grundläggande inslag i kontrafaktiska konsekvensbedömningar. Skillnaden mellan behandlings- och kontrollgrupperna motsvarar insatsens verkan eller nettoeffekt. Exakt hur verkningarna uppskattas beror dock på den valda forskningsmetoden.

När man planerar en kontrafaktisk konsekvensbedömning är det viktigt att överväga om insatsen är tillräckligt omfattande eller sannolikt kan ge effekter som är statistiskt påvisbara. Kontrafaktiska konsekvensbedömningar grundar sig på principen att mål- och kontrollgrupperna är ett urval av hela den befolkning som potentiellt behandlas och används som kontroll. För att ge resultat måste antalet individer i båda grupperna vara statistiskt signifikant.

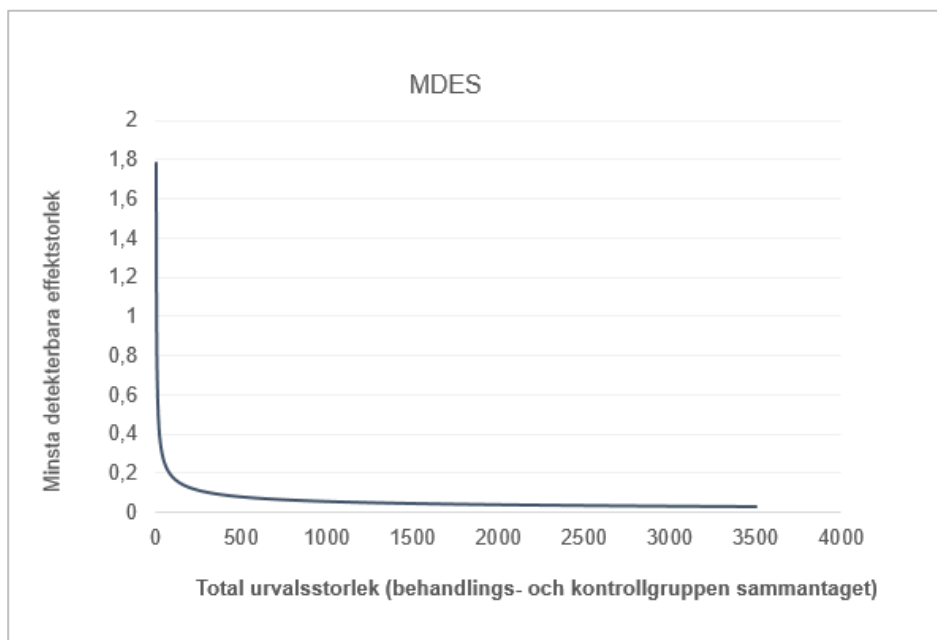
Om du är osäker på om urvalet är tillräckligt stort för analys är begreppet "minsta detekterbara effekt" ett användbart hjälpmedel⁶⁵. Enkelt uttryckt är minsta detekterbara effekt den minsta verkliga effekten som kan upptäckas i ett urval vid standardnivåer för statistisk tillförlitlighet och giltighet⁶⁶.

Bedöma urvalsstorlek och omfattningen av effekter

Diagram 4 nedan visar hur den minsta detekterbara effektstorleken varierar efter den totala urvalsstorleken (totalt antal individer i behandlings- och kontrollgrupperna) vid en nivå på 95 % för statistisk tillförlitlighet och en nivå på 80 % för statistisk giltighet. Om man går från vänster till höger minskar den minsta detekterbara effektstorleken snabbt allteftersom urvalsstorleken närmar sig 500 (250 behandlingenheter och 250 kontroller). Med andra ord blir en kontrafaktisk konsekvensbedömning mer exakt och upptäcker mindre effekter allteftersom urvalsstorleken ökar.

⁶⁵ I [Bloom H. S., 1995](#) ges praktisk vägledning om hur minsta detekterbara effekt kan upptäckas för experimentella metoder. För kvasiexperimentella metoder är det nödvändigt att justera beräkningen. Kvasiexperimentella metoder kräver vanligen större urvalsstorlekar jämfört med experimentella metoder.

⁶⁶ Statistisk tillförlitlighet är ett statistiskt mått på hur tillförlitlig uppskattningen är, medan statistisk giltighet visar sannolikheten för att ett test upptäcker en verklig effekt. Dessa två parametrar är även beroende av urvalsstorleken och kan omvänt användas för att beräkna urvalsstorleken i förhållande till lämpliga signifikansnivåer. Den statistiska tillförlitligheten bör normalt ligga på 95 % och den statistiska giltigheten bör vara 80 %.

Diagram 4 Minsta detekterbara effektstorlekar vid olika urvalsstorlekar

Anmärkning: 95 % statistisk signifikans och 80 % statistisk giltighet. Diagrammet utgår från en randomiserad utformning.

När man planerar en kontrafaktisk konsekvensbedömning är det lämpligt att börja med att definiera behandlings- och kontrollgrupperna. Uppskattningen baseras på uppskattningar av hur många enheter som ska behandlas, hur den metod som ska användas i bedömningen är utformad och storleken på de motsvarande kontrollgrupperna, som bygger på tillgängliga administrativa data eller en lämplig undersökning. Denna information kan sedan, utifrån vissa antaganden, användas för att kontrollera om de resulterande minsta detekterbara effekterna är tillräckligt signifikanta och om en kontrafaktisk konsekvensbedömning är ett lämpligt val. En viktig aspekt i detta sammanhang är om den berörda insatsen kan ge lika stora effekter som de uppskattade minsta detekterbara effekterna. Här kan en litteraturgranskning eller andra liknande utvärderingar vara till hjälp.

Urvalets sannolika storlek

Det är också värt att notera att kontrafaktiska konsekvensbedömningar ofta syftar till att bedöma effekterna på olika delpopulationer i insatsen (män/kvinnor, unga/gamla, personer med högre/lägre utbildning osv.). Dessa analyser leder oundvikligen till att storleken på behandlings- och kontrollgrupperna för varje delpopulation (män, kvinnor, unga osv.) minskar, vilket i sin tur minskar den statistiska signifikansen. Om undersökningen av effekterna på delpopulationerna är en viktig del av utvärderingen måste behandlings- och kontrollgrupperna i varje delpopulation innehålla tillräckligt många individer.

Att fastställa urvalsstorleken kräver teknisk kunskap som inte alltid finns att tillgå hos den förvaltande myndigheten. Om så är fallet bör utvärderaren beräkna urvalsstorleken och sedan diskutera med den förvaltande myndigheten om en kontrafaktisk konsekvensbedömning är genomförbar och i så fall vilken omfattning den bör ha. Det är inte möjligt att fastställa en tröskel över vilken kontrafaktiska konsekvensbedömningar alltid är möjliga, eftersom tröskeln kan variera beroende på vald metod, förväntade analyser av delpopulationer, godtagbara nivåer för statistisk signifikans och andra

faktorer. I många utvärderingar har dock ett antal på 2 000 individer sammanlagt i behandlings- och kontrollgrupperna visat sig vara tillräckligt för en rimligt tillförlitlig analys (inklusive de viktigaste delgrupperna, såsom kvinnor/män osv.). Denna siffra ska inte användas som en vetenskaplig "tröskel", utan är endast avsedd att vara till hjälp i beslutet om huruvida en kontrafaktisk konsekvensbedömning är ett lämpligt alternativ.

Urvalsstorleken påverkar resultatens stabilitet. I rutan nedan diskuteras några exempel på osäkerheter i tolkningen av resultat och hur de hänger samman med urvalsstorleken.

Ruta 12 Osäkerheter i tolkningen av resultat

Bland exemplen på utvärderingar av ESF-finansierade insatser gav utvärderingen av en av fyra åtgärder till förmån för långtidsarbetslösa i regionen Marche i Italien, åtgärden för "utbildningskuponger", inte statistiskt signifikanta resultat. Detta berodde förmodligen på att det analyserade urvalet av behandlade personer var litet jämfört med det stora antalet faktiska deltagare. En liknande situation uppstod i utvärderingen av åtgärder för att främja yrkesutbildning i regionen Podlasien i Polen. Förutom frågan om statistisk signifikans visar detta exempel att alltför små urval av behandlingsgrupper kan leda till problem i effektanalysen i form av generalisering av resultaten.

Utvärdering av språkkurser för invandrare i Tyskland visar positiva resultat i motsats till tidigare studier som bedömt effekterna av språkkursprogram för invandrare. Författaren menar att en möjlig förklaring till skillnaden är att det tyska programmet i fråga även omfattade en yrkeserfarenhetsmodul. Författaren kunde inte testa hypotesen med de uppgifter som fanns till hands eftersom det var omöjligt att skilja mellan språkkursdelen och yrkeserfarenhetsdelen av insatsen. Dessa slutsatser kan vara intressanta för framtida forskning. Exemplet visar att kontrafaktiska konsekvensbedömningar ibland måste upprepas med förbättrade och mer utförliga uppgifter eller kompletteras med andra utvärderingsmetoder för att förklara de mekanismer som avgör resultaten av en insats.

2.3.6. Checklista för att kontrollera om en kontrafaktisk konsekvensbedömning är ett genomförbart alternativ

I det här skedet är det till nytta att sammanfatta de viktigaste faktorer som den förvaltande myndigheten bör överväga och kontrollera när de utarbetar en kontrafaktisk konsekvensbedömning.

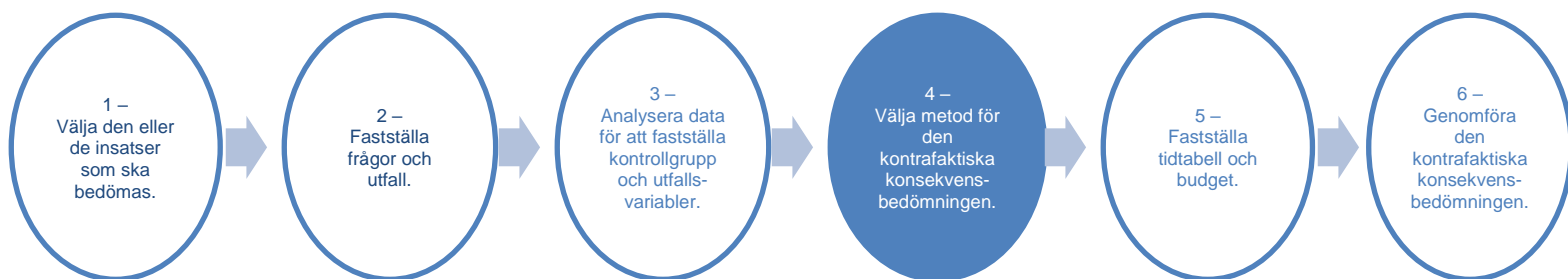
Viktiga faktorer att tänka på

- Är den insats som valts för utvärdering lämpad för en kontrafaktisk konsekvensbedömning?
 - o Är insatsen särskiljande och förhållandevis enhetlig?
 - o Visar insatsens "förändringsteori" att det finns en övertygande kausal mekanism i utfallet som kan undersökas med hjälp av en kontrafaktisk konsekvensbedömning?
 - o Är kontrollgruppen tydligt identifierbar utifrån de regler som gäller för insatsen?
- Kan effekterna på deltagarna mätas kvantitativt?
 - o Mäts insatsens utfall på ett tydligt sätt som överensstämmer med insatsens förändringsteori?
 - o Har det gått tillräckligt lång tid sedan behandlingen avslutats för att det ska vara möjligt att fastställa utfallet för deltagarna?
- Kräver utvärderingsfrågorna en mätning av insatsens genomslag (nettoeffekter)?

- Finns lämpliga uppgifter (uppgifter om behandlings- och kontrollgruppen), utfallsdata, kontextdata) redan tillgängliga eller är det möjligt att få tillgång till dem?
 - Uppfyller de tillgängliga uppgifterna dataskyddsförordningens regler? Om så inte är fallet, kan förvaltningarna ingå avtal och finns det tekniska lösningar för att göra uppgifterna förenliga?
 - Är det möjligt att definiera en lämplig kontrollgrupp utifrån de tillgängliga uppgifterna?
 - Är urvalet tillräckligt stort för att det ska vara statistiskt signifikant inom ramen för resultaten av den kontrafaktiska konsekvensbedömningen?

Om du kan svara ja på alla dessa frågor innebär det att det är möjligt att utforma och genomföra en kontrafaktisk konsekvensbedömning. Om svaret blir nej på några av frågorna innebär inte detta att det är omöjligt att genomföra en kontrafaktisk konsekvensbedömning, men det krävs ytterligare kontroller av genomförbarheten, förbättrad tillgång till data eller andra viktiga faktorer.

2.4. Metoden för den kontrafaktiska konsekvensbedömningen



Att välja metod för den kontrafaktiska konsekvensbedömningen är ett viktigt steg i utarbetandet av bedömningen eftersom den valda metoden påverkar både resultatens kvalitet och kvaliteten överlag. Generellt sett är metoderna likvärdiga, men en metod kan passa bättre mot bakgrund av den typ av uppgifter som finns tillgängliga, medan en annan kan vara lämpligare för att undersöka en viss typ av insats. Att identifiera den lämpligaste utvärderingsmetoden är därför ett viktigt steg i utformningen av den kontrafaktiska konsekvensbedömningen.

*Grundläggande
förutsättningar för
att fastställa den
lämpligaste
metoden*

De tekniska aspekterna och varje metods inneboende egenskaper undersöks i detalj i **kapitel 3**, medan vi här kortfattat diskuterar den praktiska följderna av valet av metod.

Den första metoden är randomisering, där individer slumpmässigt väljs för behandlings- eller kontrollgruppen. Beslut om att använda denna metod måste fattas innan insatsen inleds. Utvärderingen måste alltså utformas efter detta och det måste finnas kompetens för denna metod inom organisationen. Randomiseringsmetoden används sällan i utvärderingar av ESF-insatser. Som vi kommer att se är det dock möjligt att använda denna metod under vissa omständigheter och den är mycket effektiv ur metodologisk synvinkel.

För det andra kanske den förvaltande myndigheten inte har kapacitet för att utföra en kontrafaktisk konsekvensbedömning, eftersom detta kräver

avancerad teknisk kompetens och specifik erfarenhet. I en sådan situation väljer utvärderare som fått uppdraget den lämpligaste metoden. Detta kan ha en ytterligare fördel eftersom det kan användas som ett kriterium för kompetensbedömning när utvärderaren väljs. Om den förvaltande myndigheten kan svara ja på alla eller de flesta frågor i checklisten och har tillgång till nödvändig kompetens – internt eller externt – för att godkänna valet av utvärderare kan den dock ge grönt ljus för utförandet av en kontrafaktisk konsekvensbedömning.

För det tredje kan randomiserade metoder skilja sig från kvasiexperimentella metoder när det gäller varaktighet och kostnader. Om randomiserade kontrollförsök utförs bör utvärderaren kopplas in redan i början av insatsen. Kvasiexperimentella metoder är relativt lika. Utvärderaren bör inleda sitt arbete efter det att insatsen har börjat (men i tid för att organisera och förbereda de nödvändiga uppgifterna). Det är kvaliteten de tillgängliga uppgifterna som avgör tidsgång och kostnader.

Beroende på typ av insats och om det finns tillräckligt med uppgifter tillgängliga kan flera kvasiexperimentella metoder användas. Om så är fallet kan det vara bra att tillämpa flera metoder, eftersom bekräftelse av effekter med användning av olika metoder förbättrar resultatens tillförlitlighet.

2.5. Tidtabell och budget



2.5.1. Vilka resurser finns tillgängliga?

Tillgången till resurser är ett viktigt övervägande i utformningen av ett utvärderingsschema för en kontrafaktisk konsekvensbedömning. Det finns många frågor att överväga, och här delas de upp under tre rubriker: a) expertresurser, b) tid, och c) ekonomiska resurser.

Vilken extern expertis och intern personal krävs för en kontrafaktisk konsekvensbedömning?

Konsekvensbedömningar läggs oftast ut på en extern leverantör. Avtalet måste dock förvaltas inom den förvaltande myndigheten av personal med kunskap om metoder för kontrafaktiska konsekvensbedömningar. Sådan kunskap krävs för att säkerställa kvalitet och för smidiga kontakter med externa experter. Annan sakkunskap kan också krävas inom den förvaltande myndigheten, t.ex. statistisk kompetens och erfarenhet av datainsamling och datahantering. Det är viktigt att den förvaltande myndigheten redan i förväg försäkras om att det finns tillgång till lämpligt kvalificerad och utbildad personal, och att personalen har tid att hjälpa till med utvärderingen.

Intern personal

För att ta fram en ändamålsenlig kontrafaktisk konsekvensbedömning är det viktigt att välja leverantörer som har nödvändig kompetens och erfarenhet av den här typen av utvärderingar. Leverantören måste dessutom vara insatt i den förvaltande myndighetens policy och administrativa förfaranden, känna till potentiella datakällor och besitta nödvändiga språkkunskaper. Det kan vara värt att överväga vad som krävs för att utveckla en leverantörsbas för kontrafaktiska konsekvensbedömningar i ditt land (denna fråga diskuteras närmare i kapitel 4).

Extern personal

Samarbete med de personer som ansvarar för att förvalta programmet eller insatsen är en förutsättning för ändamålsenliga kontrafaktiska konsekvensbedömningar. Leverantören måste till exempel ha tillgång till register som förs av de ansvariga för insatsen. Sådana register ger information om de individer eller företag som deltog i insatsen.

Personal som förvaltar program eller insatser

De personer som ansvarar för insatsen eller programmet kan även vara behjälpliga med råd och vägledning om dessa datatyper. Det kan även behövas andra register för en konsekvensbedömning än de register som normalt upprättas för en insats.

För att lösa problemet med att uppgifter samlas in från olika källor bör du som planerar en kontrafaktisk konsekvensbedömning samarbeta med personal som hanterar officiella datakällor (arbetslöshetsregister, socialförsäkringsuppgifter, statistikkontor osv.). På så sätt kan du planera datainsamlingen i god tid.

Statistisk expertis

Vilka faktorer är viktiga för planeringen av en tidtabell för en kontrafaktisk konsekvensbedömning?

En kontrafaktisk konsekvensbedömning tar en rad olika personalresurser i anspråk och tar dessutom tid att genomföra. Utvärderingsschemat bör innehålla en preliminär tidsplan med viktiga delmål för projektet, som bör gälla både de som arbetar med insatsen och de som utför utvärderingen. Den preliminära tidsplanen bör organiseras med hänsyn till tidsgränserna för både utvärderingen och insatsen. Viktiga politiska delmål bör också beaktas.

Kontrafaktiska konsekvensbedömningar tar tid

Det är en svår balansgång att utforma en meningsfull och realistisk preliminär tidsplan för en kontrafaktisk konsekvensbedömning. Å ena sidan måste den förvaltande myndigheten (eller det förmedlande organet) ta hänsyn till viktiga datum då beslut som är beroende av utvärderingsresultaten ska fattas. Å andra sidan kommer begränsningar oundvikligen att uppstå, vilka kommer att påverka tidpunkten för de rapporter som ska utarbetas. En del resultat blir märkbara först efter flera år, och dessa aspekter bör i möjligaste mån vägas in i datainsamlingen, analysen och rapporteringen (se avsnitt 2.5.2).

Planera in en tidsförskjutning för att beakta senare effekter

Utvärderingen måste genomföras tidigt under programperioden så att ändringar kan göras och erfarenheter och lärdomar kan utnyttjas under den återstående tiden. I vissa fall kan samma eller liknande insatser stödjas under flera programperioder i följd. Resultat från kontrafaktiska konsekvensbedömningar av insatser från tidigare programperioder kan vara till mycket stor hjälp som underlag för genomförandet och

Fokus på specifika tidpunkter

utformningen av konsekvensbedömningar under senare programperioder.

Det är också viktigt att överväga förhållandet mellan tidpunkten för en kontrafaktisk konsekvensbedömning och andra utvärderingar. Teoretiska utvärderingar är ofta nödvändiga som en förberedelse för kontrafaktiska konsekvensbedömningar. När det gäller innovativa insatser (t.ex. ESF-insatserna under programperioden 2014–2020 till stöd för barn i Tjeckien, tjänster för social inkludering för personer som mottar det nya inkomststödet i Italien eller åtgärder för att stödja arbetstagare och studenter som distansarbetade under nedstängningen till följd av covid-19-krisen i många medlemsländer) kan det vara bra att rapportera om viktiga aspekter i processutvärderingen redan i ett tidigt skede för att möjliggöra förbättringar av utformningen av den kontrafaktiska konsekvensbedömningen, eller längre fram som underlag för en mer detaljerad undersökning av orsaker och effekter. Om den kontrafaktiska konsekvensbedömningen rör en mogen och pågående insats är det bättre att processutvärderingen utförs samtidigt som konsekvensbedömningen.

Sekvensera olika typer av utvärderingar

Tidtabellen påverkas även av tillgången till uppgifter. Datakällor kan ta lång tid att uppdatera, vilket ofta är fallet med skatteregister. Det kan också vara både besvärligt och kostnadskrävande att övervinna rättsliga och institutionella hinder för att få tillgång till de nödvändiga uppgifterna. Kontrafaktiska konsekvensbedömningar bör bygga på uppgifter från en rad olika källor. Det är dock både tids- och resurskrävande att kontrollera att uppgifterna går att kombinera, att de håller god kvalitet och att de kan sammanställas i en form som gör det möjligt att uppskatta effekten.

Datainsamling kan vara tidskrävande

Hur bedömer jag kostnaderna?

Det är viktigt att ta fram en preliminär budget som anger hur stora resurser den förvaltande myndigheten kan och vill lägga på en kontrafaktisk konsekvensbedömning. Budgeten består av två delar: interna och externa kostnader. De interna kostnaderna avser de interna resurser som krävs för att följa upp utvärderingen. Dessa kostnader anges vanligen inte i exakta belopp i förvaltningens budget, men är nödvändiga för att säkerställa lämplig kontroll och uppföljning av utvärderingen. Kostnaderna för att anlita de externa experter som ska utföra den kontrafaktiska konsekvensbedömningen måste noggrant uppskattas för att säkerställa utvärderingar av hög kvalitet till en rimlig kostnad. Här tittar vi främst på de externa kostnaderna.

Det bör göras åtskillnad mellan utvärderingar av rutininsatser, där utgifterna vanligen är lägre, och innovativa åtgärder eller pilotåtgärder som kräver insamling av en relativt stor datamängd, användning av nya datakällor och deltagande av nya aktörer som kan motivera de högre utgifterna. Det finns dock inga fasta regler: varje situation bedöms utifrån utvärderingsfrågorna, som kan vara mer eller mindre kostnadskrävande beroende på komplexitet och antal.

Förslag till metod för kostnadsbedömning

Valet av utvärderingsmetod gör också skillnad. Om en randomiserad metod tillämpas krävs som sagt en utvärderare som samlar in och kontrollerar uppgifter medan insatsen pågår. När det gäller kvasiexperimentella metoder beror arbetsinsatsen dock i stor utsträckning på antalet datakällor, deras kvalitet och tillgänglighet.

Kommissionen har tagit fram en vägledning⁶⁷ med uppskattningar av hur mycket en utvärdering av ett program kan kosta. Kommissionen påpekar också att utvärderingar av omfattande program och rutinprogram inte bör uppta mer än 1 % av programbudgeten, medan upp till 10 % kan läggas på utvärderingar av innovativa initiativ eller pilotinitiativ. Resursbehov för kontrafaktiska konsekvensbedömningar behandlas dock inte uttryckligen i denna vägledning. Dessa procentsiffror kan sägas vara kostnadstaket för utvärderingar som omfattar en eller flera kontrafaktiska konsekvensbedömningar.

Det är förstås omöjligt att ange en preliminär kostnad för kontrafaktiska konsekvensbedömningar av så många insatser och operativa förhållanden som ESF+-programmet innehåller för de 27 medlemsländerna. Vi föreslår i stället en metod för att beräkna en rimlig budget för kontrafaktiska konsekvensbedömningar i olika sammanhang. En enkel metod illustreras i följande tabell: raderna anger de viktigaste uppgifterna som ska utföras medan de huvudsakliga kostnaderna anges i kolumnerna.

Antalet arbetsdagar som utvärderingsgruppen måste ägna åt varje aktivitet beror på hur lång tid aktiviteten tar och hur komplicerad den är, vilket i sin tur avgör kostnaden. Kostnaderna för datainsamling i form av kvantitativa undersökningar av deltagarna och personerna i kontrollgruppen är stora. Om den kontrafaktiska konsekvensbedömningen bygger på befintliga administrativa data och uppgifterna är lättillgängliga blir de totala kostnaderna lägre. Administrativa data kräver dock mycket "förberedelser" (att korrigera saknade eller felaktiga uppgifter, justera databasens format, kontrollera administrativa regler för tillgång till uppgifter osv.), vilket kan öka kostnaderna.

⁶⁷ Se [Europeiska kommissionen, 2013](#).

Tabell 3 Struktur för de huvudsakliga kostnaderna för kontrafaktiska konsekvensbedömningar

Huvudsakliga verksamheter	Utvärderingsgrupp			Andra kostnader (Utrustning, materiel, varor, resor osv.)	Anmärkningar
	Projektledare	Ledande experter	Biträdande experter		
Planering och samordning					I planeringen ingår att bedöma om en kontrafaktisk konsekvensbedömning är ett genomförbart alternativ och hur den i så fall ska organiseras. I detta skede fastställs också metoden, i samförstånd med den förvaltande myndigheten. Under denna fas är det viktigt att lägga tid på analysen av eventuella dataluckor. Samordningsfasen omfattar hela varaktigheten för den kontrafaktiska konsekvensbedömningen. Under denna fas organiseras arbetet och kontakterna med den förvaltande myndigheten och andra berörda aktörer för att färdigställa utvärderingsfrågorna.
Litteraturgranskning					Litteraturgranskningen ligger till grund för den preliminära förändringsteorin. Under detta skede fastställs de variabler som ska användas och sammanhanget undersöks. Litteraturgranskningar kan också vara till nytta för att förfina utvärderingsfrågorna, titta på tidigare resultat och formulera hypoteser.
Insamling och bearbetning av data					De resurser som krävs för insamling och bearbetning av data varierar stort beroende på vilken eller vilka datainsamlingsmetoder som används. Den grundläggande datainsamlingen, via undersökningar, kan vara både tidskrävande och kostsam. Andra metoder kan kräva investeringar i teknik (programvara eller maskinvara) eller avtal med dataägare. Kostnaderna kan minskas om dataägarna deltar i förberedelserna.
Dataanalys					Dataanalys kräver avancerad kompetens och erfarenhet. Tidsåtgången för denna aktivitet varierar beroende på metod, antal analyser och datakvalitet.
Utarbetande av rapporter					De resurser som krävs för denna aktivitet varierar beroende på antal och typer av rapporter och andra kommunikationsverktyg. Kostnaderna kan omfatta utskrifter och grafik, utöver den tid som förberedelserna tar.
Uppföljningsmöten					Uppföljningsmöten och andra informationsaktiviteter är viktiga för att sprida resultaten. Övriga kostnader kan omfatta kontorshyra och mat.
Totalt					

Om andra referenser saknas kan personalkostnaden beräknas i förhållande till de kostnader som gäller för ESF+ i fråga om utbildningsinstitutioner eller arbetsförmedlingar i varje land för ledande och biträdande utbildare eller experter. Dessa kostnader är vanligen jämförbara med kostnader för ledande och biträdande forskare, och även om de inte nödvändigtvis är likvärdiga, är de en användbar parameter. En enkel marknadsanalys eller intervjuer med ett antal forskare kan vara till hjälp för att ange den beräknade kostnaden för denna post i den kontrafaktiska konsekvensbedömningen.

2.5.2. När bör insatsen utvärderas?

Det är viktigt att fastställa när det är lämpligast att genomföra konsekvensbedömningen under insatsen, när resultaten bör mätas och verkningarna uppskattas.

När bör nya respektive pågående insatser utvärderas?

Den lämpligaste tidpunkten för att genomföra en kontrafaktisk konsekvensbedömning beror på om det rör sig om en ny insats eller en mogen pågående insats. Om insatsen är ny dröjer det naturligtvis längre innan den mognar och blir stabil. En kontrafaktisk konsekvensbedömning som utförs för tidigt kan ge vilseledande belägg. När det gäller nya insatser är en inledande processutvärdering som utförs före den kontrafaktiska konsekvensbedömningen ofta ett användbart verktyg för att identifiera inkörningsproblem och lösningar på dessa.

Tidpunkten för nya och pågående insatser varierar

Beslutet om när det är lämpligast att utföra en kontrafaktisk konsekvensbedömning styrs av flera faktorer, bland annat följande: Säkerställa att lämpliga datakällor finns tillgängliga, inrätta en intern projektgrupp med utbildad personal och välja extern leverantör. Behoven i fråga om den beslutsprocess som utvärderingen ligger till grund för är en viktig begränsning som man bör ha i åtanke.

Resursfrågor

Valet av tidpunkt för konsekvensbedömningar av pågående insatser styrs främst av praktiska och politiska hänsynstaganden. Insatsen bör redan vara konsoliderad och ha nått en viss mognadsgrad när den kontrafaktiska konsekvensbedömningen genomförs. En ytterligare aspekt som bör övervägas är om andra reformer genomförs parallellt med den utvärderade insatsen. Verkningarna av dessa reformer kan påverka verkningarna av den berörda insatsen. Som beslutsfattare bör du överväga om andra reformer är relevanta för de politiska beslut som resultaten av den kontrafaktiska konsekvensbedömningen är tänkta att ligga till grund för.

ESF-utvärderingar inriktas vanligen på en programperiod. När det gäller stabila insatser som redan ingick i ESF-programmet under den föregående programperioden kan det dock vara värt att kombinera efterhandsutvärderingen av den föregående programperioden med en utvärdering av den löpande programperioden för att täcka ett längre tidsspänn för insatsen.

Flera programperioder

När bör resultaten mätas och verkningarna beräknas?

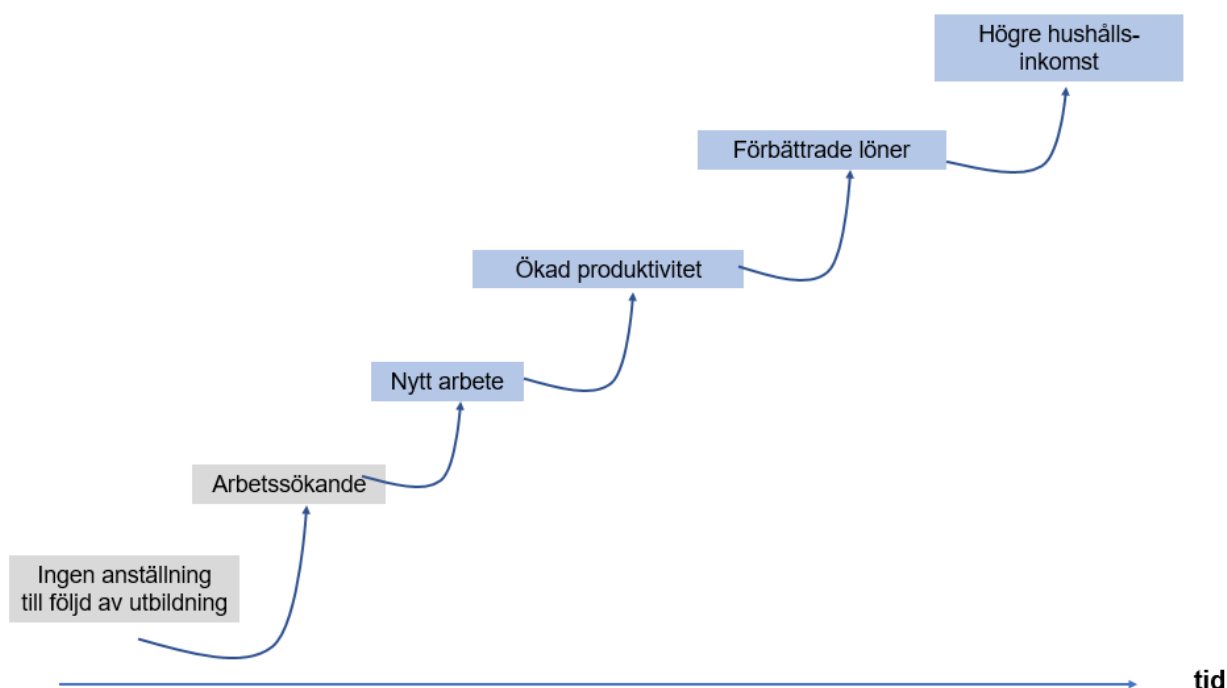
Den andra avgörande frågan när det gäller utvärderingstidpunkten är när verkningarna bör mätas och uppskattas, närmare bestämt när verkningarna av en insats uppskattas bli märkbara.

Om utvärderingen till exempel rör en insats som riktas mot arbetslösa är frågan när högre sysselsättningstal kan förväntas. Det är ett etablerat faktum att utbildningsprogram på kort sikt leder till minskad sysselsättning bland deltagarna. Detta beror på den så kallade inlåsnings-effekten. Utbildningsinsatser brukar leda till att arbetslösa deltagare slutar att söka jobb eftersom de deltar i kurserna. Personerna i kontrollgruppen är dock jobbsökande. Om verkningarna beräknas för tidigt kan de bli negativa eller underskattas. I planeringen av kontrafaktiska konsekvensbedömningar är det därför viktigt att ha en realistisk syn på när verkningarna uppstår och när de blir mätbara. En förenklad modell av verkningar ges i diagram 5 nedan.

Beslutet om när det är lämpligast att mäta resultat och uppskatta verkningar styrs även av beslutsfattarnas behov av information till vissa tidsfrister. När det gäller insatser som syftar till att öka anställbarheten på lång sikt kan det i analytiskt hänseende vara lämpligt att följa upp deltagarna efter två år eller längre efter det att de har exponerats för behandling, för att se om deras inkomster och sysselsättningsnivåer är högre än för en likvärdig grupp av personer som inte har mottagit behandling. Programförvaltare behöver däremot snabba resultat, och här krävs en kompromiss mellan när en uppföljning är rimlig med tanke på förutsättningarna för insatsen och beslutsfattarnas behov av att snabbt få tillgång till belägg.

Avvägning mellan det rimliga och det genomförbara

Figur 5 Förenklad tidsplan för resultat från ett utbildningsprogram



Om resultatmåttarna hämtas från administrativa källor (t.ex. socialförsäkringsregister med uppgifter som anställning och inkomster) är det praktiskt att spåra resultaten upprepade gånger under en längre tid och därefter uppskatta verkningarna (det kan även göras månadsvis). Risken här är att resultatens karaktär kan förändras med tiden. Om primär datainsamling i form av urvalsundersökningar krävs för att mäta resultaten kan det bli mycket kostsamt att uppskatta verkningarna med regelbundna mellanrum om inte historiska uppgifter om resultaten finns tillgängliga. Kostnaden för att hämta

Fokus på specifika tidpunkter

data från flera administrativa system och skapa en enhetlig analysdatauppsättning bör dock inte underskattas.

Som nämns i avsnitt 1.5 kan det vara lämpligt att ta fram en förändringsteori (eller interventionslogik) för att uppskatta resultaten.

Ett alternativ för dig som planerar en kontrafaktisk konsekvensbedömning utan en förändringsteori (men även om det finns en tydlig förändringsteori att bygga på) är att göra en snabb granskning av tidigare studier med utvärderingar av insatser som liknar den insats som du ska utvärdera. En noggrann granskning av resultat från tidigare studier kan ge en bra fingervisning om hur resultaten kan mätas och verkningarna beräknas.

Granskningar av andra studier som nyligen utförts

2.6. Genomförande av kontrafaktiska konsekvensbedömningar



När tidtabellen och budgeten har fastställts är planeringen avslutad och det är dags att börja genomföra den kontrafaktiska konsekvensbedömningen. Det sista steget i denna vägledning syftar till att beskriva de viktigaste stegen i utvärderingen och hur förvaltande myndigheter kan lösa eventuella problem som uppstår på vägen. Detta kapitel bygger på erfarenheter från hjälpcentralen för utvärdering, som är den kanal som Europeiska kommissionen och generaldirektoratet för sysselsättning använder för att hjälpa förvaltande myndigheter i samtliga medlemsländer. Genomförandet av utvärderingar diskuteras allmänt i "EVALSED: *The resource for the evaluation of Socio-Economic Development*"⁶⁸. I denna vägledning behandlas fyra verksamheter som är särskilt relevanta för kontrafaktiska konsekvensbedömningar:

1. Val av utvärderare.
2. Övervaka genomförandet av den kontrafaktiska konsekvensbedömningen.
3. Rapportering.
4. Användning av resultat.

2.6.1. Välja utvärderare

Det första beslutet i valet av utvärderare är om man ska använda en intern eller extern utvärderare. I det senare fallet väljs utvärderaren genom ett offentligt anbudsförfarande.

Utvärderarens oberoende ställning

Vikten av oberoende utvärderingar bekräftas i artikel 44.3 i förordningen om gemensamma bestämmelser, där följande anges: "Utvärderingarna

⁶⁸ Finns på https://ec.europa.eu/regional_policy/sources/docgener/evaluation/guide/guide_evalsed.pdf.

ska anförtros interna eller externa experter som är funktionellt oberoende.” Det första kriteriet är därför att utvärderaren ska vara fullständigt oberoende i förhållande till den förvaltande myndigheten och ESF+-programmets beslutsprocesser. Det andra kriteriet rör frågan om huruvida den kompetens och de resurser som krävs för att utföra en kontrafaktisk konsekvensbedömning finns internt tillgängliga. Detta brukar vara det vanligaste hindret för interna utvärderingar, eftersom kontrafaktiska konsekvensbedömningar kräver särskild teknisk kompetens som sällan finns tillgänglig hos förvaltningarna.

När beslutet om att använda en intern eller en extern utvärderare är fattat är nästa steg att fastställa de tekniska specifikationerna för utvärderingen. Om förvaltningen har beslutat att anlita en extern utvärderare ingår de tekniska specifikationerna i det offentliga anbudsförfarandet. När det gäller specifikationer för kontrafaktiska konsekvensbedömningar bör följande aspekter övervägas särskilt noggrant.

*Centrala delar i
specifikationerna*

- **Utvärderingens syfte** och forskningsfrågor bör ha en tydlig och konsekvent koppling till den kontrafaktiska konsekvensbedömningen.
- **Tillgängliga uppgifter** ska anges. Detta är en av de viktigaste frågorna som måste lösas i utformningen och genomförandet av kontrafaktiska konsekvensbedömningar. På så sätt blir utvärderarna insatta i eventuella problem med datatillgången och kan ta fram realistiska och exakta förslag.
- **Begäran om en kontrafaktisk konsekvensbedömning** kan vara exakt när det gäller specifikationerna, men valet av vilka metoder som ska användas bör lämnas åt utvärderaren. Det ger utvärderaren frihet att komma med lösningar och visa sin kapacitet.
- **Utvärderingsgruppen** bör besitta kompetens inom flera områden, särskilt utvärderingar i allmänhet och kontrafaktiska konsekvensbedömningar i synnerhet (hur man utformar och genomför utvärderingar), ekonometri (använda rätt metoder) och det politiska område inom vilket den utvärderade insatsen genomförs (rekonstruera förändringsteorin och tolka resultat).
- **Tilldelningskriterierna** ska prioritera kvalitet före kostnaden för förslaget. Utvärderingens kvalitet beror i hög grad på utvärderarens förmåga att utforma och genomföra den. De rättsliga ramarna för anbudsförfaranden passar vanligen inte perfekt för de behov som upphandlare av utvärderingar har, men beställaren bör så långt som möjligt anpassa de befintliga kraven för att belöna kvalitet och teknisk kapacitet.
- Under **urvalsförfarandet** bör experter på utvärderingar och kontrafaktiska konsekvensbedömningar sitta med i urvalskommittén för en heltäckande bedömning av kvaliteten på förslagen och de olika metodologiska och organisatoriska lösningarna.

I en del medlemsländer saknas erfarenhet av kontrafaktiska konsekvensbedömningar. Detta kan ge upphov till problem i urvalsförfarandet eftersom marknaden för denna tjänst är mycket begränsad. Förmodligen är det mycket få företag eller experter som uppfyller de tekniska villkoren i specifikationerna för att kunna lägga fram förslag av god kvalitet. I sådana fall kan det vara till hjälp att den förvaltande myndigheten håller öppna seminarier med intresserade

akademiker och företag för att utarbeta den kontrafaktiska konsekvensbedömningen och förbereda ansökningsomgången. Under seminarierna kan förslag från nationella experter samlas in, samtidigt som de får tid på sig att organisera sitt deltagande i anbudsförfarandet. Seminarierna bör vara offentliga för att uppfylla insynskraven och undvika intressekonflikter. Grundläggande information bör finnas tillgänglig för dem som inte hade möjlighet att delta.

2.6.2. Övervakning

När utvärderaren har valts och den kontrafaktiska konsekvensbedömningen har inletts är det dags för den förvaltande myndigheten att starta övervakningen av utvärderingsprocessen. Övervakningen handlar främst om att kontrollera och validera utvärderarens resultat, organisera statusmöten med utvärderaren och berörda aktörer och vid behov ingå avtal med dataägare.

Den förvaltande myndigheten måste ha kapacitet för att utföra dessa uppgifter. Detta innebär även att myndigheten måste ha resurser för att övervaka hela processen och kompetens för att kontrollera de tekniska lösningar som används i den kontrafaktiska konsekvensbedömningen. De förvaltande myndigheterna brukar avsätta särskild personal för utvärderingarna. Om myndigheten saknar erfarenhet av denna typ av bedömningar bör oberoende externa experter anlitas för råd och hjälp. Sådana experter kan anlitas via korttidskontrakt. Om de kommer från offentliga institutioner är tjänsterna kostnadsfria. Uppdraget brukar ta mellan sex och tolv arbetsdagar, vilket bör vara tillräckligt för att granska rapporter och delta i möten med den förvaltande myndigheten, utvärderaren och de viktigaste berörda aktörerna.

Intern och extern kompetens i övervakningsprocessen

De förvaltande myndigheterna inrättar ofta styrgrupper för att övervaka programutvärderingar. Styrgrupperna består vanligen av företrädare för den förvaltande myndigheten, tjänstemän från andra genomförandeorgan, de viktigaste berörda aktörerna (arbetsmarknadens parter, icke-statliga organisationer, andra förvaltningar) och experter från den akademiska världen och/eller offentliga institutioner (t.ex. det nationella statistikinstitutet). Styrgruppen deltar i arbetet med att ta fram utvärderingsfrågorna och diskuterar utvärderingsresultaten och hur de ska spridas och användas i beslutsprocessen. En mer flexibel och specialiserad underordnad grupp kan också inrättas för att övervaka resultaten inom den förvaltande myndigheten.

2.6.3. Rapportering

Rapportering är ett viktigt inslag i utvärderingen och den huvudsakliga kanalen för att informera om resultaten. Utvärderingsrapporter ska vara tydliga, kortfattade, förståeliga för lekmän och transparenta i sina bedömningar och politiska rekommendationer. För att visa att utvärderingen är tillförlitlig ska den metod och de uppgifter som används redovisas i rapporten. Eventuella begränsningar av analysen ska också anges.

Vanligen utarbetas tre rapporter för varje kontrafaktisk konsekvensbedömning.

- En **inledande rapport** tas vanligen fram kort efter det att avtalet har undertecknats. I rapporten redovisas den slutliga metod som har

Rapporter om framsteg och användning av resultaten

fastställts under diskussionerna med den förvaltande myndigheten samt den inledande kontrollen av tillgängliga uppgifter.

- En **interimsrapport** tas fram efter halva utvärderingen. När det gäller kontrafaktiska konsekvensbedömningar kan interimsrapporten redovisa vilka uppgifter som har samlats in, behandlings- och kontrollgruppens sammansättning samt urvalsprocessen.
- I den **slutliga rapporten** redovisas analyser och resultat av den kontrafaktiska konsekvensbedömningen. Den innehåller uppgifter om metod och data och eventuella politiska följder som framgår av beläggen diskuteras. Slutliga rapporter om kontrafaktiska konsekvensbedömningar har vissa obligatoriska inslag, nämligen följande: Utvärderingsfrågor. Förändringsteori (eller interventionslogik) och de resultat som ska bedömas. Vald metod. Använda uppgifter och behandlings- och kontrollgruppernas kännetecken. Uppskattade effekter. Svar på utvärderingsfrågorna och deras politiska följder.

Beställaren granskar alla rapporter, vid behov med hjälp av externa experter, och synpunkterna diskuteras med utvärderaren, som därefter gör nödvändiga förtydliganden och förbättringar innan rapporten godkänns formellt.

Rapporter om kontrafaktiska konsekvensbedömningar får inte vara alltför tekniska, och resultaten ska vara tydliga och förståeliga för alla aktörer och lekmän. De mer tekniska stegen i analysen kan redovisas i bilagor, så att huvudrapporten skrivs på ett enkelt och flytande språk. De metodologiska detaljerna är dock viktiga för att visa att den kontrafaktiska konsekvensbedömningen är tillförlitlig, vilket innebär att det krävs transparens även i det avseendet. Rapporten ska innehålla uppgifter om utfallsvariabler, datakvalitet, behandlings- och kontrollgruppens sammansättning och urvalsstorlek, metoden för att uppskatta effekterna och de statistiska försök som tillämpats. Enligt akademisk praxis ska dessutom de avidentifierade uppgifter och databehandlingsmetoder som använts göras tillgängliga för forskare som vill upprepa och kontrollera analyserna.

Den slutliga rapporten om en kontrafaktisk konsekvensbedömning ska innehålla vissa standardiserade minimiuppgifter. Detta har ett dubbelt syfte: 1) underlätta jämförelser av resultat mellan olika operativa program och länder, 2) samla in information för framtida metaanalyser.

Minimiuppgifter som ska ingå i den slutliga rapporten om en kontrafaktisk konsekvensbedömning

Mallen för att samla in dessa standardiserade minimiuppgifter om varje ESF-utvärdering följer rekommendationerna i pilot- och genomförbarhetsstudien [...]. Denna information kan tas med i en bilaga till den slutliga rapporten och bör även skickas till kommissionen. Ett förslag till mall för att samla in de minimiuppgifter som krävs presenteras i

Tabell 4 nedan.

Tabell 4 Grundläggande information som ska ingå i ett informationsblad om den kontrafaktiska konsekvensbedömningen

Informationskategori	Specifika uppgifter som ska samlas in och rapporteras
1. Information om insatsen	<ul style="list-style-type: none"> - Den utvärderade insatsens benämning. - Insatsens kännetecken (utbildning, sysselsättningstjänster, tjänster för social inkludering osv.). - ESF+-programmets inriktning (prioriterade områden, särskilt mål, åtgärder). - Förvaltande myndighet och/eller andra genomförandeorgan - Utgifter - Varaktighet för och omfattning av den behandling som ges via insatsen. - Deltagare, med uppgift om ålder, kön och stödberättigande status (personer som mottar arbetslöshetsbidrag, långtidsarbetslösa, unga som varken arbetar eller studerar). - Insatsens territoriella omfattning.
2. Mått på insatsens ändamålsenlighet	<ul style="list-style-type: none"> - Uppgift om huruvida den kontrafaktiska konsekvensbedömningen uppskattar en positiv och statistiskt signifikant behandlingseffekt i fråga om de relevanta utfallen. - Faktisk omfattning av den effekten. - De två ovanstående måtten (+/-) och (storlek) för uppsättningen relevanta och jämförbara utfall, sysselsättningsgrad, inkomster osv. - Indikatorer på valuta för pengarna (om detta ingår i uppskattningen).
3. Data och metod	<ul style="list-style-type: none"> - Datakälla (undersökning, administrativa data). - Observationsperiod. - Ekonometrisk metod för den kontrafaktiska konsekvensbedömningen. - Tidshorisont (vanligen kort < =12 månader från det att insatsen avslutats, medellång >12 och < =24 månader, långsiktiga effekter >24 månader). - Urvalsstorlek för behandlings- och kontrollgrupper. - Tillgång till/användning av uppgifter om situationen före programmet.

2.6.4. Använda resultaten

Användningen av resultaten är en viktig aspekt. Om resultaten inte sprids effektivt och når den avsedda målgruppen har utvärderingen en begränsad inverkan.

Spridning av resultaten av kontrafaktiska konsekvensbedömningar är en central aspekt

Spridning av resultat och output av utvärderingen omfattar vanligen följande:

- Minst en skriftlig utvärderingsrapport med ett sammandrag och en sammanfattning.
- Minst en muntlig presentation av resultaten, med hjälp av bildpresentationer och liknande verktyg.
- Ett tekniskt avsnitt (eller bilaga) i rapporten med en utförlig redogörelse för metod, huvudantaganden och strategi för statistiska analyser.

Alla utvärderingsrapporter ska offentliggöras. Detta är ett krav enligt förordningen om gemensamma bestämmelser för programperioden 2021–2027⁶⁹. Utvärderingsresultaten ska även läggas fram för och diskuteras i programmets övervakningskommitté. Detta är obligatoriskt, men är i sig inte tillräckligt för att uppmuntra till djupgående diskussioner eftersom operativa prioriteringar oftast är det viktigaste ämnet för

⁶⁹ Artikel 44.7 i förordningen om gemensamma bestämmelser.

utvärderingskommitténs möten. I rutan nedan ges ett exempel på nationell praxis för presentation och spridning av utvärderingsresultat.

Det finns ingen färdig lösning för att effektivt sprida resultat och diskutera lärdomar, utan det beror på sammanhanget. Strategin måste vara ändamålsenlig och det är viktigt att se till att aktörer utanför den förvaltande myndigheten informeras om resultaten. En lösning är att dela upp målgruppen i huvudgrupper och därefter besluta vilka kommunikationskanaler som ska användas. Målgruppen kan till exempel delas upp på följande sätt:

Många målgrupper och kommunikationskanaler

- Politiskt ansvariga och politiska företrädare på hög nivå – de arbetar oftast med kortfattade dokument (policydokument, sammanfattningar) med de viktigaste resultaten och rekommendationerna.
- Tjänstemän vid de förvaltningar som deltar i genomförandet av insatsen eller liknande politiska åtgärder. Här kan seminarier eller workshoppar kan vara effektiva instrument för att presentera resultaten av kontrafaktiska konsekvensbedömningar. Den tekniska nivån på presentationen bör anpassa till deltagarna.
- Politiska aktörer (arbetsmarknadens parter, icke-statliga organisationer, stödmottagare osv.). Sammanfattningar av rapporterna och årliga utvärderingsmöten kan vara en bra kombination för att involvera dessa aktörer i diskussionen. Om politiska aktörer deltagit i arbetet med att ta fram utvärderingsfrågorna bör utvärderingsresultaten kopplas till dessa frågor.
- Experter och akademiker – vetenskapliga konferenser eller seminarier kan vara de lämpligaste mötesplatserna för att presentera utvärderingsresultaten. Dessa faktorer validerar resultaten av utvärderingen ur en vetenskaplig synvinkel och ökar resultatens tillförlitlighet.

Metautvärderingar⁷⁰ är ett kraftfullt kompletterande verktyg för att sprida resultaten av kontrafaktiska konsekvensbedömningar. De gör det möjligt att bekräfta och generalisera de huvudsakliga resultaten, vilket inte kan göras med individuella kontrafaktiska konsekvensbedömningar. För att utföra en metautvärdering behöver du ha tillgång till ett antal tillförlitliga kontrafaktiska konsekvensbedömningar av god kvalitet om jämförbara politiska åtgärder. Metautvärderingar är fortfarande sällsynta och inriktas inte enbart på ESF-insatser. För att införa metautvärderingar över hela linjen krävs en mer utbredd användning av kontrafaktiska konsekvensbedömningar. Metautvärderingar bör dessutom planeras in på förhand för att underlätta jämförelser. Under programperioden 2021–2027 kan metautvärderingar inom ramen för ESF+ genomföras i länder som har många program eller via jämförelser mellan länder.

Metautvärderingar som verktyg för att öka spridningen och generaliseringen

⁷⁰ En metautvärdering är en systematisk jämförelse, eventuellt med stöd av statistiska verktyg, av flera kontrafaktiska konsekvensbedömningar av liknande insatser. Metautvärderingar ger indikationer på en insats ändamålsenlighet genom att undersöka effekterna av liknande åtgärder i olika sammanhang. Varje insats i jämförelsen undersöks dock inte i detalj och metautvärderingar ger allmänna politiska riktlinjer, inte för varje insats.

Ruta 13 Den polska erfarenheten av utvärderingskonferenser

Den internationella utvärderingskonferensen, som anordnas av ministeriet för utvecklingsfonder och regionalpolitik och den polska myndigheten för företagsutveckling, är ett av de viktigaste inslagen i arbetet med utvärderingar av offentliga insatser i Polen. Med tiden har konferensen blivit en etablerad plattform för diskussioner mellan politiskt ansvariga, aktörer som ansvarar för genomförandet av politiken och aktörer som arbetar med konsekvensbedömningar.

Konferensen hölls första gången 2005 och har sedan dess genomförts årligen i tio år och vartannat år sedan 2015. Konferensen 2021 var den fjortonde i ordningen. Den organiseras för närvarande som ett tvådagarsevenemang. En dag ägnas åt forskningsresultat och insatsernas effekter med politiska diskussioner. Den andra dagen inriktas på metodologiska frågor och utvärdering som en process.

Under de fjorton konferenser som hållits har kontrafaktiska metoder varit föremål för paneldiskussioner eller expertpresentationer vid ett flertal tillfällen och har behandlats ur olika perspektiv under båda dagarna. Vid 2017 års konferens presenterades till exempel uppskattade effekter under sessioner om insatsernas effekter, medan diskussioner om fördelar och nackdelar med olika uppskattningstekniker hölls separat.

Konferensen är också speciell på grund av den blandade publiken. Den vänder sig inte bara till forskare eller specialister på metodologiska frågor, utan även till dem som utformar och genomför insatser. Evenemanget har två mål: Den bidrar till att bygga upp utvärderingssystemets kapacitet genom att utveckla deltagarnas kompetens och kunskap och ger möjlighet till nätverksbyggande genom att fungera som en mötesplats för erfarenhetsutbyte mellan nationella, regionala och utländska partnerorganisationer. Varje konferens samlar cirka 300–400 företrädare från flera olika sektorer – offentliga förvaltningar, den akademiska världen, konsultföretag och icke-statliga organisationer. Det är ett stort deltagarantal med tanke på att evidensbaserade beslutsprocesser ännu inte helt har slagit rot i Polen.

Den internationella utvärderingskonferensen ingår i den offentliga förvaltningens utvärderingssystem. Kapacitetsuppbyggnadsarbetet blev framgångsrikt tack vare utbildningar som tillhandahölls av gemensamma forskningscentrumet och Världsbanken eller via finansiering från budgeten för tekniskt stöd i de operativa programmen. Utbildningar om utvärdering, som även omfattade den kontrafaktiska metoden, fanns tillgängliga såväl för personalen vid de förvaltande myndigheternas utvärderingsenheter som för andra institutioner, t.ex. lokala eller regionala arbetsförmedlingar. Konferensen utgör ett tillfälle att samlas, diskutera och utbyta erfarenheter om den senaste utvecklingen och nya rön⁷¹.

⁷¹ Information om konferensen finns här: <https://www.ewaluacja.gov.pl/strony/xiv-miedzynarodowa-konferencja-ewaluacyjna/>. Vi vill tacka Piotr Strzęboszewski och andra tjänstemän vid ministeriet för utvecklingsfonder och regionalpolitik för den information som de har lämnat om konferenssystemet.

Kapitel 3. Hur väljer jag en lämplig metod för min kontrafaktiska konsekvensbedömning?

I detta kapitel redogör vi för empiriska metoder för kontrafaktiska konsekvensbedömningar. Särskild tonvikt läggs vid den experimentella metoden (randomiserade kontrollerade studier) och de vanligaste kvasiexperimentella metoderna: matchning (utifrån benägenhetspoäng), skillnaden-mellan-skillnader, regressionsdiskontinuitetsdesign och instrumentvariabler. Varje metod använder olika sätt att skapa kontrollgrupper för att besvara den kontrafaktiska frågan: "vad skulle ha hänt med behandlingsgruppen resultatmässigt sett om den inte hade deltagit i/exponerats för insatsen"? I praktiken kan den valda utformningen skraddarsys till det aktuella sammanhanget, vilket avgörs av typ av insats och de typer av data som finns tillgängliga eller kan samlas in⁷².

Det är inte möjligt att ge detaljerad vägledning om vilken utvärderingsmetod som är lämpligast, eftersom de förvaltande myndigheterna verkar under mycket varierande förhållanden. För att välja den mest relevanta metoden för en kontrafaktisk konsekvensbedömning mot bakgrund av den situation som råder i varje fall bör de överväga vilka lösningar som har fungerat bra i tidigare utvärderingar, både inom myndigheten, inom landet och i andra medlemsländer. Det finns mycket att lära om vad som har åstadkommit inom det berörda programmet och på andra håll under liknande omständigheter. Forum för erfarenhetsutbyte om utformning och genomförande av utvärderingar kan vara användbara informationskällor i detta avseende. Granskningar av litteratur om utvärderingar av liknande insatser kan också vara en viktig informationskälla som stöd i utformningen. De experter som den förvaltande myndigheten har anlitat kan också bistå med synpunkter på vilken metod som är lämpligast för utformningen av utvärderingen. Tänk på att det även kan finnas bred sakkunskap och erfarenhet inom myndigheten som du kan utnyttja.

Välja rätt metod

3.1. Randomisering – den experimentella metoden

Viktiga egenskaper hos randomisering

- Individier som är berättigade att delta fördelas slumpmässigt till en behandlingsgrupp eller en kontrollgrupp.
- Randomisering säkerställer att båda grupperna (i genomsnitt) är identiska när det gäller alla relevanta kännetecken.
- Kontrollgruppen besvarar den kontrafaktiska frågan och skillnaden i utfall mellan behandlings- och kontrollgruppen visar insatsens kausala effekt.

Randomiserade utvärderingar kan utformas på olika sätt. Tyngdpunkten ligger här på en enkel metod med två grupper – en behandlingsgrupp och en

⁷² Andra metodiska aspekter beskrivs i [Europeiska kommissionen, 2019](#).

kontrollgrupp – för att klargöra de viktigaste principerna. Figur 1 visar en enkel randomiserad utformning.

Det viktiga är att randomisering säkerställer att de två grupperna är statistiskt likvärdiga i alla avseenden vid tidpunkten för randomiseringen. Därefter exponeras behandlingsgruppen för den berörda insatsen vars genomslag eller effekter ska mätas.

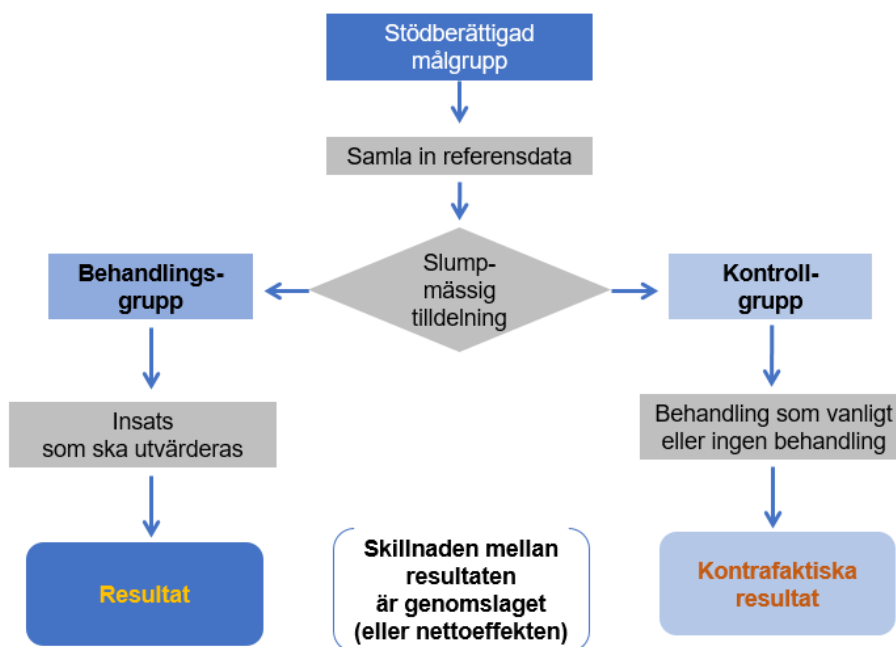
Statistiskt likvärdiga grupper

Det finns olika alternativ att välja mellan beroende på den politiska fråga som står i centrum för utvärderingen. Kontrollgruppen kan väljas som den grupp som inte ska motta någon behandling alls, eller så kan man jämföra behandlingsgruppen med en grupp som mottar en annan behandling av intresse (detta kan sägas motsvara "behandling som vanligt"). Ett ytterligare alternativ är att ha flera behandlingsgrupper som jämförs med kontrollgruppen. Det kan till exempel vara intressant att jämföra effekterna av ett ESF-finansierat utbildningsprogram med annan nationellt finansierad utbildning eller andra tjänster som tillhandahålls till samma grupper.

Ingen behandling eller annan behandling för kontrollgruppen

Eftersom behandlings- och kontrollgrupperna är statistiskt likvärdiga vid randomisering och exponeringen för behandlingar är kontrollerad kan skillnaden i resultaten tillskrivas den utvärderade insatsen (med förbehåll för statistiska osäkerheter). Det innebär att alternativa förklaringar kan uteslutas som orsaker till eventuella observerade skillnader (ett exempel ges i rutan nedan).

Figur 6 Randomiserad kontroll med två grupper



På grund av de särskilda egenskaperna hos randomiserade metoder kan de, om de genomförs korrekt, ge starka belägg för effekterna av en insats. De är därför mycket effektiva i det avseendet. Randomiserade metoder kräver dock detaljerad planering i god tid och kan vara svåra att utforma och hantera. De kan också vara svåra att genomföra korrekt för programförvaltarna. En randomiserad process kan till exempel leda till en förändring av

Randomisering ger starka belägg ...

sammansättningen av deltagare i insatsen, eftersom potentiella deltagare kan avskräckas av att randomisering kommer att tillämpas och inte vilja delta av det skälet. Personer som är föremål för randomisering kanske inte alltid motsvarar den status som de tilldelas, och det kan uppstå en rad andra problem som måste lösas. Randomiserade kontrollerade studier kan dessutom vara kostsamma att genomföra under vissa omständigheter.

Av dessa och andra skäl är det inte särskilt sannolikt att randomiserade metoder kommer att användas för utvärderingar av ESF-finansierade instrument och insatser. Vi avråder dock från att randomisering utesluts utan vidare övervägande. Randomisering har använts ofta och det finns fler exempel förutom det exempel från Förenade kungariket som diskuteras i följande ruta, bland annat GAI-experimenten från USA⁷³ (det finns många andra exempel från Nordamerika), experiment i Sverige⁷⁴ rörande ESF-insatser under programperioden 2014–2020 (se nästa ruta) och en studie från Tyskland för att bedöma effekterna av aktiva arbetsmarknadstjänster som tillhandahölls av privata leverantörer jämfört med de tjänster som tillhandahölls via de offentliga arbetsförmedlingarna⁷⁵, och många fler.

... men är svåra att utforma

Ruta 14 Exempel på en randomiserad prövning av ett ESF-projekt för ungdomar

Utvärderingen av projektet "Ung framtid" i Sverige

Det svenska projektet "Ung framtid" är ett ESF-finansierat projekt som genomfördes mellan 2015 och 2018. Projektet syftade till att hjälpa unga arbetssökande i åldern 18–24 år att etablera sig på arbetsmarknaden i mellersta Norrland, Norra Mellansverige och södra Sverige. Projektet bestod av intensifierade och individuellt anpassade stödåtgärder som riktade sig till unga arbetssökande och genomfördes av Arbetsförmedlingen i Sverige. Arbetsförmedlingen tillhandahöll individuell studie- och yrkesvägledning, information, vägledning och konkret stöd i form av matchning, utbildning och andra verksamheter. Nästan 17 000 unga deltog i projektet.

Svenska ESF-rådet gav Arbetsförmedlingen i uppdrag att genomföra utvärderingen som en del av den utvärderingsram som utformats inom ramen för ett ESF-projekt under programperioden 2014–2020. Målet var att förbättra Arbetsförmedlingens utvärderingskapacitet (evidensbaserat EU-fondsprojekt 2014–2020, se rutan i det sista kapitlet i vägledningen).

I ett försök som genomfördes mellan juni 2017 och januari 2018 med åtta lokala arbetsförmedlingskontor av de totalt 90 kontor som deltog i projektet fördelades stödberättigade ungdomar slumpmässigt till behandlingsgruppen (personer som mottog intensifierat stöd) och kontrollgrupper (personer som mottog normalt stöd)⁷⁶. Försöket omfattade 4 689 ungdomar som slumpmässigt fördelades till behandlings- och kontrollgrupperna, medan 2 972 fördelades till kontrollgruppen. Den randomiserade fördelningen gav behandlings- och kontrollgrupper som var mycket lika vid tidpunkten för randomiseringen. Detta innebar att eventuella skillnader i utfallsvariablerna mellan de två grupperna (återvändande till arbetsmarknaden, andel arbetslösa och genomsnittligt antal dagar i arbetslöshet) efter det att deltagarna gick in i ESF-insatsen kan tillskrivas verksamheterna inom projektet Ung framtid.

Studiens resultat visar att projektet Ung framtid hade en positiv inverkan på kvinnor, men inte på män. För kvinnor påvisades positiva effekter i alla bedömda utfallsvariabler. Andelen kvinnor som återvände till arbetsmarknaden var omkring 7 procentandelar högre under de första två månaderna efter det att de gått in i projektet, andelen arbetslösa minskade mer i behandlingsgruppen än i kontrollgruppen och behandlade kvinnor var arbetslösa under kortare tid än kvinnorna i kontrollgruppen⁷⁷.

⁷³ Se Riccio J., Friedlander, D., Freedman S., 1994.

⁷⁴ Se Hägglund, P., 2006.

⁷⁵ Se Krug G., Stephan G., 2011.

⁷⁶ En pilotfas genomfördes mellan april och maj 2017 för att testa ett antal verksamheter, t.ex. hantering av randomisering, unga deltagares engagemang och rapportering.

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

Randomiserade metoder kan främst särskiljas från andra metoder genom den starka betoningen på att kontrollera potentiella fel (bias) mellan behandlings- och kontrollgrupperna genom forskningsdesign. Betoningen på forskning gör metoden mycket intuitiv, men den kräver förhandsplanering. Randomiserade metoder lämpar sig ofta bättre för utvärderingar av nya pilotinsatser än befintliga insatser. Detta beror på att de kräver en viss grad av kontroll över hur deltagarna rekryteras till den utvärderade insatsen. Denna "kontroll" är ofta svårare att åstadkomma i befintliga program jämfört med nya insatser som är öppna för nya idéer.

Randomisering genom forskningsdesign

Att genomföra en randomiserad kontrollerad design innebär att en del av den stödberättigade målgruppen inte erbjuds deltagande utan fördelas till en kontrollgrupp. Fördelningen är helt slumpmässig och styrs varken av de sökande eller insatsförvaltarna. Beslutsfattare brukar därför motsätta sig randomiserade kontrollerade studier av etiska skäl, utan att överväga om de är genomförbara i praktiskt och analytiskt hänseende.

Det finns etiska invändningar ...

Det finns dock goda skäl att använda randomiserade metoder. Om randomisering ger den bästa kvaliteten och de mest tillförlitliga beläggen för hur ändamålsenliga offentliga insatser bör den metoden väljas för att bedöma effekterna av en insats. Om en viss insats inverkan är okänd på förhand är det inte oetiskt att utesluta individer, eftersom det inte kan förutsättas att de skulle ha vunnit på att delta. Randomiserade metoder användes dessutom utbrett inom medicin och andra studiefält, till exempel utbildningsforskning. Under vissa omständigheter där de tjänster och stödåtgärder som tillhandahålls via en insats är "överbokade" (dvs. att det finns fler stödberättigade individer än det antal individer som faktiskt kan delta i utvärderingen) och resurserna är knappa, kan slumpmässig fördelning av individer till insatsen från gruppen av berättigade deltagare dessutom vara det mest etiska sättet att fördela resurserna på.

... men också starka argument för användning av randomiserade metoder

3.2. Icke-randomiserade eller kvasiexperimentella metoder

Om randomisering inte är en genomförbar lösning finns det ett antal alternativa metoder i verktygslådan för kontrafaktiska konsekvensbedömningar med olika metoder för att skapa en kontrollgrupp som ska besvara den kontrafaktiska frågan. I detta underavsnitt börjar vi med att beskriva de största utmaningarna med kvasiexperimentella metoder i allmänhet och presenterar därefter fyra metoder som används allmänt i empiriska kontrafaktiska konsekvensbedömningar.

När randomisering inte är möjlig

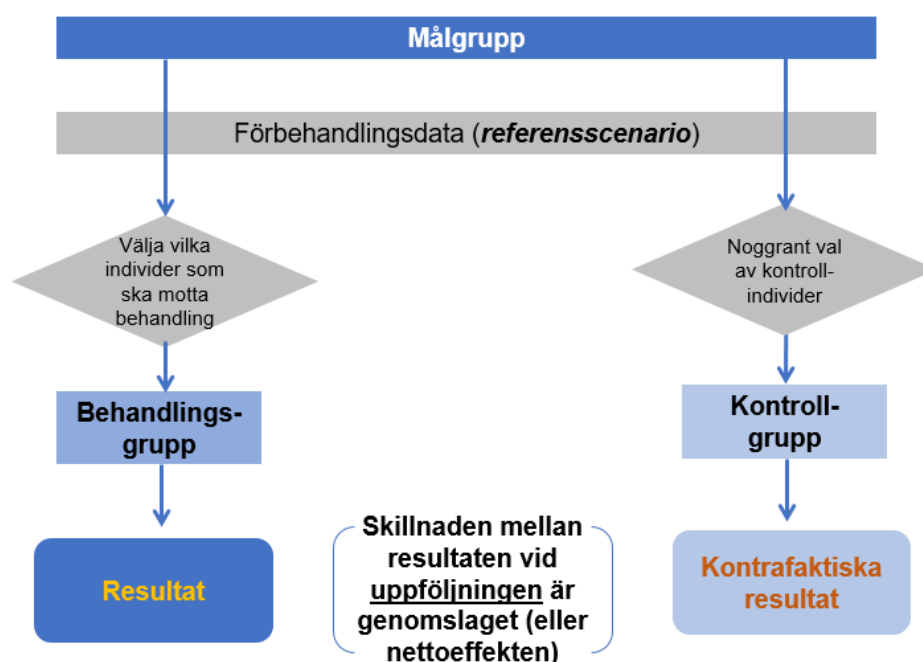
3.2.1 Mål- och kontrollgrupper utan randomisering

I kvasiexperimentella metoder jämförs insatsens målgrupper med en kontrollgrupp⁷⁸ av slumpmässigt fördelade mål eller potentiella mål som inte deltar i insatsen. Precis som i ett försök är målet att åstadkomma en opartisk uppskattning av de förändringar i utfallet som den berörda insatsen medför. Eftersom behandlings- och kontrollgrupperna inte bildas slumpmässigt

⁷⁸ I strikt bemärkelse är "kontrollgrupp" det begrepp som används för experimentell design, medan man i kvasiexperimentell design oftast talar om "jämförelsegrupper". I praktiken brukar begreppen dock användas utan åtskillnad. I denna vägledning använder vi därför begreppet "kontrollgrupp" för den grupp som används för att uppskatta den kontrafaktiska situationen, oberoende av hur den kontrafaktiska konsekvensbedömningen är utformad.

kräver kvasiexperimentella metoder mycket större uppmärksamhet på metoder som beräknar eventuella skillnader mellan deltagarna i behandlingsgruppen och potentiella kontrollindivider som sannolikt kan påverka beslutet att delta och därför påverkar resultaten. Det viktiga är att välja en rimlig kontrollgrupp. Om man inte väljer en lämplig kontrollgrupp och redovisar återstående skillnader mellan de två grupperna i analysen minskas uppskattningarnas trovärdighet, samtidigt som det blir svårt att utesluta alternativa förklaringar till observerade effekter.

Figur 7 Stiliserad kvasiexperimentell design med behandlings- och kontrollgrupper



När det gäller insatser som samfinansieras av ESF är det vanligast att använda en kvasiexperimentell utvärderingsmetod som består av två grupper, ett grundscenario och en uppföljning. En sådan metod består av en kontrollgrupp och en behandlingsgrupp som i randomisering, förutom att kontrollgruppen konstrueras (utan randomisering) av befintliga icke-deltagande grupper som sätts ihop så att de liknar behandlingsgruppen i så stor utsträckning som möjligt.

Kontroll- och behandlingsgrupperna måste likna varandra

En bra strategi för att hitta en giltig kontrollgrupp inom ramen för en kvasiexperimentell utformning är att välja kontrollindivider som har uteslutits från behandling på grundval av faktorer som inte har med deras kännetecken och potentiella resultat att göra. Under vissa omständigheter kan det, trots att kontrollgrupperna inte uttryckligen har konstruerats slumpmässigt, vara möjligt att identifiera individer eller företag i efterhand eftersom det faktum att de inte har exponerats för behandling visar sig vara slumpmässigt i förhållande till de potentiella resultaten. Om sådana omständigheter föreligger är de praktiskt taget idealiska för en kvasiexperimentell metod. Vissa personer i insatsens målgrupp kan till exempel uteslutas från deltagande i insatsen på grund av ett administrativt förbiseende eller fel. Förståelse av urvalsprocessen för behandlingen är

därför en ytterst viktig aspekt för att det ska vara möjligt att konstruera en giltig kontrollgrupp, detta kan inte nog betonas.

Det finns flera olika metoder för att ta fram en trovärdig kontrollgrupp. För det första kan en statistisk matchningsmetod tillämpas, dvs. data samlas in från både behandlade individer och från ett (vanligen mycket stort) urval av personer som inte mottagit behandling. Därefter konstrueras en kontrollgrupp från gruppen av individer som inte har mottagit behandling genom att man väljer de individer som har störst likheter med individerna i behandlingsgruppen. "Likhet" avser en uppsättning sociodemografiska kännetecken – ålder, kön, utbildning, sysselsättning eller arbetslöshet osv. – som mäts vid en tidpunkt innan behandlingsgruppen gick in i insatsen. På så sätt "matchas" obehandlade individer med behandlade individer på ett ändamålsenligt sätt. Om det finns många sociodemografiska kännetecken kan de i praktiken sammanfattas med hjälp av värdet för "benägenhetspoängen", vilket underlättar tillämpningen.

Matchning av behandlade och obehandlade individer

3.2.2 Benägenhetspoäng (*Propensity score matching*)

Viktiga aspekter av metoden med benägenhetspoäng

- Matchning imiterar randomisering genom att man i efterhand konstruerar en kontrollgrupp som är så lik behandlingsgruppen som möjligt i alla relevanta kännetecken.
- Till skillnad från randomisering kan endast observerbara kännetecken (ålder, kön, utbildning osv.) matchas, medan kännetecken som inte är observerbara (t.ex. motivation) inte kan beaktas.
- Metodens giltighet är starkt beroende av tillgång till uppgifter.

Benägenhetspoäng (*Propensity score matching*, PSM) innebär att en statistisk modell uppskattas för hela urvalet (behandlingsgruppen och potentiella kontroller), vilket ger en uppskattad benägenhet att delta för varje individ eller företag – oavsett om de faktiskt deltar eller ej⁷⁹. Behandlade individer eller företag matchas sedan på grundval av benägenhetspoängen, antingen med en individ eller ett företag som inte behandlats, eller med flera obehandlade individer eller företag⁸⁰. På så sätt identifieras en kontrollgrupp som därefter kan användas för att uppskatta den kontrafaktiska situationen. En sådan matchning säkerställer att de observerbara skillnaderna mellan behandlingsgruppen och de individer eller företag som fungerar som kontrollindivider beaktas i effektuppskattningarna. På så sätt får man fram en opartisk uppskattning av insatsens effekt, under förutsättning att alla relevanta skillnader före matchningen är möjliga att uppskatta. Om urvalet till behandling baseras på faktorer som inte är möjliga att observera är det dock osäkert om matchningen är lämplig för att eliminera fel. Det kritiska antagande som ligger till grund för matchningsmetoden är att urvalsprocessen kännetecknas av observerbara data.

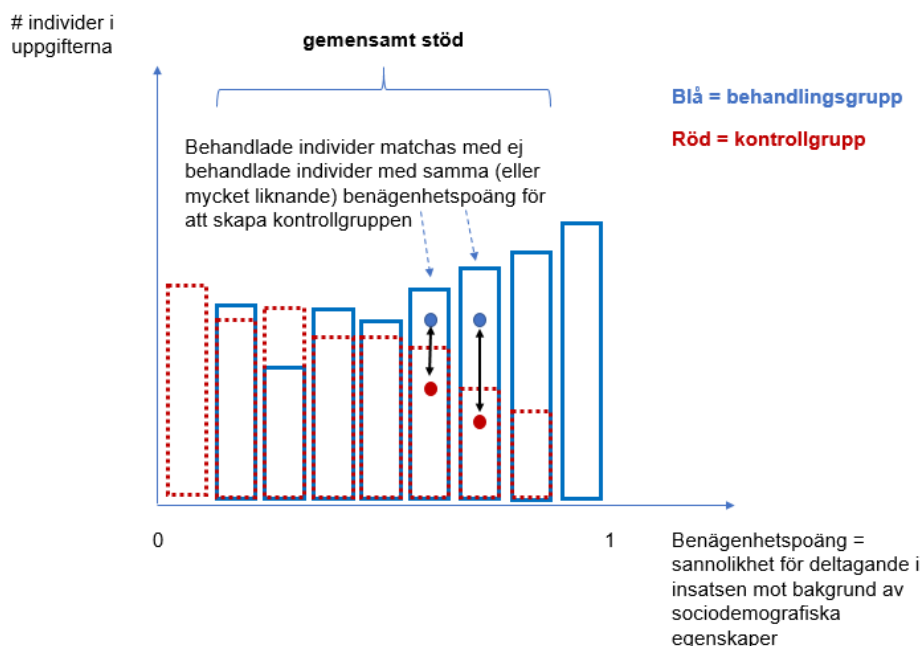
Benägenhet att delta som en metod för att fastställa behandlings- och kontrollgrupper

⁷⁹ För att förenkla diskussionen antar vi att beslutfattarna vill ta reda på hur de individer eller företag som faktiskt mottog tjänster från programmet har påverkats av behandlingen (detta motsvarar ofta en deluppsättning av den målgrupp som erbjöds möjligheten). Detta kallas för TOT-analys ("behandling av de behandlade").

⁸⁰ Det finns många metoder för matchning med benägenhetspoäng. För en lättförståelig översikt, se [Caliendo M., Kopeinig S., 2008](#).

I figuren nedan presenteras en intuitiv och förenklad version av metoden för matchning med benägenhetspoäng. Y-axeln motsvarar antalet individer i de behandlade och obehandlade grupperna, ordnade efter benägenhetspoäng på x-axeln. Behandlade individer brukar ha förhållandevis högre benägenhetspoäng, medan icke-behandlade individer brukar ha lägre poäng. Det område där benägenhetspoängen för de två grupperna sammanfaller kallas för området med gemensamt stöd.⁸¹ Behandlade fall matchas med obehandlade fall inom detta område. I diagrammet ges två exempel, men i själva verket upprepas processen till dess att varje behandlat fall har matchats med ett obehandlat fall inom området för gemensamt stöd. I figuren görs detta med hjälp av matchning med "närmaste granne". Den "närmaste grannen" till en individ i behandlingsgruppen är den individ i kontrollgruppen som har närmaste benägenhetspoängen. När de två grupperna har bildats kan deras genomsnittliga resultat jämföras för att få fram en uppskattning av effekten. I praktiken kan matchning med benägenhetspoäng vara en ytterst komplex process med ett antal frågor att ta hänsyn till. Många av dessa frågor utelämnas här för att tydliggöra de grundläggande principerna. Ett praktiskt exempel på en ESF-insats där en matchningsmetod användes presenteras i nästa ruta.

Diagram 8 Illustration av metoden med benägenhetspoäng



Metoden med benägenhetspoäng förutsätter bland annat att det är möjligt att fullständigt karakterisera valet till behandling med hjälp av observerbara data. Med andra ord får det inte finnas några ej observerade skillnader mellan behandlings- och kontrollgrupperna som är kopplade till resultaten och/eller beslutet att delta i insatsen. Rimligheten i detta antagande kan förstärkas genom att ett stort antal variabler integreras i uppskattningen av benägenhetspoängen. Valet av variabler ska baseras på befintlig kunskap

Urval baserade på observerbara data

⁸¹ Omfattningen av området för gemensamt stöd påverkar urvalsstorleken och resultatens användbarhet i politiskt hänseende, särskilt om ett stort antal behandlade individer faller utanför området för gemensamt stöd.

och teori. När det gäller arbetsmarknadsinsatser rekommenderar vi starkt att tidigare arbetsmarknadshistorik tas med i kontrollen av eventuella ej observerade skillnader⁸².

Ruta 15 Exempel på en utvärdering som tillämpar en matchningsmetod⁸³

Effekter av ESF-insatser som finansierades under programperioden 2014–2020 för långtidsarbetslösa i regionen Marche⁸⁴

En matchningsmetod användes för att utvärdera effekterna av ett antal ESF-insatser (praktikplatser, jobbstipendier, yrkeserfarenhet hos kommuner och utbildningskuponger) för långtidsarbetslösa som finansierades i regionen Marche i Italien. Insatserna riktades inte särskilt mot långtidsarbetslösa, men de utgjorde majoriteten av deltagarna i samtliga analyserade ESF-insatser.

Syftet med utvärderingen var att fastställa hur ESF-åtgärderna inverkar på sannolikheten att få jobb 6, 9, 12, 15 och 18 månader efter det att insatserna inleddes. Flera utfallsvariabler användes: sannolikhet att få anställning en viss tid efter insatserna, sannolikhet att få fast anställning och antal arbetsdagar under en period efter insatserna.

En metod med benägenhetspoäng tillämpades (matchning med närmaste granne) för att identifiera skillnader mellan behandlade och icke-behandlade grupper. De fyra ESF-insatserna analyserades separat. För varje insats bestod behandlingsgruppen av deltagare som gick in i insatsen före slutet av augusti 2019 (526 för praktikplatser, 1 058 för jobbstipendier, 236 för yrkeserfarenhet hos kommuner och 241 för utbildningskuponger). Kontrollgruppen bestod av arbetslösa som anmält sig till den offentliga arbetsförmedlingen under perioden 2016–2018 och som varit arbetslösa i minst tolv månader (77 255 poster).

Flera variabler som hämtats från behandlade och icke-behandlade individers sociodemografiska kännetecken användes för att beräkna benägenhetspoängen, t.ex. kön, ålder, medborgarskap, utbildningsnivå, boställningsort, datum när individen började delta i insatsen och sysselsättningshistorik 36 månader före insatsen.

För att mäta utfallsvariablerna för behandlings- och kontrollgrupperna (före och efter insatserna) användes data från COB, arkivet över arbetsgivares obligatoriska anmälan av anställningsavtal som skickas till den offentliga arbetsförmedlingen.

Studien gav blandade resultat. Positiva verkningar rapporterades för praktikplatser och jobbstipendier, medan yrkeserfarenhet hos kommuner visade en negativ inverkan och effekterna av utbildningskuponger var försumliga⁸⁵.

3.2.3 Skillnaden-mellan-skillnader

Viktiga egenskaper hos metoden skillnaden-mellan-skillnader

- Skillnaden-mellan-skillnader är en intuitiv metod där man jämför skillnaden mellan utfallet före/efter för behandlingsgruppen och kontrollgruppen.
- Eftersom förändringarna över tid i kontrollgrupper mäter vad som skulle ha hänt i behandlingsgruppen om insatsen inte hade jämförts (den kontrafaktiska situationen) visar eventuella ytterligare skillnader i utfallet före/efter hos behandlingsgruppen på insatsens kausala effekt.
- Detta är en enkel metod som går att använda i de flesta fall.

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

⁸³ Andra exempel på praktisk tillämpning av metoder (PSM, RDD och skillnaden-mellan skillnader finns i [Europeiska kommissionen, 2020](#).

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

⁸⁵ Såsom anges i utvärderingsrapporten måste resultaten för utbildningskuponger betraktas som preliminära, eftersom urvalet var för litet i förhållande till det totala antalet deltagare.

För uppskattningar med metoden skillnaden-mellan-skillnader (*difference-in-differences* (DID)) kan man använda resultatvariabler i form av utgångsmätningar (eller mätningar före behandling), antingen separat eller i kombination med matchning. Skillnaden i ett resultat före och efter behandling i kontrollgruppen dras sedan av från samma skillnad som observerats hos kontrollgruppen för att få fram en uppskattning av insatsens effekt. Även här är det mycket viktigt att välja en rimlig kontrollgrupp. Verkningar som beräknas med hjälp av skillnaden-mellan-skillnader observeras vanligen inom en regressionsram, som även redovisar andra observerade skillnader mellan behandlings- och kontrollgruppen. Med denna metod är det dessutom möjligt att undersöka om det finns skillnader mellan de två grupperna som inte observerats, dels skillnader som är oföränderliga över tid, och dels skillnader som varierar med tiden, men som påverkar både kontroll- och kontrollgruppen på samma sätt (t.ex. faktorer som påverkar hela ekonomin). Eftersom det är möjligt att kontrollera vissa aspekter av ej observerade skillnader mellan behandlings- och kontrollgruppen är metoden skillnaden-mellan-skillnader oftast bättre än matchning mellan olika tvärsnitt. Diagram 4 ger en visuell återgivning av metoden skillnaden-mellan-skillnader.

*Före och efter
behandling*

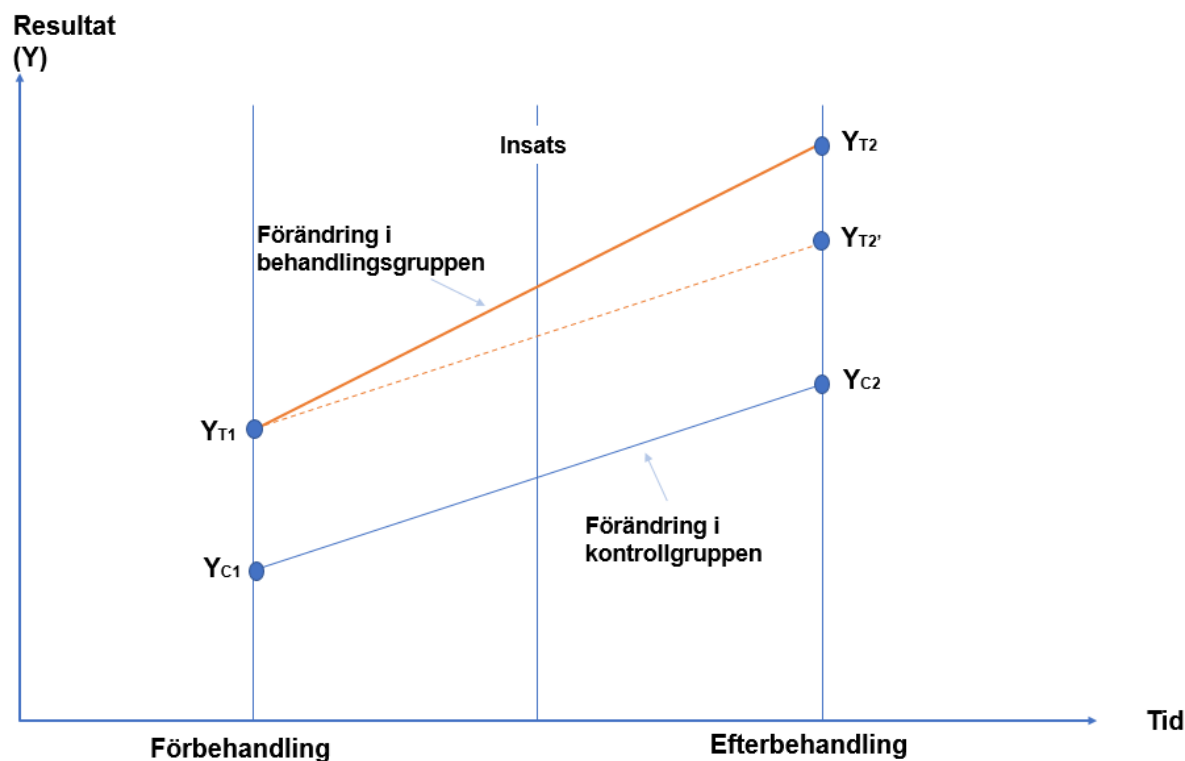
X-axeln motsvarar tidslinjen och y-axeln en skala där resultaten registreras. I detta fall kan resultaten vara löner. Genomsnittslönen för behandlingsgruppen under perioden före behandling är YT1 och YC1 för kontrollgruppen. Under perioden efter behandling är lönerna YT2 respektive YC2 för behandlingsgruppen respektive kontrollgruppen. Den heldragna övre linjen motsvarar således löneförändringarna hos kontrollgruppen, medan den heldragna undre linjen visar löneförändringar hos kontrollgruppen.

En grov uppskattning av insatsens effekt kan göras genom en jämförelse av lönerna för behandlings- och kontrollgruppen under perioden efter behandling, dvs. YT2–YC2. En sådan uppskattning är dock felaktig, eftersom löneskillnader före behandlingen inte beaktas. I metoden skillnaden-mellan-skillnader görs en feluppskattning före behandling som sedan dras av från resultatskillnaden efter behandling. På så sätt justeras löneskillnaden efter behandling (YT2–YC2) genom att den dras av från löneskillnaden före behandling (YT1–YC1). Den uppskattade effekten kan därför uttryckas mycket enkelt, enligt följande:

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

*Logiken "dubbel
skillnad"*

Om löneskillnaderna efter behandling inte justeras för de befintliga skillnaderna mellan behandlings- och kontrollgruppen blir uppskattningarna partiska. Som redan nämnts kan metoden skillnaden-mellan-skillnader också tillämpas genom att förändringen i resultat hos kontrollgruppen dras av från den förändring som observerats hos behandlingsgruppen. Den observerade förändringen hos kontrollgruppen anses motsvara det som skulle ha hänt i behandlingsgruppen utan insatsen.

Diagram 9: Förklaring av skillnaden-mellan-skillnader-metoden

I det enklaste fallet baseras huvudantagandet för metoden skillnaden-mellan-skillnader på gemensamma trender, vilket innebär att tidstrenderna i resultaten hos behandlings- och kontrollgruppen är likvärdiga utan behandling. Detta antagande kan inte testas direkt, men om flera mått för resultatvariabeln finns tillgängliga före behandling för både behandlings- och kontrollgruppen och dessa mått visar parallella trender blir metoden mer rimlig. Ett exempel på metoden skillnaden-mellan-skillnader visas i rutan nedan.

Huvudantagandet: gemensamma trender är likvärdiga i avsaknad av behandling

Ruta 16 Exempel på en utvärdering som tillämpar metoden skillnaden-mellan-skillnader

Utvärdering av åtgärder för äldre arbetstagare i Lubelskie som finansierades via ESF under programperioden 2007–2013⁸⁶

Studien syftade huvudsakligen till att utvärdera ESF-insatser för att få äldre att stanna kvar längre på arbetsmarknaden mot bakgrund av den ogynnsamma demografiska situationen i regionen Lubelskie i Polen. Insatserna finansierades via det operativa programmet "Humankapital" under programperioden 2007–2013.

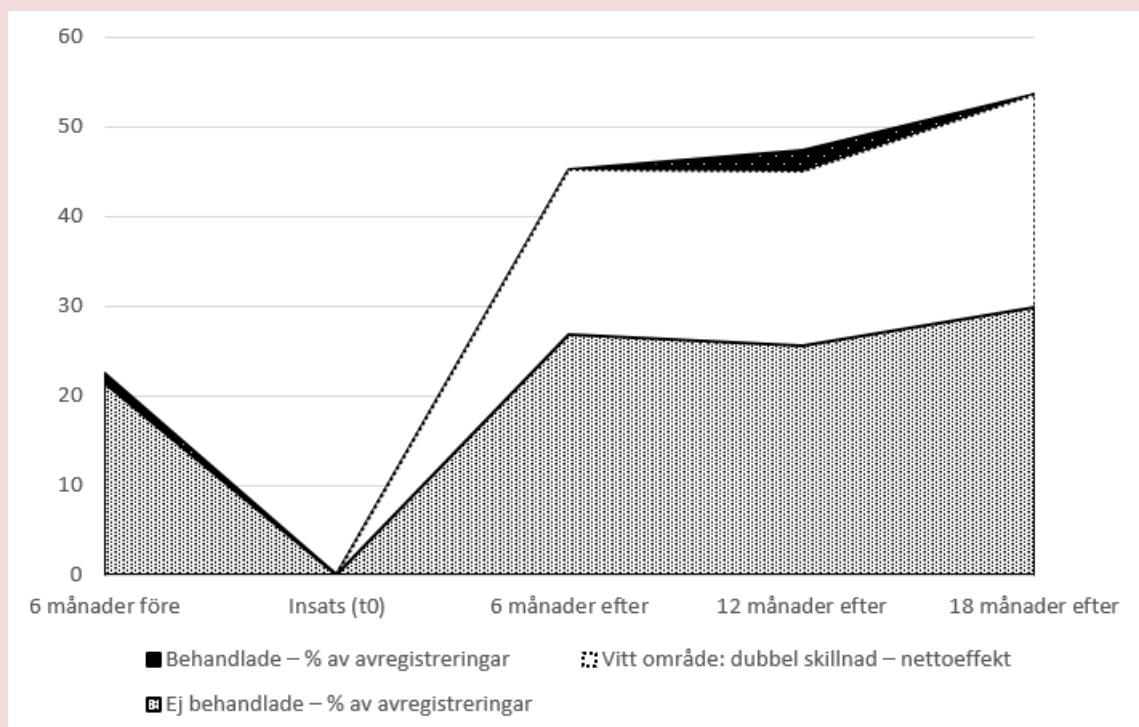
Den behandlade gruppen, ett urval på 145 individer, bestod av personer som var arbetslösa vid den tidpunkt då de gick in i insatsen. Kontrollgruppen hämtades från aidentifierade registeruppgifter från arbetsförmedlingarna, som innehöll sociodemografiska uppgifter och historik över registreringar rörande specifika personers sysselsättningsstatus (uppgifterna inhämtades från åtta arbetsförmedlingar och innehöll information om 67 102 personer). Utfallsvariabeln var "status som registrerad i arbetslöshetsregistret"⁸⁷. Den var observerbar både före och efter insatsen, vilket gjorde det möjligt att använda metoden skillnaden-mellan-skillnader. Sysselsättningsstatus mättes vid fem tidpunkter: sex månader före deltagandet i projektet, under deltagandet i projektet och sex, tolv respektive arton månader efter det att projektet avslutats.

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

⁸⁷ Att inte vara registrerad som "arbetslös" är visserligen inte samma sak som att vara "anställd", men detta antagande var nödvändigt för att utföra studien.

Urvalet för kontrollgruppen stärktes med hjälp av matchningsmetoder för att identifiera en obehandlad person för varje behandlad person baserat på ett antal observerbara kovariater, t.ex. kön, ålder, utbildningsnivå och arbetslöshetsstal för bostadsorten.

Analysen visade att stödet till arbetslösa över 50 år hade en positiv nettoeffekt på sysselsättningen: behandlade arbetslösa över 50 år hade 2,5 gånger större chans att avregistrera sig som arbetslösa jämfört med personer som inte fått behandling. Diagrammet nedan visar utfallsvariabelns utveckling för behandlingsgruppen och kontrollgruppen samt den uppskattade nettoeffekten i beaktande av det parallella antagandet enligt metoden skillnaden-mellan-skillnader.



3.2.4 Regressionsdiskontinuitetsdesign

Viktiga egenskaper hos regressionsdiskontinuitet

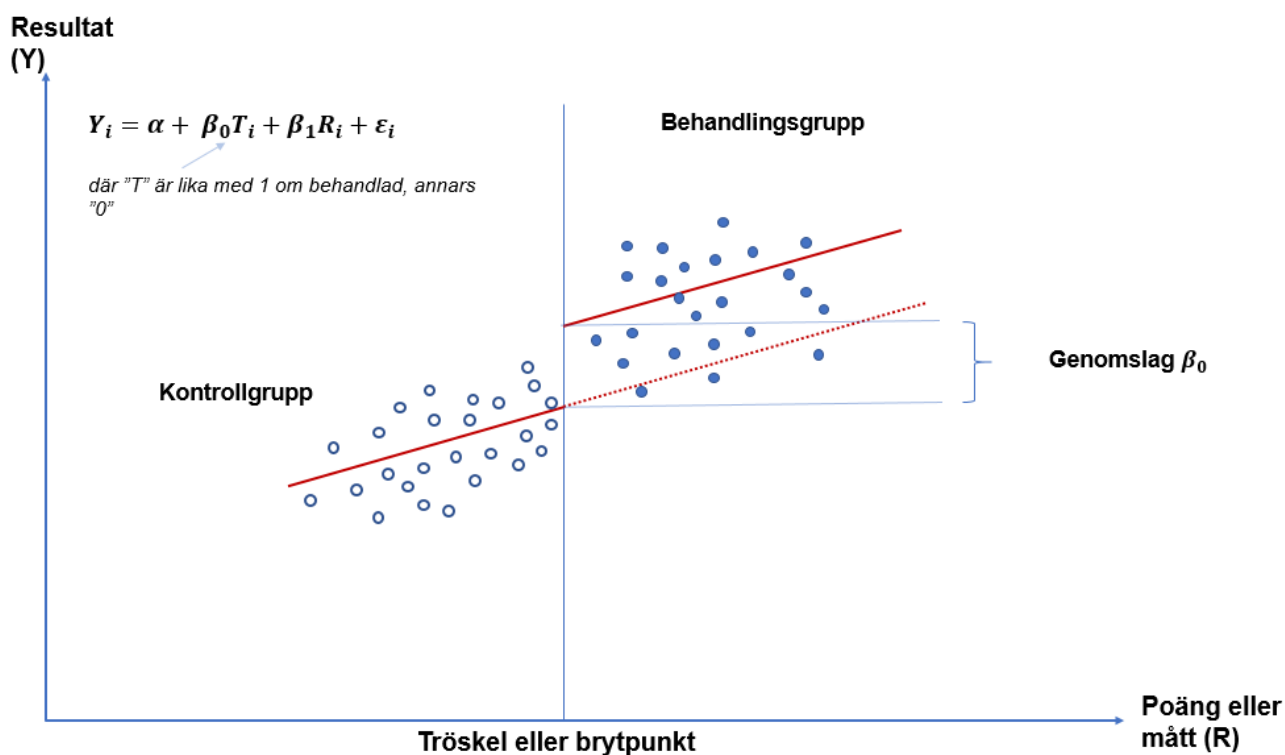
- Regressionsdiskontinuitetsdesign bygger på antagandet att ett specifikt värde ("brytpunkten") för en poäng avgör om individen deltar i en insats eller ej.
- Individer som ligger nära det värdet anses vara jämförbara, med den enda skillnaden att de individer som befinner sig på ena sidan av brytpunkten deltar (behandlingsgruppen), och de som befinner sig på andra sidan inte gör det (kontrollgruppen).
- En jämförelse mellan dessa grupper ger ett exakt och intuitivt mått på insatsens effekt, men den kräver en poängbaserad urvalsprocess.

Regressionsdiskontinuitetsmetoden kan användas när tillträdet till insatsen avgörs av en brytpunkt längs en kontinuerlig klassningsskala eller ett kontinuerligt mått. Tillträde till utbildning kan till exempel avgöras av resultaten av ett lämplighetsprov, där de som får en poäng över en fastställd tröskel (eller brytpunkt) mottar utbildning, medan de som får poäng under tröskeln inte

Brytpunkten delar individerna i behandlings- och kontrollgruppen

antas till utbildningen. För att metoden ska vara giltig bör brytpunkten fastställas utan tillgång till information om de eventuella deltagarnas poäng. De personer som får poäng som ligger precis vid brytpunkten liknar varandra, förutom att de som ligger precis över kommer att delta i insatsen medan de precis under inte kommer att göra det. Resultat för de personer som befinner sig över och under brytpunkten kan jämföras för att uppskatta insatsens effekt vid brytpunkten.

Diagram 10 Illustration av regressionsdiskontinuitetsmetoden



En regressionsdiskontinuitetsdesign (RDD) kan genomföras där brytpunkten antingen fullständigt identifierar behandlingsgruppen (full uppfyllelse), vilket innebär att man får fram en *skarp* diskontinuitet, eller, under vissa omständigheter, när inte alla individer på en sida av brytpunkten strikt följer behandlingen (*otydlig* diskontinuitet).

Skarp eller otydlig diskontinuitet

I figur 10 ovan visas ett stiliserat exempel på en regressionsdiskontinuitetsdesign. Detta är den enklaste illustrationen av en skarp diskontinuitet. Insatsen ger konstanta effekter vid varje värde i klassningen och verkningarna uppskattas med hjälp av en linjär regressionsmodell (det finns inga problem med regressionens funktionella form). I praktiken krävs en betydligt mer avancerad analys än den analys som presenteras i figur 10.

Prickarna i figur 10 motsvarar individuella enheter, t.ex. deltagare i utbildning. X-axeln visar den klassning eller det mått som har använts för att fördela deltagarna till platser i utbildningskursen. Individer som får poäng till höger av den heldragna vertikala linjen som anger tröskeln för klassningen eller måttet (t.ex. ett lämplighetsprov) får utbildning och bildar behandlingsgruppen. Individer som får poäng under tröskeln för klassningen eller måttet får inte utbildning och bildar kontrollgruppen.

Det viktigaste är att den klassning som används för att fördela målgruppen till behandling eller kontroll är en kontinuerlig kvantitativ variabel som mäts före behandling. Individerna godkänns därmed för deltagande i utbildningsprogrammet beroende på om deras poäng ligger över eller under den på förhand fastställda brytpunkten eller tröskeln.

Resultatet prickas in på y-axeln. Behandlingens inverkan identifieras genom en uppskattning av uppgifterna med hjälp av en linjär regressionsmodell (enligt antagandena ovan), dvs. genom regression av resultatvariabeln mot klassningsmättet, tillsammans med en dummyvariabel (behandlingsindikator), som definierar om poängen ligger under eller över brytpunkten (dvs. om individen ska anvisas till behandlings- eller kontrollgruppen).

En sådan regressionsekvation illustreras i figur 10. Effekten eller genomslaget av utbildningen i vårt exempel erhålls från koefficienten på behandlingsindikatorn, dvs. β_0 .⁸⁸ En sådan ekvation visar om det finns en brytning eller diskontinuitet kring brytpunkten, som i figur 10 anges genom ett skift uppåt i regressionslinjen vid tröskeln eller brytpunkten. I detta exempel observeras en positiv inverkan av utbildningen på resultatet.

Ett alternativt sätt att tolka uppskattningen av en insats genomslag är att överväga den streckade linjen till kontrollgruppen i figur 10. Detta kan anses vara en kontrafaktisk uppskattning för behandlingsgruppen – förhållandet mellan klassningen och resultatmättet som det skulle ha sett ut utan insatsen – där skillnaden mellan denna streckade linje och trendlinjen för behandlingsgruppen motsvarar behandlingens effekt eller genomslag. Tänk på att utan behandling finns det ingen diskontinuitet i linjen. I ett sådant fall förutsätter vi att resultatet varierar kontinuerligt efter klassningen eller måttet.

Regressionsdiskontinuitetsmetoden fungerar, eftersom observationerna i behandlings- och kontrollgrupperna nära brytpunkten liknar varandra, förutom det faktum att de som ligger över brytpunkten i det här exemplet får utbildning medan de som ligger under inte får det. Observationer nära brytpunkten liknar därför randomisering. Det finns dock en viktig begränsning. I de flesta tillämpningar ger uppskattningar av effekter som utförs med hjälp av en RDD-metod endast information om effekter för individer som befinner sig nära brytpunkten eller tröskeln. De individer som befinner sig längre bort från tröskeln kan vara svåra att generalisera.

God intern giltighet med generaliseringen kan vara begränsad

RDD kan vara en användbar metod när individer fördelas till en insats utifrån behov som mäts enligt en kontinuerlig klassning eller poängskala. Analysen kan dock bli komplex om brytpunkten är otydlig och en del individer inte deltagit i insatsen eller om det finns problem med den funktionella formen i regressionsmodellen. Det krävs en rad antaganden och deras giltighet kan inte alltid bekräftas.

I nästa ruta visas ett praktiskt exempel där en regressionsdiskontinuitetsmetod användes i en utvärdering av ett ungdomsgarantisystem.

⁸⁸ I ett enkelt fall skulle detta motsvara effekten av den tänkta behandlingen vid brytpunkten (se [Bloom H. S., 2009](#)).

Ruta 17 Exempel på en utvärdering som tillämpar en regressionsdiskontinuitetsmetod

Yrkesutbildning för arbetslösa ungdomar i Lettland: Belägg från en regressionsdiskontinuitetsdesign⁸⁹

Forskare från kommissionens gemensamma forskningscentrum använde en regressionsdiskontinuitetsmetod för att bedöma effekterna av ett yrkesutbildningsprogram på arbetsmarknadsresultat för arbetslösa ungdomar i Lettland. Programmet finansierades inom ungdomsgarantisystemet för perioden 2014–2020 och riktades till unga som varken arbetar eller studerar i åldern 15–29 år.

Uppgifterna hämtades från den lettiska statliga arbetsförmedlingen som gav information om både deltagare och icke-deltagare som var anmälda som arbetslösa på vissa datum. Dessa uppgifter matchades mot data från det statliga skattekontoret, som innehåller information om personers inkomst på olika datum före och efter programmet (januari 2012 och juni 2017). Efter rensningen av datauppsättningarna innehöll det slutliga urvalet för behandlingsgruppen upp till 898 individer och antalet för kontrollgruppen var 10 717.

Utvärderarna förlitade sig till en bestämmelse om att arbetslösa under 25 år skulle prioriteras i yrkesutbildningsprogrammet. Därför är ålder den löpande variabel som avgör sannolikheten för deltagande i yrkesutbildningsprogrammet, där åldern 25 år utgör den tröskel under vilken deltagandet i programmet ökar kraftigt till följd av prioritetsregeln. Eftersom individer inte har någon kontroll över sin ålder kan fördelningen till yrkesutbildningsprogrammet anses vara slumpmässig kring tröskeln. Prioritetsregeln och omständigheten att deltagandet var frivilligt innebar att forskarna fick använda en FRDD-design (*Fuzzy Regression Discontinuity Design*). Identifieringsstrategin byggde på att personer nära brytpunkten (åldern 25 år) liknande varandra förutom deltagandet i yrkesutbildningsprogrammet.

Resultaten visade att effekterna i fråga om möjlig framtida sysselsättning och månadsinkomst upp till tre och ett halvt år efter deltagandet i programmet är positiva, men inte statistiskt signifikanta. Dock observerades en positiv effekt av prioriteringsregeln på programmet. Eftersom giltigheten för FRDD-metoden bygger på att potentiella deltagare inte kan kontrollera den löpande variabeln (ålder) genomfördes flera försök för att bekräfta denna effekt.

3.2.5 Instrumentvariabler

Viktiga egenskaper hos instrumentvariabler

- Tanken med instrumentvariabler är att vissa (på förhand fastställda eller exogena) processer avgör deltagandet, men inte utgör det faktiska urvalsförfarandet.
- Den förhandsbestämda processen och dess nyckelvariabler kan användas för att beakta situationer där individen själv väljer att delta eller inte.

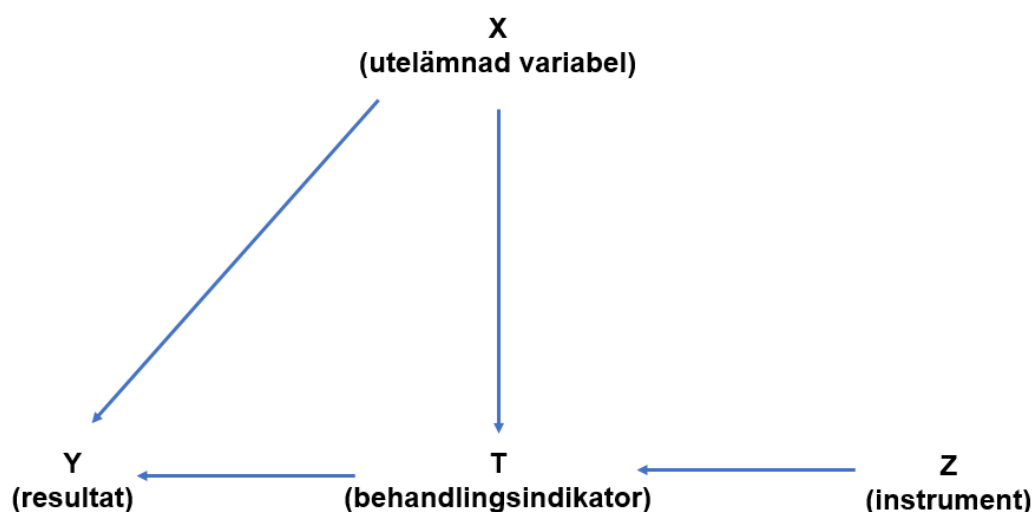
För metoden med **instrumentvariabler** (IV) bör urvalet till behandling åtminstone delvis avgöras av en exogen faktor som inte har någon annan koppling till resultaten än behandlingen. Som sådan påverkar den exogena faktorn deltagandet men har ingen direkt inverkan på resultaten. Exogena faktorer brukar vara administrativa fel eller förbiseenden, eller andra slumpmässiga variationer i den mottagna behandlingen.

Hur exogena faktorer påverkar deltagandet

Metoden med instrumentvariabler illustreras i figur 11 nedan. Fyra variabler visas i ett ytterst förenklat kausalt system. Variablerna motsvarar data som samlats in från en grupp som är den hypotetiska målgruppen för en utbildningsinsats (både de individer som mottar behandling och de som fungerar som kontroller).

⁸⁹ Se [Bratti M. m.fl., 2018](#).

Figur 11 Illustration av metoden med instrumentvariabler



"Y" motsvarar det resultat som ska undersökas. För en utbildningsinsats kan resultatet vara inkomster. "T" är en indikator som visar om en individ har inlett utbildning⁹⁰.

"X" är en utelämnad variabel som inte observeras men är kopplad både till "Y" (resultatet) och "T" (behandlingsindikatorn). Den utgör en förlängning av utbildningsprogrammet, t.ex. ett utgångsvärde för förmåga. I detta fall rör "förmåga" både deltagande i utbildningen och inkomster. Till exempel kan fler individer med förmåga i målgruppen välja att inleda utbildningen och få högre inkomst.

X-variabeln motiverar sökandet efter ett instrument och leder till att utbildningens effekt på inkomsterna blir "förvirrat". Med andra ord blir uppskattningen partisk på grund av X-variabeln, eftersom den inte observeras och inte kan redovisas direkt i analysen.

Variabeln "Z" är ett instrument som enligt Morgan och Winship⁹¹ kan ses som en chock för "T", som är oberoende av "X". Därför finns det ingen linje i figur 11 som kopplar Z till X. Variabeln Z påverkar dessutom endast Y via T, variabeln Z kan alltså inte påverka Y på något annat sätt. Detta innebär att Z kan användas för att skapa en variation i T (behandlingen) som inte är kopplad till den förvirrande variabeln X. Genom denna variation är det således möjligt att få fram ett partiskt mått av T- och Y-variablernas effekt⁹².

Ett giltigt instrument är en absolut förutsättning

Nedan följer en förenklad beskrivning av de omständigheter då en IV-metod kan användas. Vi går inte in på metodens komplexitet. I praktiken är det svårt att hitta ett övertygande instrument. De olika instrumenten är ytterst kontextbundna och de underliggande antagandena kan oftast inte testas statistiskt (närmare bestämt kan den nödvändiga kopplingen mellan Z och T testas statistiskt, men inte "oberoendet" för Z och X). En strategi kan till

⁹⁰ Med andra ord är efterlevnaden fullständig, dvs. alla individer i behandlingsgruppen deltar i utbildningen.

⁹¹ Morgan S.L., Winship C., 2014.

⁹² Den kausala effekten av T på Y beräknas med ett instrument som uppskattar förhållandet mellan Z och Y, som sedan divideras med det uppskattade förhållandet mellan Z och T.

exempel vara att använda avståndet från utbildningslokalerna (den fysiska platsen för utbildningskursen) till en deltagares bostad som ett instrument för att uppskatta utbildningens effekt på deltagarens nettoinkomster. I ett sådant fall skulle man kunna observera att deltagare som bor närmare utbildningslokalerna är mer benägna att delta i utbildningsinsatsen och att avståndet mellan deltagarens bostad och utbildningslokalerna inte har något samband med andra bestämningsfaktorer för nettoinkomster och deltagande i utbildning (t.ex. humankapitalåtgärder). Detta distansmått kan således endast påverka nettoinkomsterna genom sin effekt på utbildningen⁹³.

Instrumentvariabler kan användas i många olika sammanhang. Det finns flera uppskattningsmetoder att välja mellan beroende på responsvariabeln. Hittills har denna metod sällan använts i ESF-utvärderingar. I ruta 18 nedan illustrerar vi ett exempel på en analys av kausala effekter av en politisk åtgärd med hjälp av en IV-metod.

Ruta 18 Exempel på en studie som tillämpar en metod med instrumentvariabler

Sysselsättnings effekter av språkkurser för arbetslösa invandrare⁹⁴

Kompetens i värdlandets språk är en viktig faktor för en effektiv arbetsmarknadsintegration av invandrare.

I en färsk forskningsrapport (Lang J., 2021) använder författaren en metod med instrumentvariabler för att uppskatta den kausala effekten av yrkesinriktat språkkursprogram på sysselsättningsresultaten för de invandrare som deltog i kurserna, två år efter det att de skrev in sig. Programmet genomfördes av den federala myndigheten med ansvar för migrations- och flyktingfrågor (BAMF) och finansierades via ESF. ESF-BAMF-programmet löpte fram till slutet av 2017. Kursen bestod av flera delar: undervisning i tyska, uppbyggande av yrkeskompetens och praktikplatser. Personer med migrantbakgrund som var aktiva på den tyska arbetsmarknaden var berättigade att delta, oavsett nationalitet och invandringsdatum.

För att hantera den ej observerade heterogeniteten vad gäller språkkunskaper använder författaren exogen variation för att mäta intensiteten hos de språkkurser som erbjuds av de lokala arbetsförmedlingarna. Detta anses utgöra ett giltigt instrument eftersom arbetsförmedlingarna har stor handlingsfrihet att genomföra olika kombinationer av program beroende på respektive policy, och variabeln är exogen i förhållande till de arbetssökandes sysselsättningsresultat. De kurser som analyserades hölls under 2014.

Utvärderingen bygger på en bred uppsättning administrativa data, nämligen "integrerade anställningsbiografier" (IEB), den tyska federala arbetsförmedlingens kombinerade databas med administrativ information. IEB innehåller uppgifter om deltagande i språkkurser, detaljerad information om sysselsättningshistorik (förutom egenföretagande), jobbsökningar, överföringar under arbetslöshetsperioder och personuppgifter. Andra uppgifter (WGH, *Werdegangshistorie*) användes för att inhämta detaljerad information om deltagandet i de språkkurser som tillhandahölls av BAMF och när så var möjligt för att kompensera data som saknas i IEB med information om egenföretagare och föräldradighet.

Resultaten visar att efter en period av inläsningseffekter som i allmänhet varade i drygt ett år har invandrare som deltog i insatsen efter två år nio procentandelars större sannolikhet att få anställning än kontrollgruppen. Sannolikheten är inte begränsad till okvalificerade arbeten. Man fann även att snabbt tillhandahållande av språkundervisning (snart efter ankomsten till Tyskland) gynnar invandrares integration på arbetsmarknaden.

⁹³ Tolkningen av resultat från en sådan analys kan kompliceras om instrumentet är kopplat till variationer i behandlingseffekten (se [Bryson m.fl., 2002](#)).

⁹⁴ Se [Lang, J., 2021](#).

Tabell 5 Jämförelse av viktiga egenskaper hos de vanligaste metoderna för kontrafaktiska konsekvensbedömningar

Metod	Huvudegenskaper	Fördelar	Datakrav	Begränsningar
Randomisering – experimentell metod	<ul style="list-style-type: none"> - Individer som är berättigade att delta fördelas slumpmässigt till en behandlingsgrupp och en kontrollgrupp. - Randomisering säkerställer att båda grupperna (generellt) är identiska när det gäller alla relevanta kännetecken. - Kontrollgruppen besvarar den kontrafaktiska frågan och skillnaden i utfall mellan behandling och kontroll ger insatsens kausala effekt. 	<ul style="list-style-type: none"> - Om metoden genomförs korrekt är uppskattningarna av genomslaget "opartiska". - Resultaten är transparenta och lättförståeliga. - Resultaten är mindre mottagliga för kvalifikation och tvivel. - Metoden kan utformas på en rad olika sätt för att hantera olika politiska sammanhang och insatser. 	<ul style="list-style-type: none"> - En grundläggande förutsättning är att kontrollera urvalet till insatsen via randomisering. - Uppgifter om vilka individer som har fördelats till vilka grupper. - Insamling av referensdata är en grundläggande förutsättning. - Resultatmått ska registreras för både behandlings- och kontrollgruppen. 	<ul style="list-style-type: none"> - Det är vanligt att behandlings- och kontrollgruppen inte uppfyller fördelningsreglerna för behandlingen. - Deltagarnas samtycke är ofta ett krav. - Randomisering kan påverka valet av personer som deltar i/ansöker om att få delta i insatsen. - Om deltagarna känner till sin fördelningsstatus kan detta påverka deras beteende, vilket i sin tur påverkar resultaten. - Etiska aspekter. - Kräver mycket arbete med planering och genomförande. - Kan vara kostsam, även om detta inte nödvändigtvis behöver vara fallet.
Matchning (benägenhetspoäng)	<ul style="list-style-type: none"> - Matchning imiterar randomisering genom att man i efterhand konstruerar en kontrollgrupp som (generellt sett) är så lik behandlingsgruppen som möjligt i alla relevanta kännetecken. - Till skillnad från randomisering kan endast observerbara kännetecken (ålder, kön, utbildning osv.) matchas, medan kännetecken som inte är observerbara (t.ex. motivation) inte kan beaktas. - Metodens giltighet är således starkt beroende av tillgång till uppgifter. 	<ul style="list-style-type: none"> - Kräver god kunskap om urvalsprocesser, men inte direkt kontroll över urvalet till insatsen. - Kan tillämpas retroaktivt om de rätta uppgifterna finns tillgängliga och kan användas i en rad olika sammanhang. - Matchning är tekniskt sett en semiparametrisk uppskattningsmetod, kräver färre parametriska antaganden (standardmässiga regressionsuppskattningar krävs t.ex. inte). - Kan användas för att uppskatta flera behandlingseffekter. 	<ul style="list-style-type: none"> - Exakt identifiering av insatsens deltagare. - Datakällor finns tillgängliga för urvalet. - Tydligt koncept för deltagande och god förståelse av urval till behandling. - Det måste finnas tillgång till mycket information, som helst ska samlas in vid den referenspunkt från vilken matchningen väljs. - Resultatmått för insatsen för deltagare och kontroller. 	<ul style="list-style-type: none"> - Kräver en stor datamängd som möjliggör en fullständig karakterisering av urvalsprocessen. - Giltigheten beror på kvaliteten på kontrollerna, som därför måste väljas noggrant, och på graden av gemensamt stöd. - Bygger på antagandet att urvalet till insatsen kan karakteriseras på lämpligt sätt med observerbara data. - Det finns ett antal olika matchningsmetoder, och valet av metod kräver därför en känslighetsanalys. - Resultaten kan vara komplicerade att förklara och tolka och kan eventuellt vara tvetydiga.

Metod	Huvudegenskaper	Fördelar	Datakrav	Begränsningar
Skillnaden-mellan-skillnader (DID)	<ul style="list-style-type: none"> - Skillnaden-mellan-skillnader (DID) är en intuitiv metod där man jämför skillnaden i utfallet före/efter mellan behandlingsgruppen och kontrollgruppen. - Eftersom förändringarna över tid i kontrollgrupper mäter vad som skulle ha hänt i behandlingsgruppen om insatsen inte hade jämförts (den kontrafaktiska situationen) visar eventuella ytterligare skillnader i utfallet före/efter hos behandlingsgruppen på insatsens kausala effekt. - Detta är en enkel metod som går att använda i de flesta fall. 	<ul style="list-style-type: none"> - Gör det möjligt att kontrollera vissa aspekter av ej observerade skillnader mellan deltagare och kontroller. - Kan användas i kombination med matchning. - Fungerar med uppgifter från tiden före och efter insatsen, t.ex. paneldata (individuella uppgifter över tid) eller upprepade tvärsnittsdata (uppgifter om individer som samlas in vid olika tidpunkter). 	<ul style="list-style-type: none"> - Datakraven liknar andra metoder, men kräver även tillgång till resultatmått före insatsen. - För att pröva huvudantagandena krävs flera observationer av resultaten före behandlingen både för behandlings- och kontrollgruppen. 	<ul style="list-style-type: none"> - Kräver antaganden om gemensamma trender i resultaten mellan deltagare och kontroller. - Analysen kan vara komplex och öppen för feltolkningar. - Det måste finnas tillgång till mycket information för att pröva antagandena om gemensamma trender. - Kan inte användas för att uppskatta flera behandlingseffekter⁹⁵.
Regressionsdiskontinuitetsdesign (RDD)	<ul style="list-style-type: none"> - Regressionskontinuitet bygger på antagandet att ett specifikt värde ("brytpunkten") för en poäng avgör om individen deltar i en insats eller ej. - Individer som ligger nära det värdet anses vara jämförbara, med den enda skillnaden att de individer som befinner sig på ena sidan av brytpunkten deltar (behandlingsgruppen), och de som befinner sig på andra sidan inte gör det (kontrollgruppen). - En jämförelse mellan dessa grupper ger i allmänhet ett exakt och intuitivt mått på insatsens effekt, men den kräver en poängbaserad urvalsprocess. 	<ul style="list-style-type: none"> - Det finns både skarpa och otydliga RDD-metoder. - Kan under vissa förutsättningar visa opartiska verkningar av behandlingseffekterna. 	<ul style="list-style-type: none"> - Brytpunkterna ska väljas oberoende av värdena för klassningen av varje individ i målgruppen. - Det krävs tillgång till data om individerna både för behandlings- och kontrollgruppen när det gäller klassningen eller måttet, tröskeln eller brytpunkten och resultaten. 	<ul style="list-style-type: none"> - Denna metod är inte giltig utan ett kontinuerligt mått eller klassning som avgör behandlingen. - Analyserna kan bli komplicerade och osäkra om det uppstår problem med regressionens funktionella form, om inte alla individer deltar eller om storleken på urvalet kring brytpunkten är begränsad. - Risker kan uppstå i samband med resultatolkningen och extrapoleringen av generaliseringar.
Instrumentvariabler (IV)	<ul style="list-style-type: none"> - Tanken med instrumentvariabler är att det finns vissa (på förhand fastställda eller exogena) processer som avgör deltagandet, men inte utgör den faktiska urvalsprocessen. 	<ul style="list-style-type: none"> - Kan ge uppskattningar av hög kvalitet eller belägg för förekomsten av kausala effekter. - Löser problem med fel till följd av utelämnade variabler (eller urvalsfel). 	<ul style="list-style-type: none"> - Kräver referensdata, data om resultat och deltagande i insatsen samt att ett instrument kan identifieras. - Instrumentet måste vara kopplat till 	<ul style="list-style-type: none"> - Det kan vara svårt att hitta ett rimligt instrument. - Kan vara svår att förklara för lekmän. - Resultaten kan inte tolkas direkt, begränsade möjligheter att testa de

⁹⁵ Se [Frolich, M., 2004](#).

Metod	Huvudegenskaper	Fördelar	Datakrav	Begränsningar
	<ul style="list-style-type: none"> - Den förhandsbestämda processen kan därefter användas för att beakta situationer där individen själv väljer att faktiskt delta eller inte. 	<ul style="list-style-type: none"> - Kan tillämpas retroaktivt. 	deltagandet i insatsen och påverkar endast resultaten på denna basis. Instrumentet bör inte kopplas till några andra bestämningsfaktorer för resultaten.	underliggande antagandena.

Kapitel 4. Främja användningen av kontrafaktiska konsekvensbedömningar

Syftet med denna vägledning är att uppmuntra förvaltande myndigheter att oftare använda kontrafaktiska konsekvensbedömningar och hjälpa dem att ta fram bedömningar av god kvalitet. Den är tänkt att vägleda dig som ansvarar för planering och utförande eller beställning av konsekvensbedömningar av insatser som samfinansieras via ESF+. Hittills har vi inriktat oss på planeringen och genomförandet av kontrafaktiska konsekvensbedömningar och vi även har diskuterat ett antal frågor som behöver övervägas i detta sammanhang. Under programperioden 2014–2020 framkom dock ett antal andra mer allmänna frågor och problem som behöver åtgärdas för att förbättra utvärderingarna av ESF-insatser och öka användningen av kontrafaktiska konsekvensbedömningar.

I detta avsnitt presenterar vi några förslag på hur dessa allmänna frågor kan lösas. Vi diskuterar i synnerhet åtgärder för att lösa följande problem:

Mer allmänna frågor som behöver lösas

- Brist på kunskap om metoder för kontrafaktiska konsekvensbedömningar hos de förvaltande myndigheterna och generellt bland beslutsfattare i medlemslandet.
- Brist på externa, lämpligt kvalificerade och erfarna leverantörer som kan utföra kontrafaktiska konsekvensbedömningar.
- Rättsliga hinder i samband med kontrafaktiska konsekvensbedömningar.
- Förbättrad planering av kontrafaktiska konsekvensbedömningar.
- Utvidgning av tillämpningsområdet för kontrafaktiska konsekvensbedömningar.

4.1. Förbättra förståelsen bland berörda aktörer

För programperioden 2021–2027 anges följande i artikel 44.1 i förordningen om gemensamma bestämmelser: "Medlemsstaten eller den förvaltande myndigheten ska utvärdera programmen mot ett eller flera av följande kriterier: ändamålsenlighet, effektivitet, relevans, samstämmighet och mervärde för unionen, i syfte att förbättra kvaliteten när det gäller utformning och genomförande av program. Utvärderingarna får också omfatta andra relevanta kriterier, såsom inkludering, icke-diskriminering och synlighet, och får omfatta fler än ett program." Som vi har diskuterat i de föregående avsnitten innebär detta att det måste finnas tillräcklig utvärderingskapacitet både hos den förvaltande myndighet som lägger ut uppdraget att utföra en kontrafaktisk konsekvensbedömning och hos de utvärderare som ansöker om kontraktet.

I vissa fall har hjälpcentralen för utvärdering noterat att det saknas kapacitet för att utföra kontrafaktiska konsekvensbedömningar, dvs. att lägga ut bedömningen och bistå i genomförandet, hos en del

förvaltningar. Detta skapade svårigheter för utvärderarna eftersom förvaltningarna i dessa fall inte hade tagit fram tydliga utvärderingsfrågor, säkerställt tillgång till grundläggande data och planerat ordentligt i förväg inför genomförandet.

Ibland, och särskilt i mindre länder, saknas kapacitet och teknisk expertis för kontrafaktiska konsekvensbedömningar på konsultmarknaden. Det finns följaktligen ett utbrett behov av att stimulera både efterfrågan och utbud vad gäller kontrafaktiska konsekvensbedömningar. Situationen på utbudssidan kan förbättras när de förvaltande myndigheterna och medlemsländerna börjar utföra eller efterfråga kontrafaktiska konsekvensbedömningar. Hur snabbt den ökade efterfrågan tillgodoses beror på befintlig kompetens och erfarenhet och om det finns institutioner i medlemsländerna som kan genomföra sådana metoder. Efterfrågan kan dock även delvis stimuleras genom åtgärder för att förbättra kunskapen om och förståelsen av metoder för kontrafaktiska konsekvensbedömningar bland personalen hos de förvaltande myndigheterna.

Stimulera efterfrågan och utbud av kontrafaktiska konsekvensbedömningar

En lösning på problemet är att den förvaltande myndigheten kör utbildningskurser i metoder för kontrafaktiska konsekvensbedömningar för sin personal. Utbildningen bör inriktas på fördelarna för den förvaltande myndigheten med att införa metoder för kontrafaktiska konsekvensbedömningar. Dessutom bör ansvarsfrågor och vad som fungerar i praktiken diskuteras. Förslag på kursinnehåll ges i bilaga 2.

Utveckling av utbildning i metoder för kontrafaktiska konsekvensbedömningar

4.2. Kapacitetsuppbyggnad

En annan fråga som kom upp under programperioden 2014–2020 och som nämns i de föregående avsnitten var behovet av att bygga upp kapacitet för kontrafaktiska konsekvensbedömningar bland forskare, akademiker och konsulter ute i medlemsländerna. I vissa fall var det uppenbart att det fanns nödvändig kompetens i medlemsländerna, men att kompetensen inte kunde utnyttjas i utvärderingssammanhang (t.ex. begränsad tillgång till användbara uppgifter eller problem med att identifiera en lämplig kontrollgrupp).

Det finns flera lösningar för att utveckla utbudet av utvärderingstjänster. Många av de problem som har tagits upp gäller såväl kontrafaktiska konsekvensbedömningar som utvärderingar mer generellt. Det finns tre generella åtgärder för att förbättra utbudet av utvärderingar:

Stärka institutioner och bygga upp praktikgemenskaper

- Bygga upp relationer med akademiska institutioner, särskilt universitet.
- Utveckla och stärka en oberoende konsultgemenskap.
- Stödja utvecklingen av ett professionellt utvärderingssamfund.

Universitet

Skapandet av kopplingar till universitet är viktigt av två skäl: För det första kan akademisk personal vid universiteten besitta den kompetens och kunskap som krävs för att utföra kontrafaktiska konsekvensbedömningar. Många akademiker som sysslar med mikroekonomi samt ekonometriker, sociologer som arbetar med kvantitativ sociologi och psykologer har till exempel den typ av kompetens som behövs. I många medlemsländer

Utveckla akademisk kompetens

kan den nödvändiga kompetensen finnas tillgänglig, men de personer som besitter kompetensen har kanske inte tänkt på att de kan utnyttja den för att utvärdera insatser. Det kan saknas incitament för dem att göra detta, och det är ett problem som behöver lösas.

I en del medlemsländer är det tradition att akademiska forskare aktivt sysslar med tillämpad forskning på det politiska området. I sådana fall är akademikerna vana att samarbeta med statliga myndigheter och förvaltande myndigheter. I andra medlemsländer, där universiteten och akademikerna inte arbetar så mycket med tillämpad forskning, kan det behövas en attitydförändring. En effektiv metod för att ta fram en leverantörsbas inom universitetssektorn är att medlemsländerna och de förvaltande myndigheterna finansierar kostnaderna för kärnverksamheten hos forskningscentrum som arbetar med metoder för kontrafaktisk konsekvensbedömning.

Universitet och akademiker kan även bidra genom att utbilda nästa generation av utvärderare. Vid ett nära samarbete med universitet kan universiteten uppmuntras att ta med programutvärderingsmetoder i sina kursplaner, och som ett led i detta, att kontrafaktiska konsekvensbedömningar ingår i undervisningsprogrammet. I en del medlemsländer fyller universiteten även en funktion genom att ordna fortbildningskurser om konsekvensbedömningar i allmänhet och kontrafaktiska konsekvensbedömningar i synnerhet. Fortbildningen kan riktas till politiskt ansvariga, tekniska experter inom de förvaltande myndigheterna och andra potentiella leverantörer, t.ex. oberoende konsulter. Medlemsländerna kan överväga att finansiera sådan utbildning.

Utbilda nästa generation

Oberoende konsulter

För vissa storskaliga utvärderingar finns det en internationell marknad. Så är definitivt fallet för kontrafaktiska konsekvensbedömningar. Många medlemsländer vill dock säkerligen utveckla kapacitet för att utföra kontrafaktiska konsekvensbedömningar inom landet. Ett medel för att uppnå detta kan vara att etablera strategiska allianser mellan potentiella inhemska leverantörer och internationella konsultföretag.

Utveckling av marknaden

Nedan följer flera förslag på hur medlemsländerna kan utveckla en leverantörsbas inom landet för kontrafaktiska konsekvensbedömningar, som kan användas av förvaltande myndigheter (eller andra organ) som beställer sådana konsekvensbedömningar.

- Införa ett krav på att det alltid ska finnas med lokala konsulter i bud från konsortier.
- Utforma utvärderingskontrakt så att förhållandevis små utvärderingar med låga risker kan genomföras av nya nationella aktörer på utvärderingsmarknaden.
- Säkerställa att de tekniska och finansiella kraven i anbudsförfaranden avseende utvärderingar inte är alltför restriktiva och gör det möjligt för nya aktörer att delta.
- Prioritera tekniska kriterier och kunskapskriterier i urvalsförfarandet, i stället för komplexa administrativa förfaranden som mindre erfarna konsulter kanske inte känner till.

- Anordna informationsmöten med potentiella konsulter för att besvara frågor och främja bud i en konkurrensutsatt miljö.
- Stödja nätverksbyggande bland relativt isolerade utvärderingskonsulter för att främja teambyggande, bildande av konsortier och andra branschnätverk och organisationer, även på internationell nivå.
- Uppmärksamma dem som beställer utvärderingar på att det kan krävas mer praktiskt inriktad information till nya leverantörer så att de snabbare kan förvärva kunskap och erfarenhet.

Ruta 19 Exempel på ett projekt som inriktas på att stärka kulturen och kapaciteten när det gäller kontrafaktiska konsekvensbedömningar

Under programperioden 2014–2020 genomfördes ett ESF-projekt i Sverige i syfte att förbättra Arbetsförmedlingens utvärderingskultur och utvärderingskapacitet. Projektet, som benämndes "Evidensbaserade EU-fondsprojekt 2014–2020", löpte från 2016 till 2018 och omfattade seminarier och andra utbildningsåtgärder på utvärderingsområdet för chefer och annan personal hos Arbetsförmedlingen.

Aktiviteterna inriktades inte på tekniska frågor utan på betydelsen av utvärderingar, de viktigaste stegen i konsekvensbedömningar, betydelsen av att ha tillgång till bra uppgifter och kvalitetssäkring av utvärderingar. En webbplats inrättades där utbildningsmaterial laddades upp.

De viktigaste resultaten av projektet var att kunskapen om utvärderingar hos chefer inom Arbetsförmedlingen förbättrades. Generellt skapades en öppnare inställning till konsekvensbedömningar och projektet utmärktes av deltagarnas samarbetsvilliga inställning. Enligt den person som intervjuades är "projektets effekter på antalet genomförda konsekvensbedömningar inte så märkbara ännu". Vissa framsteg har dock gjorts med tanke på att inga kontrafaktiska konsekvensbedömningar över huvud taget genomfördes under programperioden 2007–2013 i Sverige.

Utvärderingen av projektet Ung framtid, där man använde en randomiserad metod, är ett konkret exempel i praktiken. Chefer hos arbetsförmedlingen deltog i utvärderingsprocessen⁹⁶.

Yrkesgemenskap

Det är viktigt att utveckla yrkesgemenskaper inom medlemsländerna, med utrymme för diskussioner om metoder för kontrafaktiska konsekvensbedömningar och erfarenhetsutbyte. Yrkesgemenskaper bidrar även till ömsesidigt stöd och lärande och till att upprätthålla kvalitetsnormerna. En effektiv strategi kan vara att knyta kontakter med relevanta utvärderingsorganisationer och uppmuntra dem att främja kontrafaktiska konsekvensbedömningar av ESF+-insatser i samband med utbildningar, branschkonferenser eller seminarier och informationsmöten.

Utveckla yrkesgemenskaper

Erfarenhetsutbyte

Kommissionen anser att konsekvensbedömningarna av ESF+-insatser bör bli noggrannare och kontrafaktiska konsekvensbedömningar rekommenderades allmänt under programperioden 2014–2020⁹⁷. Vi konstaterar att kontrafaktiska konsekvensbedömningar har blivit vanligare, men ett antal förvaltande myndigheter och medlemsländer

Utnyttja befintliga forum

⁹⁶ Informationen baseras på en intervju med en tjänsteman hos Arbetsförmedlingen.

⁹⁷ Enligt bilaga XI till förordningen om gemensamma bestämmelser för programperioden 2014–2020 (förordning (EU) nr 1303/2013) ska det finnas "ett verkningfullt system av resultatindikatorer för att mäta framstegen mot resultaten och för att bedöma genomslaget". Detta krav ingick i förhandsvillkoren för programperioden 2014–2020.

använder dem fortfarande endast i begränsad utsträckning. Erfarenhetsutbyte om metoder för kontrafaktiska konsekvensbedömningar är en av de mest effektiva lösningarna för att bygga upp kapacitet och stödja spridningen av kontrafaktiska konsekvensbedömningar i EU. För detta ändamål bör befintliga forum för ömsesidigt lärande utnyttjas, såsom sakkunnighetsbedömningar av politiska åtgärder för sysselsättning och social inkludering och praktikgemenskaper inom ramen för ESF+. Initiativ från enskilda förvaltande myndigheter eller medlemsländer, till exempel internationella konferenser och seminarier, kan också främja erfarenhetsutbytet.

4.3. Lösa rättsliga hinder

Ett av de största problemen för forskare som utför kontrafaktiska konsekvensbedömningar i medlemsländerna är att få åtkomst till de uppgifter de behöver. Forskarna möter ofta rättsliga hinder som syftar till att skydda identiteten för de personer som ingår i datauppsättningarna. Såsom framgår av den ovanstående diskussionen om den allmänna dataskyddsförordningens bestämmelser kan dessa problem lösas genom breda reformer och avtal som gör det möjligt att fortlöpande och under kontrollerade former ge utvärderare tillgång till data.

Undanröja rättsliga hinder för tillgång till data

Analysversioner av administrativa datauppsättningar skulle regelbundet kunna konstrueras på grundval av administrativa data från medlemsländernas myndigheter. Uppgifterna kan sedan dokumenteras och förvaras i ett arkiv med kontrollerad åtkomst. Godkända leverantörer skulle få tillstånd att hämta uppgifter från sådana dataägare. Om det inte är möjligt att inhämta uttryckligt samtycke av identifieras uppgifterna fullständigt med krypterade personliga identifierare för att efterleva dataskyddsförordningens bestämmelser. Sådana dataägare har skapats i flera länder. Om tillgång till data fortfarande är ett problem på grund av skiljande tolkningar av integritetsreglerna bör nationella initiativ inledas på statlig nivå för att främja avtal och system som kan leverera data för forskningsändamål på relativt kort tid.

Skapa analytiska datauppsättningar

Om det fortfarande råder oro över sekretessen avseende personuppgifter bör man överväga att inrätta datalaboratorier. På sätt skulle utvärderare som arbetar med administrativa datauppsättningar endast få åtkomst till register på säkra platser, med sträng övervakning och kontroll av åtkomsten till data. Uppgifterna måste behandlas och analyseras på plats. Endast analysresultat får lämna lokalen.

Skapa datalaboratorier

4.4. En framåtblickande strategi

En gemensam egenskap hos det fåtal kontrafaktiska konsekvensbedömningar som hittills har genomförts av ESF-finansierade insatser är att de har varit tillbakablickande snarare än framåtblickande. De utvärderingar av insatser som har lagts ut på expertutvärderare har inte varit utformade för att utvärderas. Ibland har det saknats eller funnits mycket lite planering för konsekvensbedömningar. Detta innebär i sin tur att utvärderarna själva måste konstruera datakällor, vilket är både tidsödande och kostnadskrävande. De får helt enkelt arbeta med de uppgifter som råkar

finnas tillgängliga i stället för datakällor som har konstruerats för att kunna konsekvensbedömas.

En framåtblickande strategi kan gå ut på att så tidigt som möjligt involvera utvärderarna i planeringen för en kontrafaktisk konsekvensbedömning. Insatserna (antingen nya eller befintliga) skulle då utformas med hänsyn till detta. Det krävs oftast mycket små ingrepp för att göra insatser mer lämpade för kontrafaktiska konsekvensbedömningar. Förhandsplanering är således ett måste för noggranna konsekvensbedömningar. Lämpligt utbildad personal bör göras delaktig eller externa experter bör anlitas redan i ett tidigt skede av insatsen eller när finansieringsbesluten fattas. En sådan strategi ger följande fördelar:

- Nödvändiga dokumentationsarrangemang kan integreras i genomförandet av program och insatser.
- Nödvändiga datakällor kan identifieras i ett tidigt skede. Åtkomst- och dataskyddsfrågor kan lösas i god tid.
- Insamlingen av referensdata kan förfinas och undersökningar kan göras vid behov.
- Praktiska frågor, t.ex. hur deltagarna rekryteras till insatserna, kan lösas så att rekryteringsprocesserna anpassas bättre till behovet av noggranna utvärderingar.

Utvärderare som är utbildade i metoder för kontrafaktiska konsekvensbedömningar (vare sig interna utvärderare hos de förvaltande myndigheterna eller externa experter) bör delta i processen för att utforma nya ESF+-insatser och i beslut om befintliga insatser. På så sätt kan konsekvensbedömningar planeras in redan i början av programperioden. En sådan strategi innebär dessutom stora fördelar för evidensbaserat beslutsfattande.

4.5. Utvidgning av tillämpningsområdet för kontrafaktiska konsekvensbedömningar

En sista och mycket viktig fråga rör kontrafaktiska konsekvensbedömningars politiska täckning. Det finns flera aspekter att beakta: de politiska åtgärder som omfattas, de analyserade utfallen och hur fullständig tolkningen av resultaten är.

Under programperioden 2014–2020 inriktades de allra flesta kontrafaktiska konsekvensbedömningar på aktiv arbetsmarknadspolitik och politikens effekter på sysselsättningen. Detta beror förmodligen på att det i allmänhet är lättare att få tillgång till administrativa data om sysselsättning och dessa data ger den viktigaste utfallsvariabeln för många aktiva arbetsmarknadspolitiska åtgärder.

Utbildningspolitik och politiska åtgärder för social inkludering har inte, eller mycket sällan, varit föremål för kontrafaktiska konsekvensbedömningar, trots att de är viktiga för ESF-strategin och i fördelningen av medel. Några av de få tillgängliga exemplen från Spanien, Polen och Portugal redovisas i ruta 20 nedan.

I framåtblickande strategier involveras utvärderarna redan från början

Mer fokus på kontrafaktiska konsekvensbedömningar i politiken för social inkludering och utbildning

En rad kombinerade initiativ krävs för att motverka bristen på kontrafaktiska konsekvensbedömningar och förbättra kunskapen om dessa viktiga politikområden:

Involvera ytterligare aktörer och utveckla lämpliga datakällor

- Kontrafaktiska konsekvensbedömningar av åtgärder för social inkludering eller utbildning bör planeras och förberedas för att säkerställa tillgång till data i tid. Nödvändiga samarbeten mellan olika aktörer bör inledas för att förbättra datatillgången. Behandlings- och kontrollgrupper bör tas fram redan i ett tidigt skede.
- Delaktighet för dataägare i utformningen av kontrafaktiska konsekvensbedömningar bör främjas. Förvaltningar med ansvar för utbildnings- och socialpolitik, om det är andra myndigheter än den förvaltande myndigheten, bör göras mer delaktiga i processen, med målet att skapa ett starkare och mer allmänt åtagande för utvärdering och tillgång till data.
- Institutionella avtal och programvaruverktyg för att anpassa administrativa data om sociala förhållanden och utbildning bör främjas så att dessa uppgifter kan användas i kontrafaktiska konsekvensbedömningar. Det nationella statistikkontoret bör delta i den processen, som en "överbryggande" institution och kontaktpunkt. Detta är en viktig aspekt, eftersom statistikkontoren har den kompetens som krävs och redan tar fram många administrativa datauppsättningar för att producera nationell statistik.

Insamling av relevanta uppgifter från deltagare i behandlings- och kontrollgrupper bör stödjas. Insamlingen bör ske redan i början av insatsen för att skapa en konsekvent datauppsättning med data både från tiden före och efterinsatsen. Om administrativa data inte finns tillgängliga eller inte är relevanta kan en riktad undersökning ge de uppgifter som behövs. Undersökningen bör inledas i början av insatsen och kan även omfatta kontrollgruppen för att säkerställa nödvändiga jämförelser.

Ruta 20 Exempel på utvärderingar på utbildningsområdet

ESF-insats mot tidiga skolavhopp i Asturien⁹⁸

I samband med en allmän utvärdering av det operativa programmet "Principado de Asturias ESF 2014–2020" genomfördes en särskild konsekvensanalys av programmet "Diversificación Curricular y de Mejora del Aprendizaje y del Rendimiento" (PMAR). PMAR är en ESF-åtgärd som genomfördes inom ramen för det integrerade programmet¹⁰ i med målet att motverka tidiga skolavhopp. Enligt PMAR delas eleverna upp i grupper (8–15 elever, men grupperna kan vara större eller mindre om särskilda omständigheter föreligger) för att undervisas i modersmål, samhällskunskap, naturvetenskap och matematik samt främmande språk (undervisningen i de övriga ämnena skedde i den vanliga klassen). Åtgärden genomfördes under läsåret 2016–2017 i gymnasieskolans (ESO) andra och tredje läsår. Alla elever som hade gått om minst ett år på någon nivå och elever som hade avslutat första läsåret i gymnasiet men inte hade betyg för att börja det andra läsåret var berättigade att delta. Det slutliga urvalet gjordes utifrån en individuell bedömning av eleverna (en akademisk och psykopedagogisk bedömning) som utfördes av lärarna.

Sammanlagt 1 053 elever deltog, 512 från det andra året i ESO och 541 från det tredje. Eleverna jämfördes sedan med andra elever med inlärnings svårigheter⁹⁹ för att välja jämförbara kontrollindivider i förhållande till de stödberättigande kriterier som gällde för eleverna i behandlingsgruppen (totalt 3 852 individer).

⁹⁸ Se [Diaz J.M. m.fl., 2019](#).

⁹⁹ Dessa identifierades som elever som under läsåret 2015–2016 gick första och andra året i gymnasiet, slutförde läsåret med underkänt i ett eller flera ämnen och hade gått om minst ett år under sin skolgång.

Man valde att tillämpa en matchningsmetod och använde administrativa data från SAUCE-databasen (*Sistema Informático para la Administración Unificada de Centros Educativos*) som tillhandahölls i oidentifierad form av utbildningsministeriet. Uppgifterna användes för att beräkna benägenhetspoängen¹⁰⁰ och mäta utfallsvariabeln, som var elever som hade godkända betyg för att fortsätta nästa läsår. Utvärderingen visade på en positiv och statistiskt signifikant effekt av insatserna: antagningsgraden för andraårseleverna i behandlingsgruppen var 18 procentandelar högre än för kontrollgruppen, medan skillnaden för tredjeårselever i behandlingsgruppen var omkring 16 procentenheter. Effekterna var mer positiva för flickor än för pojkar i båda grupperna.

ESF-insatser för att förbättra yrkesutbildningen i Podlasien¹⁰¹

Studien undersökte effekterna av programmet "Bra yrke – fantastiskt liv", som syftade till att främja yrkesutbildning i regionen Podlasien i Polen. Projektet riktades till grundskolepersonal, elever och föräldrar. Ett antal olika aktiviteter genomfördes både för institutionerna och människorna (skraddarsydd marknadsförings- och kommunikationsstrategier för skolor, utbildning för lärare, samarbete med företag och ledare i det lokala näringslivet, mentorsprogram för elever och föräldrar, yrkes- och karriärvägledning för elever). Projektet genomfördes 2017–2019.

Behandlingsgruppen bestod av studenter som deltog i det utvärderade projektet. Av 9 403 elever valdes 1 500 genom ett stratifierat slumpmässigt urval och 200 elever deltog i studiens undersökning. Uppgifter om de studenter som deltog i projektet hämtades från övervakningsdata. Kontrollgruppen angavs som elever med liknande kännetecken som elever i behandlingsgruppen, men som inte deltog i projektet. En kontrollgrupp som var ungefär dubbelt så stor som behandlingsgruppen planerades och valdes slumpmässigt från det nationella folkbokföringsystemet (PESEL – ett register där varje polsk medborgare har ett unikt nummer). Den planerade metoden kunde dock inte genomföras på grund av svårigheter orsakade av covid. Till slut användes "snöbollstekniken" (*chain-referral*) som primär urvalsmetod – 401 intervjuer genomfördes varav 384 användes i studien.

Data om utfallsvariablerna¹⁰² både för behandlings- och kontrollgrupperna samlades in via enkäter med personliga intervjuer, telefon- eller webbintervjuer. Studien använde viktning av benägenhetspoäng som kontrafaktisk metod för att bedöma genomslaget (endast på allmän nivå, inte för särskilda delgrupper). Den kontrafaktiska analys som genomfördes inom ramen för studien gav inga definitiva slutsatser. Resultatens statistiska signifikans var otillfredsställande.

Utvärdering av ett bidragssystem för högskolestudier riktat till mindre privilegierade studenter i Portugal¹⁰³

I en andra utvärdering, ännu ej offentliggjord, undersöks ett bidragssystem för studenter från låginkomsthushåll för att främja tillträde till högskoleutbildning och öka deltagandet. Bidragssystemet stöds av ESF (sedan programperioden 2007–2013) i regionerna Norte, Centro och Alentejo, och av staten i de övriga regionerna. Stödberättigade studenter är studenter med resurser under en viss tröskel (6 800–7 900 euro i inkomst per capita) och en minimipoäng under det föregående studieåret för studenter som inte är förstaårsstudenter). Samma regler för stödberättigande gäller för inskrivning i såväl offentliga som privata universitet. Bidragsbeloppet står i proportion till hushållsinkomsten.

Omkring 70 000 studenter har deltagit i programmet varje år sedan 2011. Utvärderingen omfattar dock perioden från 2012 och framåt, eftersom det finns tillgång till data från det året. Utvärderingen inriktas dessutom på studenter som ansöker om bidraget för första gången och går första året i en universitetsutbildning (kandidatexamen eller masterexamen). Stödberättigande avgörs av inkomstkriteriet¹⁰⁴.

¹⁰⁰ Följande variabler användes för att matcha de två grupperna: Studieresultat under föregående läsår, typ av skola (allmän eller privat), deltagande i andra kurser under föregående läsår (ja eller nej), kön, medelinkomst i bostadskommunen, bostadsort och födelseland (Spanien eller övriga världen).

¹⁰¹ Se [Zub M. m.fl., 2020](#).

¹⁰² Följande utfallsvariabler användes (baserat på enkätsvaren): inskrivning vid en teknisk högskola vid tidpunkten för studien, inskrivning vid en handelshögskola (grundläggande utbildning) vid tidpunkten för studien, inskrivning vid antingen en teknisk högskola eller en handelshögskola (grundläggande utbildning) vid tidpunkten för studien, ansökan till högskola, ansökan till teknisk högskola, ansökan till handelshögskola (grundläggande utbildning), ansökan till en yrkeshögskola (antingen en teknisk högskola eller en handelshögskola (grundläggande utbildning)), om skolan uppfyllde studentens förväntningar i mycket hög grad, om yrket verkligen motsvarade studentens intressen.

¹⁰³ Se Guthmuller S., Meroni E.C, ännu inte publicerad.

¹⁰⁴ En analys av hela urvalet av studenter planeras (inte bara förstaårsstudenterna), men har inte slutförts ännu.

Två administrativa datauppsättningar användes: uppgifter om elever som sökte 2012–2018 (från generaldirektoratet för högre utbildning (DGES)), som med en unik identifierare slogs ihop med en annan datauppsättning som innehöll information om akademiska karriärer och karriärutveckling (från generaldirektoratet för statistik och vetenskap (DGEEC)). Det slutliga analysurvalet bestod av 156 002 studenter, varav 130 602 mottog behandling och 25 400 inte. Med hjälp av inkomstgränsen (den löpande variabeln) kunde utvärderaren tillämpa en regressionsdiskontinuitetsdesign och flera utfallsvariabler mättes: på kort sikt, procentandelen studenter som fortfarande var inskrivna samt poäng i slutet av första året, på längre sikt, sannolikheten för att ta examen, slutlig poäng och antal år till examen.

Under den inledande fasen visade bidraget positiva effekter för inskrivningstal, men inte för antalet tagna poäng. På längre sikt tar fler studenter i behandlingsgruppen examen, på kortare tid och med högre poäng än kontrollgruppen. Effekterna var högre för manliga studenter bosatta i områden som omfattas av ESF samt för studenter inskrivna i masterprogram och vid statliga universitet.

Vi har redan konstaterat att metautvärderingar är mycket användbara och ger bra förklaringar. En allmän ökning av kontrafaktiska konsekvensbedömningar, vilket i sin tur skulle kunna öka antalet metautvärderingar, skulle vara ett tillskott i en gemensam debatt om specifika åtgärder som involverar flera förvaltande myndigheter och medlemsländer. Såsom nämns ovan kan den metodologiska komplexiteten hos dessa typer av utvärderingar lösas med bättre planering av kontrafaktiska konsekvensbedömningar på såväl nationell som internationell nivå.

Öka antalet metautvärderingar

När det gäller ESF-insatser mäts resultaten i kontrafaktiska konsekvensbedömningar vanligen med kvantitativa eller binära (ja/nej) variabler, till exempel inkomst eller sysselsättningsstatus. "Mjuka utfall", som handlar om deltagarnas självuppfattning och kapacitet, är dock viktiga mått för resultaten av insatser. "Anställbarhet", självkänsla och kompetens är till exempel grundläggande förutsättningar för att få jobb eller vara aktiv på arbetsmarknaden. I många fall kan analyser av mjuka utfall vara till ännu mer nytta än arbetslöshetssiffror eftersom de visar hur insatserna tillgodoser individuella behov.

Överväga "mjuka utfall"

Kontrafaktiska konsekvensbedömningar kan tillämpas på mjuka utfall på under förutsättning att uppgifter samlas in om situationen både före och efter insatsen för såväl behandlings- som kontrollgruppen enligt insatsens förändringsteori. Sådana undersökningar måste planeras redan i ett tidigt skede och omfatta alla berörda aktörer, särskilt stödmottagarna, som har direkt och löpande kontakt med deltagarna. Ett exempel på en kontrafaktisk konsekvensbedömning av "mjuka utfall" redovisas i ruta 21 nedan.

Ruta 21 Exempel på bedömning av effekter på "mjuka utfall" i Tyskland

Bidrar sysselsättningsåtgärder till långtidsarbetslösa sociala integration och välbefinnande?¹⁰⁵

Studien undersöker effekterna av sysselsättningsåtgärder som riktades till en grupp av utsatta personer, i det här fallet långtidsarbetslösa, som hade fått socialbidrag i minst fyra år och hade hälsoproblem eller barn eller båda. Programmet genomfördes 2015–2018 och hade omkring 20 000 deltagare. Åtgärden subventionerade reguljära anställningsavtal på upp till 36 månader med 30 timmars arbetstid per vecka (främst hos offentliga arbetsgivare eller välgörenhetsorganisationer).

Eftersom det uttryckliga syftet med den 36 månader långa sysselsättningsåtgärden var att främja social integration av målgrupperna försökte författarna bedöma genomslaget utifrån dessa förutsättningar.

För analysen användes en omfattande datauppsättning med administrativa data (tyska "integrerade

¹⁰⁵ Se [Ivanov B. m.fl., 2020](#).

anställningsbiografier”) och en enkät bland av programdeltagarna och kontrollindivider. Enkäten gjorde det möjligt att samla in information om subjektiva mått på livskvalitet: tillfredsställelse med tillvaron, psykisk hälsa, sociala samhörighetskänsla och social status. För att samla in information om dessa mått användes Likertskalor som överensstämde med andra undersökningar (t.ex. den nationella PASS-undersökningen) i syfte att få fram jämförbara resultat. De viktigaste stegen i undersökningen var följande: Personer som mottog och inte mottog behandling identifierades och matchades med varandra i de administrativa uppgifterna. De resulterande behandlings- och kontrollgrupperna identifierades (i tre omgångar). Den slutliga datauppsättningen byggdes upp genom att uppgifterna rensades och fall med saknade uppgifter togs bort. Urvalet bestod av 2 531 matchade par i fas 1, 1 191 i fas 2 och 450 i fas 3 (av ett totalt urval på 62 000 personer, varav 12 400 behandlade och 49 600 kontroller).

Författarna använde benägenhetspoäng och mätte effekterna 7, 18 och 29 månader efter det att deltagarna gått in i programmet. De fann att insatserna hade positiva effekter på välbefinnandet, men i olika grad, tillfredsställelsen med livet ökade till exempel avsevärt, medan social status endast förbättrades något. Effekterna tenderade dock att minska under programmets gång och detta kan eventuellt förklaras av en ökning av både antalet deltagare som lämnade programmet och av antalet kontroldeltagare som hittade arbete. De positiva effekterna var större för de mest utsatta personerna.

Som vi framhåller i början av denna vägledning är kontrafaktiska konsekvensbedömningar ett kraftfullt verktyg för att bedöma ”vad” en ESF+-insats har åstadkommit. För att förstå ”varför” eller ”hur” de uppmätta effekterna har uppstått krävs andra instrument.

*Kontrafaktiska
konsekvensbedömningar
integrerade med
teoretiska utvärderingar*

I vissa fall, där insatserna är välkända och redan har utvärderats grundligt, är det tillräckligt med en relativt begränsad analys av genomförandet och några intervjuer med stödmottagarna för att identifiera de huvudsakliga orsakerna till de mätta effekterna. För mer komplexa, mindre kända eller innovativa insatser är det ofta nödvändigt att integrera kontrafaktiska konsekvensbedömningar med andra utvärderingar.

En utvärdering av genomförandet eller processen före den kontrafaktiska konsekvensbedömningen kan visa hur insatsen fortskred, problem i genomförandet och hur de olika aktörerna samarbetade för att insatsen skulle bli lyckad. Resultaten av sådana utvärderingar kan sedan användas i den kontrafaktiska konsekvensbedömningen och visa vilka effekter som bör undersökas, hur kontrollgruppen bör vara sammansatt i förhållande till deltagarnas kännetecken och den rätta tidpunkten för en kontrafaktisk konsekvensbedömning.

En teoretisk utvärdering kan utföras parallellt med en kontrafaktisk konsekvensbedömning. De två utvärderingarna bidrar ömsesidigt till varandra: den teoretiska utvärderingen ger förklaringar av de kontextuella, sociala och individuella mekanismer som påverkade och framkallade effekterna, medan den kontrafaktiska konsekvensbedömningen beräknar nettoeffekter och visar om insatsen gav väsentliga utfall. Vi rekommenderar att teoretiska utvärderingar och kontrafaktiska konsekvensbedömningar integreras i effektutvärderingar för att ge en heltäckande översikt av effekterna och deras orsaker.

Ordlistor

Förkortningar

Allmänna dataskyddsförordningen	Europaparlamentets och rådets förordning (EU) 2016/679
Eruf	Europeiska regionala utvecklingsfonden
ESF	Europeiska socialfonden
EU	Europeiska unionen
Förordningen om gemensamma bestämmelser	Europaparlamentets och rådets förordning (EU) 2021/1060
GD EMPL	Generaldirektoratet för sysselsättning, socialpolitik och inkludering
GD REGIO	Generaldirektoratet för regionalpolitik
IEB	integrerade anställningsbiografier (<i>Integrated Employment Biographies</i>)
JLD	<i>Jobseeker longitudinal dataset</i>
PSM	benägenhetspoäng (<i>Propensity score matching</i>)
RCT	randomiserad kontrollerad studie (<i>Randomised controlled trial</i>)
RDD	regressionsdiskontinuitetsdesign (<i>Regression discontinuity design</i>)

Definitioner

Term	Definitioner
Behandlingsgrupp	En grupp av personer, företag eller andra enheter som gynnas eller omfattas av en insats (de åtgärder som erbjuds eller de faktiska mottagarna av åtgärden).
Benägenhetspoäng (<i>Propensity score matching, PSM</i>)	PSM innebär att en statistisk modell uppskattas för hela urvalet (behandlingsgruppen och potentiella kontroller), vilket ger en uppskattad benägenhet att delta för varje individ eller företag – oavsett om de faktiskt deltog eller ej. Behandlade individer eller företag matchas sedan på grundval av benägenhetspoängen, antingen med en individ eller ett företag som inte behandlats, eller med många obehandlade individer eller företag.
Deltagare	En deltagare är en fysisk person som drar direkt nytta av en insats utan att vara ansvarig för att inleda eller genomföra insatsen. När det gäller ESF avser deltagare personer som får stöd via ESF-insatser.
Effektivitet	Effektivitet definieras som en given output som uppnås till minsta kostnad, eller motsvarande, att outputen maximeras vid en viss resursnivå. Effektiviteten kan fastställas genom en kostnads-nyttoanalys eller en kostnadseffektivitetsanalys.
Extern utvärdering	En utvärdering som utförs av en extern och oberoende utvärderare på grundval av ett anbudsförfarande.
Genomslag	När det gäller kontrafaktiska konsekvensbedömningar avser genomslag nettoeffekter, som definieras som skillnaden mellan genomsnittlig behandling och kontrafaktiska resultat. I denna vägledning används begreppen "genomslag" eller "effekter" och "nettoeffekter" utan åtskillnad.
Icke-randomiserad eller kvasiexperimentell design	Strategier för kontrafaktiska konsekvensbedömningar där kontrollgrupperna konstrueras med andra metoder än randomisering.
Insats	Avser generellt insatser i ESF:s operativa program eller projekt som samfinansieras av ESF.
Intern utvärdering	En utvärdering som görs internt, dvs. som beställs direkt från en oberoende offentlig institution eller enhet (från den förvaltande myndigheten eller det förmedlande organet) utan anbudsförfarande, eller som genomförs i form av en utvidgad övervaknings- och analysprocess.
Kontrafaktisk analys	En resultatmässig jämförelse mellan vad som faktiskt skedde och vad som skulle kunna ha skett i avsaknad av insatsen. Eftersom skillnaden mellan faktiska och kontrafaktiska resultat avgör insatsens kausala effekt omfattar en kontrafaktisk analys alla metoder för att bedöma omfattningen av en observerad förändring som kan tillskrivas den utvärderade insatsen.
Kontrafaktisk konsekvensbedömning	En typ av konsekvensbedömning som syftar till att identifiera de kausala effekterna av en insats. De kontrafaktiska resultaten uppskattas och dras därefter av från de genomsnittliga observerade resultaten för de enheter som mottar behandling. Uppskattningar av kontrafaktiska resultat görs vanligen av en kontrollgrupp som väljs ut noggrant för att vara så lik behandlingsgruppen som möjligt.
Kontrollgrupp	En grupp av personer, företag eller andra enheter som liknar behandlingsgruppen i så stor utsträckning som möjligt men inte mottar behandling, och från vilka kontrafaktiska uppskattningar av resultat fås fram. I strikt bemärkelse avser begreppet "kontrollgrupp" experimentella förhållanden (se "Randomisering" nedan), medan begreppet "jämförelsegrupp" avser kvasiexperimentella förhållanden, men i praktiken används båda termerna utan åtskillnad.
Matchning	En metod där insatsurval och kontrollurval matchas med varandra på grundval av deras observerade kännetecken.
Metoden med instrumentvariabler (IV)	Enligt denna metod bör urvalet till behandling åtminstone delvis avgöras av en exogen faktor (eller ett exogent instrument) som inte har någon annan koppling till resultaten än behandlingen. Den exogena faktorn påverkar således deltagandet, men har ingen direkt inverkan på resultaten.
Output	Output anses utgöra allt som produceras/levereras direkt till följd av genomförandet av en ESF-insats, mätt i fysiska eller monetära enheter. Output mäts huvudsakligen i antal

Term	Definitioner
	personer och enheter som får stöd, levererade varor och tjänster samt genomförda projekt.
Processutvärdering	Processutvärderingar inriktas på programmets genomförande, omfattande men inte begränsat till, hur tjänsterna levereras, skillnader mellan den avsedda målgruppen och den målgrupp som får tillgång till stödet, tillträde till programmet och förvaltningsmetoder.
Randomisering	En metod där individer i en målgrupp slumpmässigt fördelas till ett antal behandlings- eller kontrollförhållanden. Randomisering säkerställer att grupperna är statistiskt likvärdiga i alla avseenden vid tidpunkten för randomiseringen.
Referensdata	Uppgifter om de variabler som mäts innan en enhet (en individ eller ett företag) exponeras för en insats. Ofta samlas uppgifter in om situationen före behandling både för behandlings- och kontrollgruppen för att mäta insatsens resultat.
Regressionsdiskontinuitetsdesign (RDD)	Denna metod användas när tillträdet till insatsen avgörs av en brytpunkt längs en kontinuerlig klassningsskala eller ett kontinuerligt mått. Metoden utgår från det faktum att de personer som ligger precis vid brytpunkten har mycket stora likheter, förutom att de som ligger på ena sidan av brytpunkten deltar medan de på den andra sidan inte gör det. Resultat för de personer som befinner sig över och under brytpunkten kan jämföras för att beräkna insatsens effekter.
Relevans	Relevans avser lämpligheten hos de uttryckliga målen för en insats med avseende på de socioekonomiska problem som insatsen är avsedd att lösa ¹⁰⁶ .
Resultat	De effekter som en insats har på deltagare eller enheter, till exempel när det gäller deras sysselsättningssituation, inkomster, poäng i standardiserade utbildningsprov, vinster osv. Effekterna kan mätas på kort sikt eller på längre sikt. Inom ramen för ESF+ anses kortsiktiga mätningar av resultatindikatorer utgöra mätningar omedelbart efter deltagandet (fyra veckor), medan mätningar på längre sikt är mätningar sex månader efter insatsen ¹⁰⁷ . När det gäller konsekvensanalyser innebär mätningar på längre sikt ofta en längre period, 24 månader eller mer efter insatsen.
Skillnaden-mellan-skillnader (DID)	I sin enklaste form, dras skillnaden i ett resultat före och efter behandling i kontrollgruppen av från samma skillnad som observerats hos kontrollgruppen för att få fram en uppskattning av insatsens effekt. Effekter som beräknas på grundval av skillnaden-mellan-skillnader härrörs vanligen från en regressionsram.
Stödmottagare	Enligt artikel 2 i Europaparlamentets och rådets förordning (EU) 2021/1060 (förordningen om gemensamma bestämmelser) är en stödmottagare "ett offentligt eller privat organ, en enhet med eller utan ställning som juridisk person, eller en fysisk person, med ansvar för att inleda eller för att både inleda och genomföra insatser". När det gäller finansieringsinstrument (som inom ESF+ ibland används för t.ex. mikrokrediter till egenföretagare), det organ som genomför holdingfonden eller, när det inte finns någon holdingfondstruktur, det organ som genomför den särskilda fonden eller, när den förvaltande myndigheten förvaltar finansieringsinstrumentet, den förvaltande myndigheten.
Utfall	Sannolika eller uppnådda effekter av en insats output på kort och medellång sikt ¹⁰⁸ . Utfall betyder ungefär "resultat" men används oftare i konsekvensbedömningar.
Utvärderingsplan	Enligt artikel 44 i Europaparlamentets och rådets förordning (EU) 2021/1060 ska medlemsstaten eller den förvaltande myndigheten upprätta en utvärderingsplan som kan omfatta fler än ett program. Medlemsstaten eller den förvaltande myndigheten ska lägga fram utvärderingsplanen för övervakningskommittén senast ett år efter beslutet om godkännande av programmet.
Ändamålsenlighet	Avser "måluppfyllelse" och utvärderas genom en jämförelse mellan vad som åstadkommits och vad som planerats (eller med utgångsläget), eller genom en jämförelse av vad som observeras efter det att åtgärden genomförts med vad som skulle ha hänt utan åtgärden (kontrafaktisk situation).

¹⁰⁶ Se [Europeiska kommissionen, 2013](#).

¹⁰⁷ Se [Europeiska kommissionen, 2021a](#) och [Europeiska kommissionen, 2018](#).

¹⁰⁸ Se [OECD, 2010](#).

Litteraturförteckning

- Ashenfelter, O. (1978). *Estimating the effect of training programmes on earnings*, *Review of Economics and Statistics*, 6, s. 47–57. <https://www.istor.org/stable/1924332?origin=crossref>
- Axdorph E, Egebark J., Lundström T., Özcan G. (2019) *Effekter av förstärkta förmedlingsinsatser för unga arbetssökande – Resultat från utvärderingen av Ung framtid*. Arbetsförmedlingen
- Baran J. m.fl. (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. m.fl. (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. m.fl. (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. m.fl. (2018a). *Badanie efektów wsparcia zrealizowanego na rzecz osób młodych w ramach Programu Operacyjnego Wiedza Edukacja Rozwój. I 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". I *Politica Economica/Journal of Economic Policy*, band XXXIV(3), s. 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), s. 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. m.fl. (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. m.fl. (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. m.fl. (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. m.fl. (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. m.fl. (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. m.fl. (2018). *Vocational Training for Unemployed Youth in Latvia: Evidence from a Regression Discontinuity Design* IZA DP nr 11870. <http://ftp.iza.org/dp11870.pdf>
- Bratu C. m.fl. (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). "Evaluating what Works for whom in active Labour market policies". *European Journal of Social Security*, 17(4), s. 436–452. <https://doi.org/10.1177/138826271501700403>
- Bryson, A., Dorsett, R. och Purdon, S. (2002). *The use of propensity score matching in the evaluation of active labour market policies*, Department for Work and Pensions, arbetsdokument nr 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%20use%20of%20propensity%20score%20matching%20in%20the%20evaluation%20of%20active%20labour%20market%20policies.pdf)
- Caliendo M., Kopeinig S., (2008). "Some practical Guidance for the implementation of propensity score matching". *Journal of Economic Surveys*, band 22, upplaga 1, s. 31–72. <https://doi.org/10.1111/j.1467-6419.2007.00527.x>
- Caliendo, M., Mahlstedt, R., & Mitnik, O. A. (2017). *Unobservable, but unimportant? The relevance of usually unobserved variables for the evaluation of labour market policies*. *Labour Economics*, 46, s. 14–25. <https://doi.org/10.1016/j.labeco.2017.02.001>
- Card, D., Ibarra, P. och Villa, J. M. (2011). *Building in an evaluation component for active labour market programs: a practitioner's guide*, diskussionsunderlag nr 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*, band 16, upplaga 3, juni 2018, s. 894–931, <https://doi.org/10.1093/jeea/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. m.fl. (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)* för Europeiska kommissionen, generaldirektoratet för sysselsättning, socialpolitik och inkludering. <https://op.europa.eu/en/publication-detail/-/publication/d0c1a558-077d-11eb-a511-01aa75ed71a1/language-en/format-PDF/source-173162502>
- Europeiska kommissionen (2007): *Indicative Guidelines on evaluation methods: evaluation during the programming period*. Arbetsdokument nr 5. GD Regionalpolitik. https://ec.europa.eu/regional_policy/sources/docoffic/2007/working/wd5_ongoing_en.pdf
- Europeiska kommissionen (2013): *EVALSED: The resource for the evaluation of Socio-Economic Development*. Uppdaterad version. https://ec.europa.eu/regional_policy/sources/docgener/evaluation/guide/guide_evalsed.pdf
- Europeiska unionen (2017) *Better Regulation Toolbox, Tool #47 on Evaluation criteria and*

questions https://ec.europa.eu/info/sites/default/files/file_import/better-regulation-toolbox-47_en_0.pdf

Europeiska 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>

Europeiska 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>

Europeiska 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>

Europeiska 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>

Europeiska kommissionen (2021). Europaparlamentets och rådets förordning (EU) 2021/1060 av den 24 juni 2021 om fastställande av gemensamma bestämmelser för Europeiska regionala utvecklingsfonden, Europeiska socialfonden+, Sammanhållningsfonden, Fonden för en rättvis omställning och Europeiska havs-, fiskeri- och vattenbruksfonden samt finansiella regler för dessa och för Asyl-, migrations- och integrationsfonden, Fonden för inre säkerhet samt instrumentet för ekonomiskt stöd för gränsförvaltning och viseringspolitik.

Europeiska 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)*, för Europeiska kommissionen, generaldirektoratet för sysselsättning, socialpolitik och inkludering.

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

Frolich, M (2004) "Programme evaluation with multiple treatments", *Journal of Economic Surveys*, 18(2), s. 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.* Gemensamma forskningscentrumets tekniska rapport. Ännu inte offentliggjord.

Hägglund, P. (2006). *A description of three randomised experiments in Swedish labour market policy*, Institutet för arbetsmarknads- och utbildningspolitisk utvärdering, rapport 2006:4.

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

Heckman, J. J., Ichimura, H., Smith, J. och Todd, P. (1998) *Characterizing selection bias using experimental data.* I *Econometrica*, band 66, nr 5. s. 1017–1098.

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

Hunger K. och 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), s. 945–960 DOI: 10.1080/01621459.1986.10478354.

<http://people.umass.edu/~stanek/pdf/cause/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 och Eurocentre (2020) *Study supporting the 2020 evaluation of promoting social inclusion, combating poverty and any discrimination by the European Social Fund (Thematic Objective 9)* för Europeiska kommissionen, generaldirektoratet för sysselsättning, socialpolitik och inkludering. <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). *Misure 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/sv/publication-detail/-/publication/84cc9eb9-b33d-11e9-9d01-01aa75ed71a1>
- Ivanov B. m.fl., 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. m.fl., 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>
- Kluge, J., Lehmann, H. och Schmidt, C. M. (2008). *Disentangling treatment effects of Active Labor Market Policies: The role of labor force status sequences*. Labour Economics, 15(6), s. 1270–1295. <https://doi.org/10.1016/j.labeco.2007.12.002>
- Krug, G och Stephan, G. (2011) *Is contracting-out intensified placement services more effective than in-house production? Evidence from a randomized field experiment*, LASER diskussionsunderlag – underlag nr 5, <http://doku.iab.de/externe/2011/k110912303.pdf>
- Lammers M. och 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*, utarbetad för den sjätte europeiska konferensen om utvärdering av sammanhållningspolitiken, Warszawa, 30 november. https://ec.europa.eu/regional_policy/archive/conferences/evaluation2009/abstracts/martini.doc
- Morgan, S. L. och Winship, C. (2014) *Counterfactual and causal inference: Methods and*

principles for social research. Andra upplagan. Cambridge och 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. m.fl (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. och 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. och Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalised causal inference*, Boston, USA: Houghton Mifflin Company.

Stern E. m.fl. (2012). *Broadening the range of designs and methods for impact evaluations*, Department for International Development of the UK, diskussionsunderlag 38. <https://assets.publishing.service.gov.uk/media/57a08a6740f0b6497400059e/DFIDWorkingPaper38.pdf>

WK Kellogg Foundation (2004). *Logic Model Development Guide*. <https://www.wkcf.org/resource-directory/resources/2004/01/logic-model-development-guide>

Zub M. m.fl. (2020). *Ocena wpływu wsparcia RPOWP na popularyzację szkolnictwa zawodowego w województwie podlaskim. Raport końcowy*.

<http://files.evaluationhelpdesk.eu/Evaluations/PLE359.pdf>

Bilagor

Bilaga 1. Läs mer

Nedan anges förslag på litteratur för anställda vid förvaltande myndigheter som är intresserade av att gå in djupare på de frågor som tas upp i den här vägledningen. Det finns ett rikt utbud av litteratur om utvärdering. Denna förteckning är avsedd att lyfta fram tillförlitliga och allmänna diskussioner som direkt ger användbar information för planering av kontrafaktiska konsekvensbedömningar. Efter varje citat följer en kort beskrivning av de viktigaste källorna.

Allmän utvärdering

- Rossi, Peter H., Mark W. Lipsey och Gary T. Henry. (2018). *Evaluation: A Systematic Approach*. Åttonde upplagan. Thousand Oaks, CA: SAGE Publications.

The classic textbook on evaluation practice and methods. Innehåller metoder och exempel.

- Gertler, Paul J., Martinez, Sebastian. Premand, Patrick, Rawlings, Laura B., Vermeersch, Christel M. J. (2016). *Impact Evaluation in Practice*, andra upplagan. Washington, DC: Interamerikanska utvecklingsbanken och Världsbanken. © World Bank.
<https://openknowledge.worldbank.org/handle/10986/25030> (finns på engelska, portugisiska och spanska).

Liksom denna vägledning inleds handboken med en klassisk utvärdering (RCT) och diskuterar därefter alternativ. Handledningen är skriven för programförvaltare i låginkomstländer, men diskussionen är relevant och direkt tillämplig på situationen i EU-medlemsländerna.

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

Liksom Gertler m.fl. (2016) som nämns ovan innehåller denna handbok, som visserligen handlar om konsekvensbedömningar av insatser i låg- och medelinkomstländer, en utförlig diskussion om allmänna metoder för kontrafaktiska konsekvensbedömningar. Mer teknisk än Gertler m.fl. (2016), men innehåller ändå många praktiska insikter som kan vara av intresse för dig som ESF-programförvaltare.

- Europeiska 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>

I denna vägledning från kommissionen presenteras nya och mer avancerade metoder för kontrafaktisk konsekvensbedömning, t.ex. dynamisk matchning och syntetiska kontroller.

- Csillag Marton, Kreko Judi och Scharle Agota. (2020). *Counterfactual evaluation of youth employment policies*. Utarbetad inom ramen för projektet "Youth Employment PartnerSHIP".
<http://yepartnership.ibs.org.pl/content/uploads/2021/02/Methodological-guide.pdf> (finns på engelska, spanska, ungerska (översättning pågår), italienska och polska).

Detta är en stegvis introduktion till kontrafaktiska konsekvensbedömningar av arbetsmarknadsåtgärder för unga, med fokus på administrativa data. Frågeställningarna är uppbyggda kring praktiska problem som uppstått i fyra kontrafaktiska konsekvensbedömningar avseende stöd för anställning av ungdomar i Spanien, Ungern, Italien och Polen.

- HM Treasury (Förenade kungariket) (2020). *The Magenta Book: Guidance for evaluation*. London: myndigheten. <https://www.gov.uk/government/publications/the-magenta-book>

”Magentaboken” beskriver olika utvärderingsmetoder. Dessa dokument är intressanta som exempel på myndigheters interna utvärderingspolicy.

Randomiserade kontrollerade studier

- White H. (2013) ”An introduction to the use of randomised control trials to evaluate development interventions”, *Journal of Development Effectiveness*, 5:1, s. 30–49, DOI: 10.1080/19439342.2013.764652
<https://www.tandfonline.com/doi/pdf/10.1080/19439342.2013.764652>
- White, H., Sabarwal S. och T. de Hoop, (2014). *Randomized Controlled Trials (RCTs), Methodological Briefs: Impact Evaluation 7*, Unicef Office of Research, Florens.
https://www.unicef-irc.org/publications/pdf/brief_7_randomized_controlled_trials_eng.pdf

Två icke-tekniska introduktioner till logiken hos av randomiserade kontrollerade studier, med en diskussion av flera utformningar och en kritisk granskning av denna metod.

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

En stegvis vägledning om hur man utformar och genomför randomiserade kontrollerade studier för utvärderingar av sociala program. Handboken bygger på konkreta randomiserade kontrollerade studier som utförts av Abdul Latif Jameel Poverty Action Lab.

Skillnaden-mellan-skillnader

- Card, David, Pablo Ibararán och Juan Miguel Villa. (2011). *Building in an Evaluation Component for Active Labor Market Programs: A Practitioner’s Guide*. IZA diskussionsunderlag nr 6085. Bonn: Forschungsinstitut zur Zukunft der Arbeit.
<http://ftp.iza.org/dp6085.pdf>
Contrasts Diff-in-Diff with RCT.
- Lechner, M. (2011) *The Estimation of Causal Effects by Difference-in-Difference Methods. Foundations and Trends in Econometrics*. Band 4, nr 3 (2010), s. 165–224 DOI: 10.1561/0800000014. https://michael-lechner.eu/ml_pdf/journals/2011_Lechner_DiD_2011_ECO%20403%20Lechner_darf%20aufs%20Netz.pdf

Denna forskningsrapport innehåller en djupgående diskussion om metoden skillnaden-mellan-skillnader (DID) och tar upp några av de viktigaste frågorna när det gäller genomförandet. Förlängningar av metoden, t.ex. icke-linjära tillämpningar och benägenhetspoäng med användning av skillnaden-mellan-skillnader, diskuteras också.

- Fredriksson A. och de Oliveira G. M. (2019) ”Impact evaluation using Difference-in-Differences”. *RAUSP Management Journal*, band 54 nr 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 översikt av DID-metoder med praktiska rekommendationer.

- Card, David och 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), s. 774–775. <https://davidcard.berkeley.edu/papers/njmin-aer.pdf>

Det klassiska exemplet på tillämpning av en skillnaden-mellan-skillnader-teknik.

Instrumentvariabler

- Morgan, Stephen L. och Christopher Winship. (2014). *Counterfactuals and Causal Inference:*

Methods and Principles for Social Research. Andra upplagan. Cambridge och New York: Cambridge University Press.

Detta är en något teknisk granskning av metoder för kontrafaktisk konsekvensbedömning, med sociologisk terminologi. Kapitel 9, "Instrumental Variable Estimators of Causal Effects" (s. 291–324) ger en översikt av logiken och förfarandena för uppskattning av instrumentvariabler.

- Kuhn, Andreas, Jean-Philippe Wuellrich och Josef Zweimüller. 2010. *Fatal Attraction? Access to Early Retirement and Mortality*. IZA diskussionsunderlag nr 5160. Bonn: Forschungsinstitut zur Zukunft der Arbeit. <http://ftp.iza.org/dp5160.pdf>

Här används regionala variationer i förändringen av pensionsåldern i Österrike som en instrumentvariabel i en studie av effekterna av förtidspensionering på arbetstagarnas hälsa.

- Galiani, S., Rossi, M. A. och Schargrotsky, E. (2011). "Conscription 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>

Nyskapande och illustrativ användning av ett slumpmässigt uttagningslotteri (för militärtjänstgöring) som instrumentvariabel. En mycket lättläst studie, som är prototypisk för en forskningsgren där lotteriprocesser används som instrument.

Matchning

- Heinrich, Carolyn, Alessandro Maffioli och Gonzalo Vázquez. 2010. *A Primer for Applying Propensity Score Matching. Impact-Evaluation Guidelines Technical Notes* nr IDB-TN-161. Washington: Interamerikanska utvecklingsbanken. <https://publications.iadb.org/publications/english/document/A-Primer-for-Applying-Propensity-Score-Matching.pdf>

Liksom vägledningen om regressionsdiskontinuitet nedan riktas denna vägledning till kunniga utvärderingsansvariga.

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

En klassisk text om de val som måste göras i genomförandet av matchningar med benägenhetspoäng (uppskattningar, matchande algoritmer, kvalitetsbedömning av matchningen, de uppskattade behandlingseffekternas känslighet).

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

Liknar den föregående artikeln.

Regressionsdiskontinuitetsdesign (*Regression discontinuity design*)

- Jacob, Robin, Pei Zhu, Marie-Andrée Somers och 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

Mycket lättillgänglig och grundlig diskussion om metoden för regressionsdiskontinuitet som innehåller en noggrant utvald bibliografi.

- Lee, D. S. och Lemieux T. (2009), *Regression discontinuity designs in economics*, NBER si 14723, National Bureau of Economic Research, Cambridge, MA, 2009. http://www.nber.org/papers/w14723.pdf?new_window=1

Rapporten är en slags användarhandledning om regressionsdiskontinuitet, med en

beskrivning av metodens logik och exempel på olika metoder för uppskattning av regressionsdiskontinuitetsdesign.

Bilaga 2. Förslag till upplägg för onlinekurser om kontrafaktiska konsekvensbedömningar

En introduktionskurs om kontrafaktiska konsekvensbedömningar för förvaltande myndigheter och tjänstemän kan omfatta följande:

- Introduktion till utvärderingsmetoder: Processutvärderingar (varför de är viktiga och vilket samband de har med kontrafaktiska konsekvensbedömningar) och effektutvärderingar (kontrafaktiska och teoretiska metoder).
- Vad är en kontrafaktisk konsekvensbedömning? Hur bör utvärderingsfrågorna utformas i kontrafaktiska konsekvensbedömningar? För vilka utvärderingskriterier ger kontrafaktiska konsekvensbedömningar de bästa belägen? Varför är kontrafaktiska konsekvensbedömningar viktiga?
- Hur fungerar de? (behandlingsgrupper och kontrollgrupper, jämförelser "före och efter", inferensprinciper).
- Översikt av metoder för kontrafaktiska konsekvensbedömningar (egenskaper, varje metods starka och svaga sidor).
 - o Randomiserad kontrollerad studie.
 - o Matchning.
 - o Skillnaden-mellan-skillnader.
 - o Regressionsdiskontinuitetsdesign.
 - o Instrumentvariabler.
 - o Andra möjliga metoder (t.ex. tekniker för rumsliga analyser).
- Datakrav.
 - o Krav för behandlingsgrupper och kontrollgrupper.
 - o Möjliga källor (ESF-övervakning, administrativa datauppsättningar, undersökningar).
 - o Allmänna dataskyddsförordningens regler.
- Översikt av genomförandet.
 - o Planering av kontrafaktiska konsekvensbedömningar (utvärderingsplan, genomförbarhet).
 - o Beställning av kontrafaktiska konsekvensbedömningar (specifikationer, urvalskriterier och metoder).
 - o Hantering av kontrafaktiska konsekvensbedömningar (kontroll av resultat och samarbete med utvärderaren).
 - o Spridning av resultat av kontrafaktiska konsekvensbedömningar typer av målgrupper och kommunikationskanaler).

En kurs som utformas enligt ovan bör även omfatta praktiska exempel som kan diskuteras vid seminarier (två–tre dagar långa). En lösning för en kurs som denna är att tillämpa en problembaserad inlärningsmetod och använda en politisk åtgärd i ett program som ett konkret fall att utveckla.

Bilaga 3. Kontrafaktiska konsekvensbedömningar – exempel som nämns i vägledningen

Tabell 6 Kännetecken hos de kontrafaktiska konsekvensbedömningar som nämns som exempel i vägledningen

Land	Titel – år	Utvärderat politikområdet och observerad period	Metod	Behandlingsgrupp och kontrollgrupp – definition och storlek	Utfallsvariabler	Observation stidens längd	Uppgifter
Sverige	Effekter av intensifierat stöd till unga jobbsökande – resultat från utvärderingen av Ung framtid (2019)	Intensifierat stöd från Arbetsförmedlingen till jobbsökande. Juni 2017–maj 2018	Randomiserad kontrollerad studie (RCT)	Behandlingsgrupp: Ungdomar 18–24 år som har valts slumpmässigt som mottagare av intensifierat stöd. Kontrollgrupp: Ungdomar 18–24 år som har valts slumpmässigt som mottagare av ordinarie stöd. Behandlingsgrupp: 2 792 – Kontrollgrupp: 1 897	– Andel arbetslösa – % av utflöde till arbete – Antal dagar i arbetslöshet	12 månader efter inträde	– Register från arbetsförmedlingen för utfallsvariabler. – Undersökning som utförs av handläggare vid arbetsförmedlingen och mellanhänder för att mäta stödets intensitet.
Italien (Marche)*	Effekter av ESF-insatser som finansierades under programperioden 2014–2020 för långtidsarbetslösa i regionen Marche (2020)	Praktikplatser, yrkeserfarenhet, jobbstipendier, utbildningskuponger för arbetslösa (bland annat långtidsarbetslösa). 2017–juli 2019	Benägenhet spoäng (<i>Propensity score matching, PSM</i>)	Behandlingsgrupp: Behandlade långtidsarbetslösa. Kontrollgrupp: Långtidsarbetslösa som var inskrivna hos den offentliga arbetsförmedlingen under perioden 2016–2018. Behandlingsgrupp: 526 för praktikplatser, 1 058 för jobbstipendier, 236 för yrkeserfarenhet hos kommuner och 241 för utbildningskuponger – Kontrollgrupp: 77 255.	– Sannolikhet att få anställning en viss tid efter insatserna – Sannolikhet att få fast anställning – Antal arbetsdagar under en period efter insatserna	6, 9, 12, 15 och 18 månader efter inträde i insatserna	– ESF-övervakningsdata för behandlingsgruppen. – Administrativa data från arbetsförmedlingen om registrerade personer. – Administrativa data om anställningsavtal som ingåtts i regionen (<i>Comunicazioni Obbligatorie</i> – COB).
Polen (Lubelskie)*	Utvärdering av åtgärder för äldre arbetstagare i Lubelskie som finansierades via ESF under programperioden 2007–2013 (2015)	Flera typer av aktiva arbetsmarknadspolitiska åtgärder. 2007–2013	PSM + skillnaden-i-skillnader (DID).	Behandlingsgrupp: Urval av arbetslösa äldre än 50 år som fick stöd Kontrollgrupp: Arbetslösa över 50 år som inte deltog i någon åtgärd och var inskrivna hos en regional arbetsförmedling vid den tidpunkt då behandlingsgruppen deltog i det utvärderade programmet. Behandlingsgrupp: 145 – Kontrollgrupp: 67102.	– Sannolikhet att upphöra att vara registrerad som arbetslöshet	6, 12 och 18 månader efter inträde i insatsen	– Administrativa uppgifter från den offentliga arbetsförmedlingen.

Land	Titel – år	Utvärderat politikområde och observerad period	Metod	Behandlingsgrupp och kontrollgrupp – definition och storlek	Utfallsvariabler	Observation tidens längd	Uppgifter
Lettland	Yrkesutbildning för arbetslösa ungdomar i Lettland: Bevis från en regressionsdiskontinuitetsdesign 2018.	Yrkesutbildning som riktades till unga i åldern 15–29 år som varken arbetar eller studerar. 2014–2015	Regressions diskontinuitetsdesign (RDD)	Behandlingsgrupp: Arbetslösa som deltog i yrkesutbildning inom ett år från registreringsdatumet. Kontrollgrupp: Arbetslösa registrerade mellan juni 2013 och december 2015 som inte deltog i någon åtgärd. Behandlingsgrupp: 898 – Kontrollgrupp: 10 717.	– Sannolikhet för att personen fick anställning en viss tid efter insatserna – Inkomst	12–36 månader efter det att deltagare gått in i programmet	– Administrativa data från den lettiska statliga arbetsförmedlingen – information om både deltagare och icke-deltagare som var anmälda som arbetslösa på vissa datum. Administrativa data från det statliga skattekontoret. Information om information om sysselsättningsstatus på olika datum samt inkomstnivåer.
Tyskland	Sysselsättningseffekter av språkkurser för arbetslösa invandrare (2021).	Språkkurser (samt yrkeserfarenhet för några av de behandlade individerna) för utlandsfödda jobbsökande. 2014	Instrumentvariabler (IV)	Behandlingsgrupp: Utlandsfödda deltagare som inledde språkundervisning för yrkesändamål under 2014. Kontrollgrupp: Slumpmässigt urval av icke-deltagare, med minst en period utan tyskt medborgarskap och som inte deltog i några aktiva arbetsmarknadspolitiska åtgärder eller integrationskurser. Behandlingsgrupp: 8 968 – Kontrollgrupp: 26 463.	– Sannolikhet att få anställning (fast anställning, fast heltidsanställning, fast anställning mer än sex månader, anställning med inkomst över fattigdomsgränsen, kvalificerad anställning) – Daglig och ackumulerad inkomst från fast anställning – Ackumulerat antal dagar i fast anställning	24 månader efter inträde i insatsen	– "Integrerade anställningsbiografier" (IEB) för identifiering av båda grupperna och för utfallsvariablerna. – Uppgifter som samlats in i efterhand under möten med jobbsökande och handläggare (<i>Werdegangshistorie</i> , WGH) för mer information om perioder av egenföretagande eller föräldraledighet och utbildning och anställning utomlands.
Italien (Trento)	Är utbildningsprogram värda kostnaden? Belägg från en kostnadsnyttoanalys (2018).	Yrkesutbildningskurser för arbetslösa. 2010–2011	PSM.	Behandlingsgrupp: Arbetslösa deltagare i yrkesutbildningskurser, finansierade med provinsiella medel och av ESF. Kontrollgrupp: Personer som var anmälda som arbetslösa hos den lokala arbetsförmedlingen vid startdatumet för utbildningskurserna. Behandlingsgrupp: 818 (provinskurser, 114 (ESF-kurser) – Kontrollgrupp: 17 236 (provinskurser) 1 152 (ESF-kurser).	– Sannolikhet att få anställning – Inkomster	3, 6, 12, 18, 24 och 36 månader efter inträde i insatserna	– ESF-övervakningsdata för behandlingsgruppen. – Administrativa data från arbetsförmedlingen om registrerade personer. – Administrativa data om anställningsavtal som ingåtts i regionen (<i>Comunicazioni Obbligatorie</i> – COB). – Administrativa data om deklarationsuppgifter från INPS. – Administrativa data från provinsens utbildningsdepartement.

EUROPEISKA KOMMISSIONEN

Land	Titel – år	Utvärderat politikområde och observerad period	Metod	Behandlingsgrupp och kontrollgrupp – definition och storlek	Utfallsvariabler	Observationstidens längd	Uppgifter
Tyskland	Bidrar sysselsättningsåtgärder till långtidsarbetslösa sociala integration och välbefinnande? (2020)	Sysselsättningsåtgärder (subventionerade jobb i upp till 36 månader) för utsatta personer (långtidsarbetslösa). 2015–juni 2017	PSM.	Behandlingsgrupp: urval av deltagare som besvarade alla frågor i frågeformuläret. Kontrollgrupp: Icke-deltagare som var inskrivna hos arbetsförmedlingen, hade samma stödberättigande kriterier som kontrollgruppen och besvarade hela enkäten. Behandlingsgrupp: 2 531 i första omgången, 1 191 i andra omgången, 450 i tredje omgången – Kontrollgrupp: som ovan.	– Sannolikhet att få anställning – Subjektiva mått av tillfredsställelse med tillvaron, psykisk hälsa, sociala samhörighetskänsla och social status	7, 18 och 29 månader efter inträde i programmet	– "Integrerade anställningsbiografier" (IEB) för identifiering av båda grupperna och för utfallsvariablerna. – Undersökning för de subjektiva utfallsvariablerna (mjuka utfall).
Spanien (Asturien)*	Utvärdering av ESF-OP Asturien (2019)	– Undervisningsmetoder för grupper av elever med risk för tidigt skolavhopp (andra och tredje året i gymnasiet). 2016–2017	PSM.	Behandlingsgrupp: elever som deltog i insatserna. Kontrollgrupp: elever som var jämförbara när det gällde kriterierna för stödberättigande. Behandlingsgrupp: 1 053 – Kontrollgrupp: 3 852.	– Sannolikhet för att bli uppflyttad till nästa årskurs	12 månader efter inträde	– Administrativa data om utbildning: SAUCE-databasen (utbildningsministeriet).
Portugal	Utvärdering av ett bidragssystem för högskolestudier riktat till mindre privilegierade studenter i Portugal (2020) – ännu ej offentliggjord	Bidrag till högskolestudenter från låginkomsthushåll. 2021–2018	RDD.	Behandlingsgrupp: studenter som söker till första läsåret i en Master- eller kandidatutbildning. Kontrollgrupp: studenter som inte är bidragsberättigade eftersom deras inkomst är högre än den fastställda tröskeln. Behandlingsgrupp: 130 602 – Kontrollgrupp: 25 400.	– Sannolikhet för inskrivning i slutet av det första året – Antalet tagna poäng vid utgången av det första året – Sannolikhet för utexaminering, – Antal år fram till examen – Slutligt betyg vid examen	12–48 månader från bidraget	– Administrativa data om studenter som ansökte om bidraget (DGES). – Administrativa data med information om akademiska karriärer och karriärutveckling (DGEEC).
Polen (Podlasien)*	Bedömning av effekterna av det operativa programmet för främjande av yrkesutbildning i Podlasien 2014–2020 (2020)	ESF-programmet "Bra yrke – fantastiskt liv" – främjande av yrkesutbildning med inriktning på grundskolan. 2017–2019	PSM.	Behandlingsgrupp: urval av deltagande elever. Kontrollgrupp: elever med liknande kännetecken som eleverna i behandlingsgruppen, men som inte deltog i projektet. Behandlingsgrupp: 200 – Kontrollgrupp: 384	– Inskrivning vid en teknisk högskola vid tidpunkten för studien – Inskrivning vid en handelshögskola (grundläggande utbildning) vid tidpunkten för studien – Inskrivning vid antingen en teknisk högskola eller en handelshögskola (grundläggande utbildning) vid tidpunkten för studien – ansökan till högskola, teknisk högskola, handelshögskola	12–36 månader efter insatserna	– Övervakningsdata. – Det nationella folkbokföringssystemet (PESEL) för urval av kontrollgruppen. – Undersökning för att mäta utfallsvariablerna.

EUROPEISKA KOMMISSIONEN

Land	Titel – år	Utvärderat politikområde och observerad period	Metod	Behandlingsgrupp och kontrollgrupp – definition och storlek	Utfallsvariabler	Observationstidens längd	Uppgifter
					(grundläggande utbildning), yrkeshögskola		
Italien (Piemonte)*	Sysselsättningseffekt av kuponger för arbetsförmedlingstjänster (<i>Buono servizi lavoro</i> – BSL, finansierade via ESF-OP Piemonte, 2014–2020) (2019)	Kuponger för arbetsförmedlingstjänster (yrkesrådgivning, utbildning, förberedelse) för utsatta arbetslösa personer som varit arbetslösa i minst sex månader. 2018	PSM.	Behandlingsgrupp: Arbetslösa som mottagit behandling. Kontrollgrupp: Arbetslösa personer inskrivna vid arbetsförmedlingen med liknande kännetecken (30 år och arbetslösa i sex månader). Behandlingsgrupp: 8 125 – Kontrollgrupp: 130 000.	– Sannolikhet att få anställning – Sannolikhet att få fast anställning	6, 12 respektive 16 månader från insatserna	– ESF-övervakningsdata för behandlingsgruppen. – Administrativa data från arbetsförmedlingen om registrerade personer. – Administrativa data om anställningsavtal som ingåtts i regionen (<i>Comunicazioni Obbligatorie</i> – COB).
Tyskland*	Utvärdering av programmet för integration av långtidsarbetslösa i Tyskland under programperioden 2014–2020 (2019)	Integrationsåtgärder för långtidsarbetslösa som varit arbetslösa i minst 24 månader (subventionerade arbeten). 2015–2017	DID (avsikten att motta behandling uppskattades).	Behandlingsgrupp: Långtidsarbetslösa som var potentiellt stödberättigade under programperioden (från 2015). Kontrollgrupp: Personer som uppfyllde kriterierna för stödberättigande före genomförandet av programmet (2010–2012) Målgrupp och kontrollgrupp: 134 515 personer för 237 874 arbetslöshetsperioder (35 % för målgruppen och 65 % för kontrollgruppen).	– Sannolikhet att få anställning (anställning med socialbidrag)	24 månader från insatserna	– Administrativa övervakningsdata från Federala administrativa myndigheten (ZUWES-databasen). – ”Integrerade anställningsbiografier” (IEB) för utfallsvariabler.
Tyskland (Baden-Württemberg)*	Andra utvärderingsrapporten om specifikt mål A1.1 i ESF-OP Baden-Württemberg, 2014–2020 (2020)	Åtgärder för integration av arbetslösa, särskilt långtidsarbetslösa och andra utsatta personer. Mars 2016–december 2017	PSM.	Behandlingsgrupp: deltagare som identifierats i IEB-databasen Kontrollgrupp: personer som uppfyllde kriterierna för stödberättigande under samma period. Behandlingsgrupp: 1 578 (av totalt 1 800) – Kontrollgrupp: i.u.	– Sannolikhet att få anställning (anställning med socialbidrag) – Sannolikhet att fortsätta att motta arbetslöshetsförmåner	15 månader från insatserna	– Övervakningsdata. – ”Integrerade anställningsbiografier” (IEB) för utfallsvariabler.

Anmärkning: * avser utvärderingar som har identifierats i hjälpcentralen.

KONTAKTA EU

Personligt besök

Det finns hundratals Europa direkt-kontor i hela EU. Hitta ditt närmaste kontor på https://europa.eu/european-union/contact_sv

Ring eller mejla

Europa direkt svarar på dina frågor om EU. Kontakta tjänsten på något av följande sätt:

- Ring det avgiftsfria numret 00 800 6 7 8 9 10 11 (en del operatörer kan ta betalt för samtalet)
- Ring det vanliga numret +32 22999696.
- Skriv ett meddelande via kontaktformuläret https://europa.eu/european-union/contact_sv

INFORMATION OM EU

Online

På Europa-portalen finns det information om EU på alla officiella EU-språk https://europa.eu/european-union/index_sv

EU-publikationer

Du kan ladda ner eller beställa EU-publikationer på <https://op.europa.eu/sv/publications>. Om du behöver flera exemplar av en kostnadsfri publikation kan du kontakta Europa direkt eller ditt lokala informationskontor (se https://europa.eu/european-union/contact_sv).

EU-lagstiftning och därmed sammanhängande handlingar

Tillgång till rättsinformation från EU, inbegripet all EU-lagstiftning sedan 1952 på samtliga officiella språk, ges via Eurlex: <http://eur-lex.europa.eu>

Öppna data från EU

På EU:s portal för öppna data (<http://data.europa.eu/euodp/sv>) finns dataserier från EU. De får laddas ned och användas fritt för kommersiella och andra ändamål.



Publications Office
of the European Union