



A2009:002

Metoder för att utvärdera den regionala tillväxtpolitiken

Lärdomar och råd för utvärderare och uppdragsgivare

Torbjörn Danell, Thomas Ericson, Gustav Hansson,
Gustav Lidén, Jonas Månsson och Fredrik Svensson

Metoder för att utvärdera den regionala tillväxtpolitiken

Lärdomar och råd för utvärderare och uppdragsgivare

Torbjörn Danell
Thomas Ericson
Gustav Hansson
Gustav Lidén
Jonas Månsson
Fredrik Svensson

ITPS, Institutet för tillväxtpolitiska studier
Studentplan 3, 831 40 Östersund
Telefon 063 16 66 00
Telefax 063 16 66 01
E-post info@itps.se
www.itps.se
ISSN 1652-0483

För ytterligare information kontakta Gustav Hansson
Telefon 063 16 66 39
E-post gustav.hansson@itps.se

Förord

ITPS har i 2008 års regleringsbrev (uppdrag nr 6) fått i uppdrag att utveckla metoder för att utvärdera den regionala tillväxtpolitiken utifrån ett ekonomiskt, miljömässigt och socialt perspektiv, med fokus på effekter och måluppfyllelse.¹ Målgruppen för den här rapporten är Regeringskansliet samt utvärderare och beställare av utvärderingar, lokalt, regionalt och nationellt.

Inom ramen för uppdraget har ITPS även uppdragits att utveckla indikatorer för uppföljning av regional utveckling samt att ta fram kunskapsöversikter av hållbar regional tillväxt samt politikutvecklingen inom området. Dessa rapporter är publicerade som enskilda rapporter.² Den här rapporten utgör slutrapport av uppdraget och drar således erfarenheter och lärdomar från de ovan nämnda rapporterna. Rapporten beskriver även lärdomar och erfarenheter från ITPS metautvärdering av strukturfonderna, samt ITPS effektutvärdering av det regionala utvecklingsbidraget.³ Rapporten beskriver också olika angreppssätt för att utvärdera komplexa program och horisontell samverkan.

Rapporten har skrivits av Torbjörn Danell, Thomas Ericson, Gustav Hansson, Gustav Lidén, Jonas Månsson och Fredrik Svensson. Torbjörn Danell har varit projektledare i ett initialt skede, medan Gustav Hansson har varit projektledare under genomförandet av rapporten. Ett stort tack för kommentarer och förslag riktas till Lennart Delander (Växjö universitet) samt till Örjan Hag och Maria Eriksson (Näringsdepartementet).

Östersund, februari 2009

Brita Saxton

Generaldirektör

¹ Se bilaga 1.

² ITPS (2008a, b, c, d), ITPS (2009).

³ ITPS (2008e), Gadd et al (2008).

Innehåll

Sammanfattning	7
1 Inledning	9
2 Allmänna utgångspunkter för att utvärdera den regionala tillväxtpolitiken	11
2.1 Inledning	11
2.2 Utvärderingens plats i processen	11
2.3 Olika typer av utvärderingar	13
2.4 Förutsättningar för tillförlitliga utvärderingar	19
2.5 Att utvärdera åtgärder inom den regionala tillväxtpolitiken	21
2.6 Sammanfattning	23
3 Indikatorer	25
3.1 Inledning	25
3.2 Vad är en indikator?	25
3.3 Ett lämpligt urvalsförfarande för indikatorer	26
3.4 Sammanfattning	28
4 Flernivåanalys och det regionala utvecklingsbidraget	31
4.1 Inledning	31
4.2 Att utvärdera det regionala utvecklingsbidraget.....	31
4.3 Flernivåanalys.....	32
4.4 Sammanfattning	35
5 Lärdomar från metautvärdering	37
5.1 Inledning	37
5.2 Metautvärdering	37
5.3 Erfarenheter från ITPS metautvärdering av strukturfondsprogrammen	39
5.4 Sammanfattning	44
6 Metoder för att utvärdera komplexa program	45
6.1 Vad utmärker komplexa program?	45
6.2 Komplexa aspekter på den regionala tillväxtpolitiken	45
6.3 Utvärdering av komplexa program	47
6.4 Sammanfattning	49
7 Metoder för att utvärdera horisontell samverkan	51
7.1 Inledning	51
7.2 Relevanta begrepp och former av horisontell samverkan.....	51
7.3 Konsekvenser för utvärdering.....	53
7.4 Två exempel på horisontell samverkan och uppföljning och utvärdering av dessa	55
7.5 Sammanfattning	57
8 Sammanfattande slutsatser och rekommendationer	59
Referenser	62
Bilaga 1: Regleringsbrevsuppdraget	66
Bilaga 2: Utvärderingar utgivna av ITPS	67

Sammanfattning

Den här rapporten beskriver olika metoder och angreppssätt för att utvärdera den regionala tillväxtpolitiken. Rapporten utgör slutrapport för ITPS regleringsbrevsuppdrag nummer 6, 2008. Målgruppen för rapporten är utvärderare samt beställare av utvärderingar. Rapporten består av ett teorikapitel samt fem kapitel med fem olika metoder eller angreppssätt speciellt lämpliga för utvärderingar av den regionala tillväxtpolitiken.

Utvärderingar är i allmänhet inriktade mot att antingen studera *utfallet* eller *genomförandet* av en åtgärd. Utvärderingar som studerar utfallet delas in i mål-, effekt- och effektivitetsutvärderingar. En målutvärdering svarar på frågan om målen har uppfyllts; en effektutvärdering om måluppfyllelsen är en effekt av åtgärden; en effektivitetsutvärdering om de uppnådda effekterna värderas högre än kostnaderna för åtgärden, medan slutligen en genomförandeutvärdering besvarar frågan om hur själva genomförandet kan ha påverkat måluppfyllelse, effekt eller effektivitet. En utfallsutvärdering kan med fördel kombineras med en genomförandeutvärdering, eftersom den senare kan vara till hjälp att tolka utfallsutvärderingens resultat.

Utvärderingar av åtgärder inom den regionala tillväxtpolitiken bör ta särskild hänsyn till den regionala heterogeniteten. Rapporten ger därför rekommendationer beträffande användandet av regionala indikatorer samt rörande en statistisk metod – flernivåanalys. Indikatorer och flernivåanalys är båda användbara verktyg för att ta hänsyn till den regionala heterogeniteten.

En viktig aspekt i utvärderingssammanhang är att kvalitetssäkra arbetet med utvärderingar. Rapporten ger därför även rekommendationer som avser utnyttjandet av ”metautvärdering” som metod för sådan kvalitetssäkring. Metautvärderingar är en bra metod för att kvalitetssäkra arbetet med utvärdering av program eftersom metoden innefattar ett tydligt lärandeperspektiv så att misstag inte upprepas.

Tillförlitliga och användbara utvärderingar förutsätter en välgjord datainsamling av åtgärdens utfall, samt en dokumentation av genomförandet. En adekvat datainsamling förutsätter att de individer, företag osv. vilka åtgärden gäller, kan identifieras och följas över tid. Om inte nödvändig information finns i befintliga register, kan data behöva samlas in redan innan åtgärden har implementerats. Ett sätt att minska risken för svårutvärderade åtgärder är att utvärderaren kopplas in tidigt i processen för att garantera dokumentationen och därmed även utvärderbarheten.

Då åtgärdsprogram går mot att bli allt mer komplexa, med många olika aktörer och målgrupper, beskriver även rapporten olika angreppssätt för att utvärdera komplexa program och horisontell samverkan. Att utvärdera sådana program kräver en öppenhet mot nya metoder, samt en kombination av både kvalitativa och kvantitativa metoder.

1 Inledning

Den regionala tillväxtpolitiken består av en stor mängd olika typer av åtgärder såsom till exempel regionala projekt, företagsstöd, samt stöd till kommersiell service. Att adekvat kunna utvärdera den regionala tillväxtpolitiken är naturligtvis av stor betydelse och ITPS fick därför år 2008 i uppdrag att utveckla metoder för sådan utvärdering.⁴ Målgruppen för den här rapporten är Regeringskansliet samt utvärderare och beställare av utvärderingar, lokalt, regionalt och nationellt. Rapporten syftar till att ge stöd till utvärderare av regional tillväxtpolitik samt att ge beställare av utvärderingar en djupare förståelse för utvärderingar och utvärderingsmetoder.

Vetenskapligt arbete utgår i regel från en specifik frågeställning och därefter väljs den metod som i just det fallet är mest lämplig för att besvara den eller de aktuella frågorna. Att föreslå att en viss bestämd utvärderingsmetod alltid bör användas är därför naturligtvis inte lämpligt eller realistiskt. Den här rapporten beskriver lärdomar och erfarenheter från det utvärderingsarbete som ITPS har bedrivit och som är tillämpbara vid de specifika förutsättningar som en utvärderare av regional tillväxtpolitik står inför. Rapporten består av ett teoretiskt kapitel samt fem metodkapitel.

Kapitel 2 innehåller en teoriöversikt om utvärderingar i allmänhet samt en orientering om vilka speciella utmaningar som kan finnas vid utvärderingar av regional tillväxtpolitik. Detta kapitel definierar och beskriver skillnader mellan olika typer av utvärderingar samt vad som behövs för att utföra tillförlitliga och användbara utvärderingar.

Efter Kapitel 2 följer fem kapitel innehållande beskrivningar och lärdomar av olika metoder och angreppssätt för utvärderingar. Dessa kapitel kan grupperas i två kategorier: Kapitel 3, 4 och 5 beskriver metoder som nyligen använts av ITPS och erfarenheter av dessa. Kapitel 6 och 7 beskriver möjliga angreppssätt för att utvärdera komplexa program som kännetecknas av vertikal och horisontell samverkan. Kapitel 3 till 7 är skrivna oberoende av varandra, vilket gör att den läsare som endast vill fördjupa sig inom något eller några av kapitlen kan göra så utan att missa sammanhanget.

Kapitel 3 behandlar framtagandet av indikatorer för regional utveckling. Kapitlet är en vidareutveckling av ITPS indikatorrapport ”Indikatorer för uppföljning av regional utveckling”⁵ och beskriver erfarenheter och lärdomar från arbetet med att skapa sådana indikatorer. Kapitlet beskriver indikatorers för- och nackdelar när de används vid utvärderingar samt ger kvalitetskriterier för hur indikatorer ska väljas ut för att vara användbara och tillförlitliga.

Kapitel 4 ger en kort introduktion till flernivåanalys. Flernivåanalys är ett verktyg för att ta tillvara och kontrollera för den regionala heterogeniteten i analysen. Kapitlet bygger på lärdomar från en av ITPS utvärderingar (en effektutvärdering av det regionala utvecklingsbidraget⁶). *Kapitel 5* beskriver hur metoden ”metautvärdering” kan användas för att kvalitetssäkra utvärderingsarbete. En metautvärdering kan innebära utvärdering av andra utvärderingar och därmed ge uppslag till hur utvärderingsarbete kan förbättras. Metautvärdering kan emellertid också innebära utvärdering av resultat från andra utvärderingar. Kapitlet bygger på lärdomar och erfarenheter från ITPS metautvärdering av

⁴ Se bilaga 1 för en beskrivning av uppdraget.

⁵ ITPS (2008a).

⁶ Gadd et al (2008).

de utvärderingar som genomförts av strukturfonderna samt från en upprepad utvärdering baserad på data från de tidigare utvärderingarna – en sekundäranalys⁷.

Ett ofta rekommenderat arbetssätt är att olika aktörer, organisationer och myndigheter ska samverka för att på så sätt effektivisera sitt arbete och uppnå synergieffekter. Åtgärdsprogram går därför mot att bli allt mer komplexa, med många olika aktörer och målgrupper. Det kan då röra sig om en samverkan vertikalt (det vill säga en samverkan mellan aktörer på olika nivåer) eller horisontellt (det vill säga en samverkan mellan aktörer på samma nivå). *Kapitel 6* beskriver olika angreppssätt för att utvärdera komplexa program med vertikal samverkan, medan *Kapitel 7* beskriver olika sätt att utvärdera horisontell samverkan. Då kapitlet inte bygger på någon av ITPS utförd utvärdering är framställningen mer resonerande och diskuterande.

Rapportens avslutande *Kapitel 8* innehåller sammanfattande slutsatser och rekommendationer.

⁷ ITPS (2008e).

2 Allmänna utgångspunkter för att utvärdera den regionala tillväxtpolitiken

2.1 Inledning

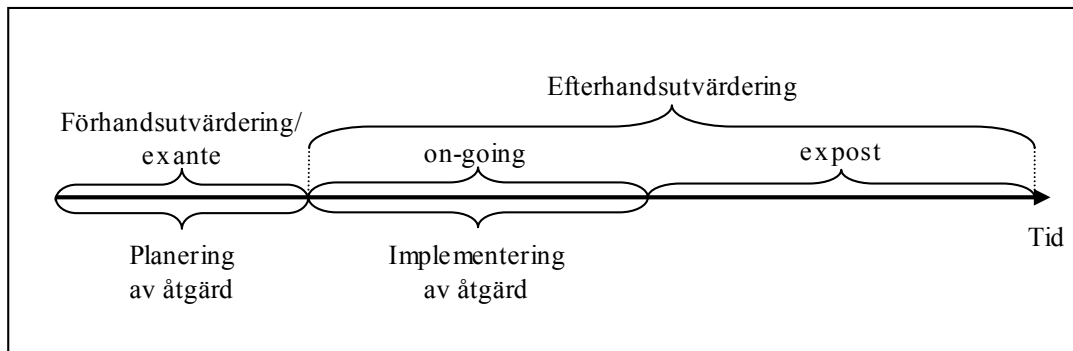
Målet med en utvärdering är att den ska kunna användas som beslutsunderlag. Det är därför viktigt att utvärderingar vilar på en solid vetenskaplig grund. Ett vetenskapligt arbete utgår i regel från en eller flera specifika frågor, varefter den metod väljs som i det aktuella fallet är den mest lämpliga för att besvara frågeställningen. Att föreslå en generell utvärderingsmetod som alltid bör användas är därför naturligtvis inte lämpligt eller realistiskt. Det finns däremot allmänna utgångspunkter som bör beaktas för att utvärderingar ska bli tillförlitliga och användbara. Syftet med det här kapitlet är att definiera och beskriva skillnader mellan olika typer av utvärderingar samt att beskriva vad som behövs för att utföra utvärderingar av god kvalitet vad gäller tillförlitlighet och användbarhet.

De olika åtgärderna inom den regionala tillväxtpolitikens kan grovt delas in i tre undergrupper: strategier, program och projekt. Även om dessa kan te sig helt olika kan samma allmänna utgångspunkter för utvärderingar tillämpas. Den regionala tillväxtpolitiken ställs emellertid inför specifika problem såsom till exempel regional heterogenitet som vid en utvärdering kan behöva beaktas särskilt. Det här kapitlet är därför disponerat att först allmänt redogöra för olika angreppssätt för utvärderingar av tillväxtpolitiska åtgärder och att därefter beskriva hur man i utvärderingar kan beakta den regionala dimensionen.

2.2 Utvärderingens plats i processen

En utvärdering kan ske vid olika tidpunkter i en åtgärdsprocess: före åtgärdens start (förhandsutvärdering/ex ante-utvärdering), under tiden åtgärden genomförs (on-going utvärdering) eller efter avslutad åtgärd (ex post-utvärdering), se Figur 2-1.

Figur 2-1 Utvärderingens plats i processen.



En *förhandsutvärdering* utförs vid planeringen av åtgärden och fram tills dess att åtgärden börjar. En utvärdering som genomförs under tiden som en åtgärd genomförs benämns *on-going*, medan en utvärdering som genomförs efter att åtgärden avslutats benämns *ex post-utvärdering*. Inom utvärderingslitteraturen används även begreppet *efterhandsutvärdering* vilket innefattar all typ av utvärdering som utförts efter att åtgärden har påbörjats.

2.2.1 Förhandsutvärderingar

En förhands- eller ex ante-utvärdering genomförs innan en åtgärd har börjat och är tänkt att ge underlag för beslut *om* och *hur* åtgärden ska genomföras.

När en åtgärd planeras är det viktigt att tänka igenom vilka resultat som förväntas uppnås och redan i det skedet är det lämpligt att involvera utvärderare och fundera över hur projektet ska utvärderas. Utvärderarens uppgift är då bland annat att granska målformuleringar i planen. Vid utformningen av målformuleringar samt vid en bedömning av dessa kan de så kallade SMART-kriterierna för målformuleringar vara användbara. Det rör sig om kriterier som i ”största möjliga utsträckning” ska följas vid målformuleringar i budgetpropositionen och i regleringsbrev. Enligt SMART-kriterierna ska målen vara:⁸

- (S) Specifika** De ska tydligt anges vad som ska uppnås
- (M) Mätbara** De ska vara möjliga att följa upp med hjälp av resultatindikatorer, nyckeltal eller liknande
- (A) Accepterade** De ska vara accepterade och uppfattas som relevanta av dem som ska genomföra den aktuella verksamheten
- (R) Realistiska** De ska vara möjliga att uppnå
- (T) Tidsbestämda** Tidpunkten då målen ska vara uppnådda ska anges

I kommersiella verksamheter spelar förhandsutvärderingar av enskilda projekt, till exempel i form av investeringskalkyler, ofta en mycket mer framträdande roll än vad de gör i offentliga verksamheter. Det torde i rätt hög grad bero på att såväl målen som beslutsprocesserna kan vara förhållandevis komplicerade inom offentliga verksamheter vilket gör det svårare där att åstadkomma beslutsunderlag som täcker alla de aspekter som alla deltagare i de offentliga beslutsprocesserna fäster avseende vid. ”Förhandsutvärderingar” i offentliga verksamheter får därför ofta formen av överväganden och bedömningar som återfinns i exempelvis utredningar, propositioner, yttranden, muntliga diskussioner och överläggningar vilka dock sällan benämns som utvärderingar.

Ibland bygger förhandsutvärderingar på resultat av tidigare utvärderingar eller forskningsresultat. Det sägs då att den planerade åtgärden är ”evidensbaserad”. I andra fall genomförs omfattande förhandsutvärderingar där det studeras hur åtgärden och förväntade resultat hänger ihop. En utvärdering av det slaget sägs utvärdera logiken i åtgärden.

En annan användbar metod för förhandsutvärderingar är olika typer av simuleringar. I simuleringsmodeller används befintliga data som genom diverse antaganden, baserade på tidigare inhämtad kunskap, programmeras att simulera tänkta utfall och scenarion. I en tidigare rapport från ITPS⁹ ges exempel på olika simuleringsmodeller som används i Sverige och hur dessa skulle kunna justeras för att ta tillvara den regionala dimensionen.

⁸ *Ds 2000:63 sid. 53.*

⁹ *ITPS (2007c).*

2.2.2 Efterhandsutvärderingar

Efterhandsutvärderingar är ett samlingsnamn på de utvärderingar som utförs efter att åtgärden påbörjats. Efterhandsutvärderingar delas upp i on-going och ex post-utvärderingar.

Efterhands- och ex post-utvärdering likställs ofta vilket kan vara olämpligt. Som tidigare nämnts är en efterhandsutvärdering en utvärdering som genomförs någon gång efter att åtgärden *startat*, medan en ex post-utvärdering genomförs efter att åtgärden har *avslutats*. I de fall då en åtgärd genomförs på försök under en relativt kort tidsperiod är det ofta rimligt att inte förvänta sig några resultat förrän försöket har slutförts varför en utvärdering i det fallet kommer att ske ex post. Innebär däremot åtgärden en ”permanent” förändring så kommer det definitionsmässigt inte att finnas något som kan benämnas ex post vilket då betyder att en utvärdering kommer att bli av efterhandskaraktär. När det rör sig om program och projekt som i och för sig är tidsbegränsade men som sträcker sig över en längre tid kan det vara både lämpligt och relevant att göra utvärderingar av utfallet även innan åtgärderna är slutförda. Att utvärdera åtgärder som inte har något slutdatum kan vara svårt. En möjlighet är att studera målvariablerna före och efter politikförändringen (till exempel med hjälp av metoden regression discontinuity).

Ex post-utvärderingar kan göras vid olika tidpunkter efter avslutad åtgärd. Då åtgärden antas ge omedelbara och hållbara resultat kan mätningar göras i direkt anslutning till åtgärdens avslut. Till exempel om syftet är att öka den formella kunskapen inom något område, så kommer de kunskaper som förvärvats genom åtgärden troligen att vara desamma på kort och lång sikt. För många åtgärder är det emellertid så att mätbara resultat uppkommer först efter en tid. Det kan även vara viktigt att studera utfall vid flera tillfällen för att se hur hållbara resultaten av åtgärden är. Det kan till exempel vara lämpligt att göra en första eftermätning som mäter kortsiktiga effekter, och sedan en andra eftermätning som mäter mer långsiktiga effekter.

I samband med programperioden 2007–2013 för EU:s strukturfonder har Europeiska kommissionen lanserat ett nytt utvärderingsbegrepp som vi redan introducerat i anslutning till Figur 2-1 – ’on-going evaluation’.¹⁰ Termen ”on-going” syftar på att utvärderingar görs under tiden som åtgärden genomförs. Denna typ ersätter det som under tidigare programperioder benämnts ”mid-term” eller ”halvtidsutvärderingar” vilket syftar på mitten av programperioden. En on-going utvärdering kan däremot ske närsomhelst under programperioden. En annan skillnad är att on-going ansatsen avser att ge större möjligheter att styra och påverka inriktning och genomförande av den aktuella åtgärden och innebär därmed att det rör sig om en mer efterfrågestyrd utvärdering. Ett viktigt syfte med on-going utvärderingar är därför att möjliggöra återkopplingar relativt tidigt i processen.

2.3 Olika typer av utvärderingar

Utvärderingar kan delas in i två grupper: utfalls- och genomförandeutvärderingar. Utfallsutvärderingar studerar resultat av åtgärder, medan genomförandeutvärderingar studerar besluts- och genomförandeprocesserna i sig själva. Dessa båda typer av utvärderingar kan med fördel kombineras då en genomförandeutvärdering kan bidra till att förklara varför ett observerat utfall av en åtgärd har blivit som det blivit.

¹⁰ *European Commission (2006).*

2.3.1 Utfallsutvärderingar

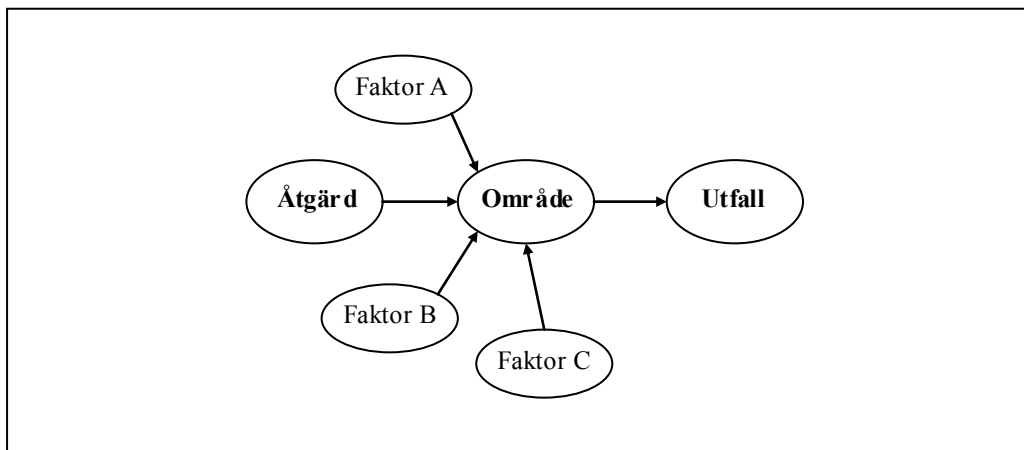
Utfallsutvärderingar består av mål-, effekt- och effektivitetsutvärderingar.

Målutvärdering

En målutvärdering eller en måluppfyllelseutvärdering svarar på frågan om, och i vilken grad, som specificerade mål för en åtgärd har uppnåtts.

En målutvärdering har den begränsningen att den endast utvärderar måluppfyllelsen men däremot inte i vilken utsträckning som måluppfyllelsen har orsakats av den aktuella åtgärden. En målutvärdering svara således endast på om målen är uppfyllda, inte om åtgärden har orsakat att målen är uppfyllda. Anta exempelvis att det har genomförts en åtgärd (till exempel företagsstöd) i ett visst område (till exempel glesbygd). I berörda företag kommer det att efter en tid kunna observeras ett visst utfall i form av till exempel en ökad vinst eller ökat antal anställda. Det utfallet kan förstås bero på åtgärden men också på andra faktorer så som illustreras i Figur 2-2. Det skulle därför vara grovt missledande att enbart på basis av en målutvärdering säga att åtgärden har medfört, eller fått till effekt, en ökning av vinsten eller av antalet anställda.

Figur 2-2 Åtgärden och övriga faktorer påverkan på utfallet.



En målformulering kan ofta vara komplicerad och kanske också motsägelsefull varför avgränsningar kan vara nödvändiga. Ett alternativ är att välja en avgränsning som innebär att endast något eller ett fåtal mål utvärderas. Det är dock viktigt att dessa avgränsningar motiveras väl. Det finns inte något entydigt svar på frågan om vilka mål som målutvärderingar bör inriktas på, detta måste bestämmas från fall till fall. Det kan också vara så att målformuleringarna är av den arten att en utvärdering inte kan ge direkta och entydiga svar på frågor om måluppfyllelse. Utvärderaren måste då nöja sig med att försöka ge indikationer på i vilken grad målen är uppfyllda. Målutvärderingar och indikatorer kan således vara nära besläktade. I Kapitel 3 behandlas indikatorer mer ingående.

Målutvärderingar har visserligen klara begränsningar men kan ändå ge värdefull information. Eftersom en målutvärdering endast studerar måluppfyllelsegraden, men inte hur denna har påverkats av den aktuella åtgärden, kan en sådan utvärdering endast ge en misstanke om eller indikation på att åtgärden har orsakat måluppfyllelsen. Denna "misstanke" kan vara mer eller mindre stark beroende på den specifika åtgärden och målen. En mål-

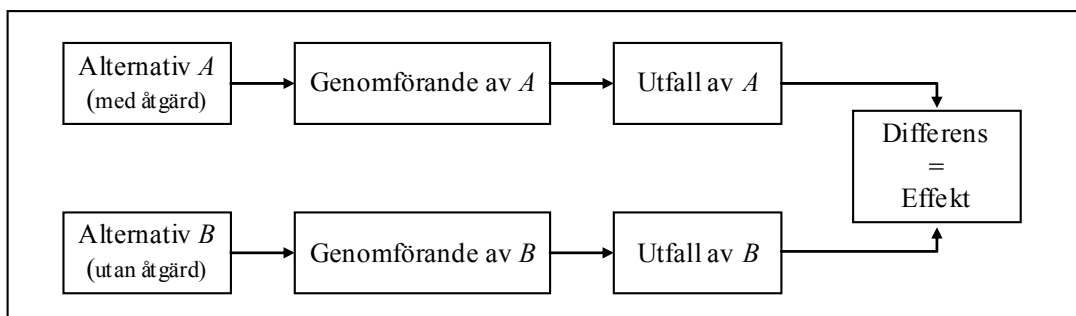
utvärdering av en åtgärd där andra faktorer påverkan på graden av måluppfyllelse kan anses ringa ger således en starkare misstanke att åtgärden har orsakat det observerade utfallet, än ett fall då andra faktorer påverkan antas vara stark. Men det bör betonas att vid en sådan *tolkning* av en målutvärderings resultat är man inne på att dra en slutsats om *orsakssambandet* bakom det observerade utfallet. Man är då på väg att ta steget över till den typ av utvärderingar som vi nu ska övergå till att beskriva.

Effektutvärdering

En effektutvärdering svarar på frågan om i vilken utsträckning som ett observerat utfall har orsakats av den åtgärd som utvärderingen gäller. Detta förutsätter en jämförelse mellan alternativa förlopp där det ena alternativet representeras av den aktuella åtgärden. Vid jämförelsen utnyttjas någon metod som gör det möjligt att, i termer av Figur 2-2, särskilja hur utfallet ha påverkats av faktorerna A, B och C respektive av den studerade åtgärden.

Anta att det finns två alternativ A och B, såsom i Figur 2-3. Låt alternativ A beskriva det förlopp då en viss åtgärd implementeras och alternativ B det förlopp då samma åtgärd inte genomförs – det kontrafaktiska förloppet. I en effektutvärdering är det skillnaden mellan utfallen av båda dessa förlopp som definieras som effekten av åtgärden.¹¹

Figur 2-3 En effekt är differensen mellan två alternativa förlopp.



En effekt kan också illustreras genom att jämföra Figur 2-3 med Figur 2-2: Differensen mellan alternativ A och B i Figur 2-3 är detsamma som differensen mellan att i Figur 2-2 implementera och inte implementera åtgärden. Att identifiera en effekt innebär således att påverkan från övriga faktorer hålls konstant.

En av svårigheterna med en effektutvärdering är att definiera det kontrafaktiska utfallet, det vill säga utfallet vid Alternativ B (utan åtgärd) i Figur 2-3. Hur detta görs beror av utvärderingens design eller utformning. En som regel mycket tillförlitlig utformning bygger på randomisering vilket innebär att slumpen får avgöra vilka i målgruppen (exempelvis individer, företag och regioner) som får del av åtgärden (Alternativ A i Figur 2-3) och vilka som inte får det (Alternativ B). Det kan emellertid finnas etiska invändningar mot en sådan design. De etiska invändningarna är dock oftast av mindre betydelse vid genomförandet av pilotprojekt. Att använda sig av randomiserade utvärderingar vid pilotprojekt är inte ovanligt i USA, där det till och med händer att myndigheter ställer krav på att utvärderingar ska baseras på slumpmässiga urval.

¹¹ Vedung (1998) sid. 137ff.

I vissa fall har implementeringen eller olika typer av oberoende händelser skapat en slumpvis variation när det gäller vilka som fått del av en viss åtgärd. I dessa fall talar man om naturliga experiment snarare än randomiserade. Sådana naturliga experiment kan skapa goda förutsättningar för att genomföra tillförlitliga utvärderingar. Som regel är dock utvärderingar av samhälleliga åtgärder hänvisad till så kallade kvasi-experimentella metoder där en fiktiv kontrafaktisk situation skapas med hjälp av ekonometriska metoder. Ett exempel på kvasi-experimentella metoder är matchningsmetoder som tas upp i Kapitel 4.

En effektutvärdering syftar ofta till att mäta effekten av en åtgärd på graden av måluppfyllelse i ett eller flera avseenden, det vill säga mäta skillnaden mellan åtgärden och jämförelsealternativet vad gäller måluppfyllelse. Ett vidare syfte är att mäta effekter som är intressanta (och kan väntas vara av någon mera betydande storleksordning) och då även sådana som kanske inte täcks av existerande målspecifikationer. Det kan röra sig om indirekta effekter och sådana kan vara antingen önskade eller oönskade. I vilket fall som helst görs det emellertid vid en effektutvärdering inte något försök att väga uppmätta positiva effekter mot den studerade åtgärdens effekter på kostnader. Detta är en uppgift för effektivitetsutvärderingar.

Effektivitetsutvärdering

I ekonomiska sammanhang avser begreppet ”effektivitet” förhållandet mellan insatserna i en verksamhet och verksamhetens resultat.¹² En effektivitetsutvärdering sätter således värdet av en åtgärds effekter i relation till kostnaden för åtgärden. Effektivitet betyder i vårt sammanhang detsamma som optimalitet – i samhälleliga sammanhang föreligger effektivitet när samhällets resurser utnyttjas på bästa möjliga sätt.

En effektivitetsutvärdering av en åtgärd innebär därför en granskning av alternativa sätt att uppnå vissa mål – en vägning av fördelar mot nackdelar av att övergå från säg alternativ *B* till alternativ *A*. Om fördelarna av en sådan övergång är större än nackdelarna, så är *A* att föredra från effektivitetssynpunkt – av de båda alternativen utnyttjas resurserna på bästa sätt vid *A*, det vill säga en övergång från *B* till *A* ger ökad effektivitet.

Privata företag är (något förenklat) effektiva då de uppnår målet om största möjliga vinst. För offentlig verksamhet är effektivitetsbegreppet mer komplicerat, vilket beror på att målen för offentlig verksamhet ofta är mer varierande och svårpreciserade och ibland även kontroversiella. Som utgångspunkt kan dock den enkla formuleringen användas att det överordnade målet är att göra samhället så bra som möjligt för människorna. En effektivitetsutvärdering av offentlig verksamhet avser åtgärdens för- och nackdelar för samhället det vill säga för de direkt och indirekt berörda människorna. ”Samhället” är kollektivet av individer, inte staten eller den offentliga sektorn. Att med utgångspunkt från samhälleliga målsättningar väga samman för- och nackdelar av en åtgärd brukar betecknas som en samhällsekonomisk lönsamhetskalkyl (cost-benefit kalkyl). De positiva effekterna ger upphov till samhällsekonomiska intäkter och de negativa effekterna ger upphov till samhällsekonomiska kostnader. Samhällsekonomiska intäkter är i princip allt som individer anser vara positivt och beredda att offra något för att få del av. Samhällsekonomiska kostnader är allt som individer uppfattar som negativt och är beredda att offra något för att slippa. En samhällsekonomisk kalkyl innebär ett försök att i görligaste mån göra positiva och negativa effekter av en åtgärd jämförbara. En praktisk måttstock är att använda det enhetliga måttet – pengar. Enligt Layard och Glaister är det ”... i sådana sammanhang absolut inte

¹² *Nationalencyklopedin (2008).*

*nödvändigt att använda pengar som värdemätare. Det kunde lika gärna vara liter spannmål, men pengar är praktiskt.*¹³

Rent formellt påminner därför en samhällsekonomisk lönsamhetskalkyl i hög grad om en företagsekonomisk kostnads- och intäktskalkyl. Skillnaden ligger framför allt i att de poster som ingår i en samhällsekonomisk kalkyl inte alltid svarar mot sådana transaktioner som i kommersiella kalkyler (och i vanligt språkbruk) förknippas med begreppen kostnader och intäkter. Eftersom samhällsekonomiska kalkyler handlar om för- och nackdelar för direkt och indirekt berörda människor ska det inte hysas några illusioner om att det är möjligt att värdera alla effekter i kronor på ett rimligt och tillförlitligt sätt. Varje samhällsekonomisk lönsamhetskalkyl är därför endast en partiell kalkyl. Det bästa är om alla effekter som människor fäster avseende vid har blivit mätta och värderade i samma mått, kronor, så att de har kunnat vägas samman. Sådana kalkyler är emellertid mycket svåra eller omöjliga att åstadkomma. Även om inte alla effekter har kunnat mätas, ger effektivitetsutvärderingar ett förbättrat beslutsunderlag där åtminstone en del betydelsefulla effekter har kunnat skattas och värderas på ett rimligt och tillförlitligt sätt.

2.3.2 Genomförandevärderingar

Vid en genomförandevärdering studeras besluts- och genomförandeprocesserna i sig själva. Ett av syftena med en genomförandevärdering är att få kunskap om hur utfallet av den aktuella åtgärden kan ha påverkats av planeringen och de grundläggande beslutsprocesserna och av själva genomförandet. En kombination av utfalls- och genomförandevärderingar är således värdefull som beslutsunderlag.

Utfallsutvärderingar betraktar genomförandeprocessen som en ”svart låda” där endast vad som stoppas in i och vad som kommer ut ur lådan är av intresse. Om enbart resultaten från utfallsutvärderingar läggs till grund för beslutsfattande kan det leda till nedläggning av verksamheter som i grunden är rätt tänkta men som kan fungera på ett bra sätt först sedan de utvecklats och förbättrats. Det är till exempel inte säkert att det vore klokt att avveckla en ny bilmodell om de första proven visar att bilen har lätt att välta vid en undanmanöver, eller att lägga ner utvecklingen av ett jaktflygplan på grund av en misslyckad landning.

I genomförandevärderingar kan, som redan antytts, även planeringsstadiet och de grundläggande beslutsprocesserna inkluderas. En genomförandevärdering av planeringsstadiet kan innefatta följande frågor:

- Varför aktualiserades den studerade åtgärden?
- Vilka aktörer medverkade i beslutsprocessen?
- Hur och varför medverkade de?
- Hur var verksamheten tänkt att fungera?
- På vilka förutsättningar var dessa föreställningar baserade?

¹³ *Layard och Glaister (1994) sid. 2, vår översättning.*

En utvärdering av själva genomförandet kan innefatta följande frågor:

- Hur organiserades verksamheten?
- Var ansvarsfördelningen tydlig?
- Var direktiven tydliga att följa?
- Fanns det intressekonflikter mellan olika aktörer?
- Vilka problem och svårigheter uppkom?
- Vilka förhållanden och orsakssamband kan ligga bakom de problem och svårigheter som uppstod?
- Har olika grupper av aktörer olika åsikter om hur verksamheten har fungerat?
- Ändrades organisationen, bland annat roll- och ansvarsfördelningen mellan olika aktörer, över tiden?
- Vad innebar och varför genomfördes dessa förändringar?
- Vilka resurser sattes in och vilka återverkningar fick detta för andra verksamheter?

Genomförandevärderingar kan ofta med fördel kombineras med måltvärderingar av olika slag av prestationsmått eller resultatindikatorer. De frågor som då kan ställas är exempelvis:

- Har verksamheterna varit till fullo inriktade mot uppställda mål?
- Har den aktuella målgruppen nåtts?
- Överensstämmer verksamhetens innehåll med vad som planerats?

Detta är tre frågor som gäller verksamhetens prestationer. Om målen ej är uppfyllda, kan genomförandevärderingen ge förklaring till varför de inte blev det samt eventuellt ge underlag för förslag på förbättringar.

Genomförandevärderingar kan vara tämligen enkla eller lika komplexa och omfattande som effekt- och effektivitetsvärderingar. Hur rigorös och sofistikerad insamling och analys av information som krävs beror i hög grad på den specifika frågeställningen. Skälet till att det ovan räknas upp en lång rad frågeställningar är att det är svårt att ge några allmänt giltiga riktlinjer för upplägg av och tillvägagångssätt för genomförande-utvärderingar.

En sammanfattning av de olika typerna av utvärderingar finns i Tabell 2-1. De olika slag av utvärderingar som presenterats kan sammanfattningsvis beskrivas så här: En måltvärdering svarar på frågan om målen uppfyllts; en effektutvärdering visar om måluppfyllelsen är en effekt av åtgärden; en effektivitetsvärdering besvarar frågan om värdet av de uppnådda effekterna överstiger kostnaderna för åtgärden; medan slutligen en genomförande-utvärdering ger kunskaper om hur själva genomförandet kan ha påverkat måluppfyllelse,

effekt och effektivitet. I Bilaga 2 presenteras de olika typerna av utvärderingar som är utgivna av ITPS.

Tabell 2-1 Olika typer av utvärderingar.

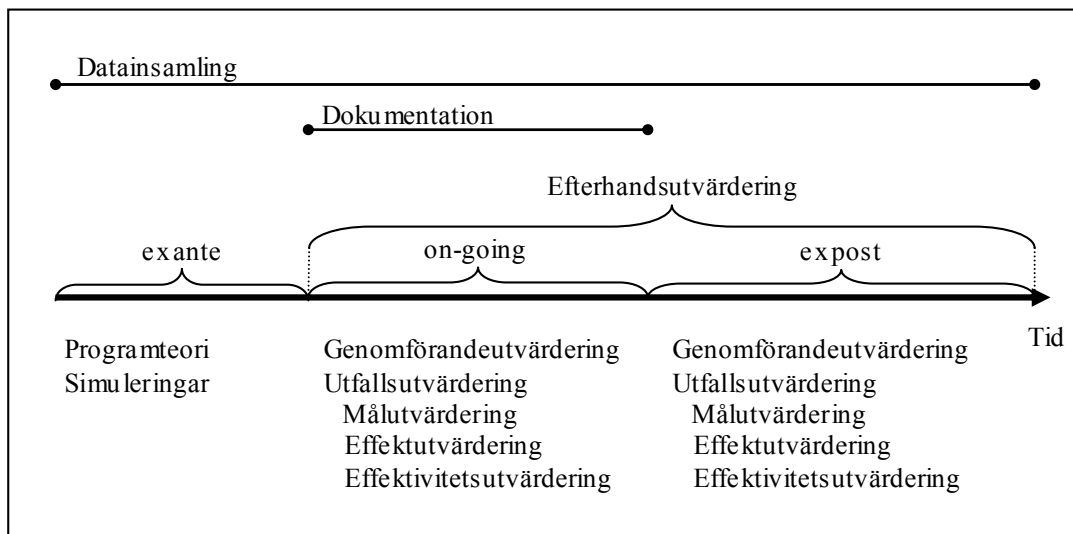
Utfallsutvärdering			Genomförandevärdering
Målutvärdering	Effektutvärdering	Effektivitetsutvärdering	
Är målen uppfyllda?	Är måluppfyllelsen en effekt av åtgärden?	Värderas de uppnådda effekterna högre än kostnaderna för åtgärden?	Hur har genomförandet påverkat måluppfyllelsen?
Studerar utfallet			Studerar genomförandet

2.4 Förutsättningar för tillförlitliga utvärderingar

Många gånger framställs frågan om en utvärdering ska fokusera på genomförande eller utfall, som ett val mellan användningen av kvantitativa respektive kvalitativa metoder. Det centrala i utvärderingssammanhang är emellertid inte ett sådant val, utan beslut rörande utvärderingstyp ska styras av vilken eller vilka frågor som utvärderingen ska ge svar på. I utfallsutvärderingar betraktas, som tidigare nämnts, själva processen som en 'svart låda' varför aktiviteter och händelser inom projektet inte beaktas. Om en sådan utvärdering inte kan påvisa några effekter av en åtgärd kan det inte utan vidare tolkas som att det alternativa handlings sättet är bättre. Resultatet kan ju vara en direkt orsak av problem vid genomförandet av åtgärden som till exempel samordningsproblem, intressekonflikter, resursbrist, byte av projektledare och oklar ansvarsfördelning. Att ha kunskaper om sådana händelser och omständigheter inom projektet som kan påverka utfallet är därför viktigt för att rätt kunna tolka den information som ges av en utfallsutvärdering. Omvänt är informationsvärdet av en väl genomförd genomförandevärdering begränsat om man inte vet något om utfallet vad gäller mål, effekter eller effektivitet. Idealet är därför en komplett utvärdering som innefattar både en genomförande- och utfallsutvärdering, ex ante, on-going och ex post.

Figur 2-4 illustrerar vilka typer av utvärderingar som kan utföras i olika skeden av ett program eller projekt, samt när datainsamling och dokumentation bör ske. En adekvat datainsamling och dokumentation är grundförutsättningar i en utvärdering. Att samla in och ställa samman en bra dokumentation är förstås naturligt för den som är involverad i och har en positiv syn på en åtgärd och därför vill övertyga andra om dess förträfflighet. I de fall då det för att genomföra en bra ex post-utvärdering fordras tillgång till data både från tiden innan och efter det att åtgärden har implementerats, måste utvärderaren ta ställning till om det rör sig om primärdata eller om de datauppgifter som behövs finns i befintliga register. Finns det inte registerdata att tillgå måste det ske en särskild datainsamling vilken då måste genomföras innan åtgärden har satts igång. Detsamma gäller genomförandevärdering vilka är beroende av dokumentation av genomförandet. Att i efterhand dokumentera planering och utförande kan visa sig synnerligen svårt, speciellt om det sker efter att det har gått en längre tid. Själva analysen av informationen kan dock, som visas i Figur 2-4, mycket väl ske ex post.

Figur 2-4 En schematisk bild av utvärderingscykeln.



ITPS utvärderingsarbete har vid flera tillfällen försvårats på grund av begränsningar i data. Dessa begränsningar har gällt dels själva tillgången till data, dels datas kvalitet. I ITPS utvärdering av expertskatten¹⁴ fick ITPS av sekretesskäl inte tillgång till den data som behövdes för utvärderingen. Den utvärderingen blev då begränsad till en målutvärdering. I andra fall har sekretesskäl hindrat ITPS från att använda viss data, men där ändå uppdraget har kunnat slutföras med andra data.¹⁵ Detta har då lett till begränsningar i vilka variabler samt vilka tidsperioder som har kunnat studeras.

Att få tillgång till data behöver dock inte alltid betyda att utvärderbarheten är säkrad. ITPS har erfarit fall där data varit av sådan kvalitet att en effektutvärdering inte är möjligt att utföra, utan en målutvärdering har fått göras i stället.¹⁶ I andra fall är dock situationen sådan att det t.ex. inte går finna ett kontrafaktiskt tillstånd, vilket försvårar utvärderingen.¹⁷

Institutet för arbetsmarknadspolitisk utvärdering (IFAU) har skrivit en rapport angående dataproblem vid utvärdering av arbetsmarknadspolitik.¹⁸ I denna rapport listas en rad problem med dokumentation av implementeringen av olika arbetsmarknadspolitiska åtgärder, vilket har gjort att det i flera fall inte har varit möjligt att utvärdera åtgärden överhuvudtaget. Rent generellt handlar en adekvat datainsamling för utfallsutvärderingar om att utfallet av de åtgärder som genomförs kan hänföras till åtgärdernas målgrupper som kan vara individer, företag, andra organisationer, regioner eller kommuner. Som nämndes ovan kan det vara aktuellt att hämta information om relevanta utfall även från perioden före implementeringen av en åtgärd. När det gäller effektutvärderingar är det nödvändigt att kunna studera det kontrafaktiska utfallet det vill säga utfallet vid det alternativ som den aktuella åtgärden jämförs med.

¹⁴ ITPS (2005).

¹⁵ Se ITPS (2003b), ITPS (2007b). I stället för att använda IFDB data har MMP data använts. IFDP data skulle kunna användas om rutinerna kring länkning och identifiering fungerade bättre.

¹⁶ Se ITPS (2003c), ITPS (2004), ITPS (2008e).

¹⁷ Se ITPS (2007a).

¹⁸ IFAU (2000).

Det första steget i en utvärderingsprocess är att fastställa utvärderingsbarheten av den åtgärd det gäller. Först därefter är det möjligt att svara på frågan om vilken eller vilka slags utvärderingar som det är möjligt att göra. Visar det sig att det inte är möjligt eller mycket svårt att få tillförlitliga svar på viktiga frågor, kan det bli aktuellt för utvärderaren att gå tillbaka till uppdragsgivaren för att ta en diskussion om tänkbara ändringar i uppdragets upplägg eller inriktning. Det är i det här skedet oerhört viktigt att kommunikationen mellan utvärderare och uppdragsgivare fungerar väl, så att utvärderingen svarar mot uppdragsgivarens förväntningar. Ett sätt att minska risken för svårutvärderade åtgärder är att utvärderaren kopplas in tidigt i processen för att garantera datainsamling och dokumentation och därmed även utvärderbarheten.

2.5 Att utvärdera åtgärder inom den regionala tillväxtpolitiken

Då många av grunderna för genomförandet av utvärderingar av politiska åtgärder inte skiljer sig åt mellan olika politikområden har ovanstående avsnitt beskrivit allmänna utgångspunkterna för utvärderingar. I det här avsnittet beskrivs emellertid aspekter som är specifika för utvärderingar av den regionala tillväxtpolitiken.

2.5.1 Olika åtgärder – olika angreppssätt

De olika åtgärderna inom den regionala tillväxtpolitikens kan grovt delas in i tre olika undergrupper: strategier, program och projekt. Även om dessa kategorier kan te sig helt olika, är utvärderingsgrunderna desamma. Det som skiljer dem åt är i stället möjligheterna att tillämpa olika utvärderingsmetoder.

När det gäller program består dessa ofta av en stor mängd projekt. De problem som en utvärderare ställs inför när det gäller att utvärdera hela program kan vara många. Regionala program tenderar att bedrivas på många olika platser och utförs i samverkan mellan många olika aktörer. Kausaliteten, orsakssambandet mellan åtgärder och utfall, kan vara oklar och det kan även finnas flera olika kausala samband som är verksamma samtidigt. Program kan därför vara svåra att effektutvärdera i sin helhet. Det kan däremot finnas större möjligheter att göra effektutvärderingar av enskilda projekt vars resultat kan användas som indikation på programeffekter. Kapitel 6 beskriver mer ingående olika angreppssätt för att utvärdera mer komplicerade program, och Kapitel 7 redogör för olika angreppssätt för att utvärdera horisontell samverkan.

Enligt regleringsbrevsuppdraget ska den här rapporten även behandla metoder för att utvärdera den regionala tillväxtpolitiken utifrån ett ekonomiskt, miljömässigt och socialt perspektiv.¹⁹ Utvärderingsgrunderna för dessa tre hållbarhetsdimensioner är i stort sett desamma som har beskrivits ovan i avsnitten 2.2–2.4 och kommer därför inte att behandlas ytterligare. Det som däremot kan vara värt att poängtera är att dessa tre aspekter måste beaktas i målformuleringar för strategier, program och projekt. I ITPS ex ante-bedömning av de regionala tillväxtprogrammen (RTP)²⁰ framkom att det fanns en stor variation i hur väl den miljömässiga dimensionen integrerades i analysen. Flera program saknade miljödimensionen helt, medan man i andra program hade försummat att integrera miljödimensionen med andra mål.

¹⁹ Se bilaga 1.

²⁰ ITPS (2003a).

2.5.2 Att ta hänsyn till den regionala dimensionen

Åtgärder inom de flesta politikområden har en regional dimension som det vid utvärderingar är viktigt att ta hänsyn till. Vid utvärderingar av åtgärder inom den regionala tillväxtpolitiken är den regionala dimensionen extra tydlig och bör speciellt beaktas. Att ta hänsyn till den regionala dimensionen kan grovt beskrivas som att man vid utvärderingar beaktar fem olika aspekter:

- (i) Gränslös pendling
- (ii) Regioners påverkan på andra regioner
- (iii) Regionala individuella effekter
- (iv) Regioners olikhet angående organisation och arbetssätt
- (v) Regioners olika storlek och påverkan på landet som helhet

(i) Gränslös pendling

Regioner är uppdelade efter olika administrativa gränser, men arbetspendlingen är gränslös. Aktiviteten i en region kan därför delas upp mellan dagbefolkning och nattbefolkning. Förekomsten av pendling mellan regioner gör att regionerna inte kan studeras som om de vore oberoende av varandra.

Ett sätt att ta hänsyn till denna problematik är att anpassa måtten till vad som utvärderingen avser att studera. För regionala lönesummor finns exempelvis både dag- och nattlönesumma. Daglönesumma används framför allt som en mått på en regions produktion, medan nattlönesumma mer beskriver vad som där konsumeras.²¹

Ett annat sätt att ta hänsyn till pendlingsproblematiken är att studera fiktiva regioner vars gränser är baserade på till exempel arbetskraftspendling. I ITPS årsbok Regionernas tillstånd²² används så kallade FA-regioner (funktionella analysregioner). Indelningen i FA-regioner är baserad på pendlingsstatistik och är gjord av NUTEK på ett sådant sätt att befolkningen kan bo och arbeta i en FA-region utan alltför tidsödande resor. Indelningen gör att regionerna blir relativt oberoende av varandra, samt att de på så sätt fångar en sorts ”vardagsregioner” där majoriteten av människorna kan få arbete och där majoriteten av regionens företag kan rekrytera sin arbetskraft. En alternativ indelning i funktionella regioner är SCB:s indelning av landet i lokala arbetsmarknader som bygger på statistik om faktisk pendling över kommungränser och där en lokal arbetsmarknad är en region med i huvudsak intern pendling.

(ii) Regioners påverkan på andra regioner

Förutom att pendling över regiongränser gör att regioner är beroende av varandra, finns det andra skäl till att regioner inte kan antas oberoende av varandra. Det är till exempel rimligt att anta att utvecklingen inom ett företag eller en region påverkas av utvecklingen inom närliggande företag eller regioner. För att ta tillvara den här typen av effekter kan rumslig

²¹ Se till exempel ITPS (2008b, 2009).

²² ITPS (2008g).

ekonometri (spatial econometrics) användas. ITPS har i en tidigare rapport²³ beskrivit hur analys med rumslig ekonometri kan tillämpas.

(iii) Regionala individuella effekter

En ofta återopad aspekt att ta hänsyn till när det gäller regional tillväxtpolitiska åtgärder är den regionala heterogeniteten det vill säga det förhållandet att regioner kan vara mer eller mindre olika varandra. En region kan till exempel ha särdrag som sätter sin prägel på, får effekt på olika företeelser (så att till exempel arbetslösheten regelmässigt är jämförelsevis låg där). Det rör sig då om en så kallad fast effekt som mycket väl kan påverka utfallet av en åtgärd på sätt som inte är fallet i andra regioner. Man kan till exempel tala om specifika Jämtlands- eller Dalarnaeffekter. Det kan emellertid också finnas ett samspel mellan en åtgärd och dess kontext, en så kallad interaktionseffekt, som resulterar i att åtgärdens utfall skiljer sig mellan regioner.

Regionala individuella effekter kan skattas med hjälp av data över tid och rum i regressio-
ner med paneldata (tidsseriedata från olika regioner). Ett annat och mer effektivt sätt att skatta regionala effekter är att använda flernivåanalys som presenteras och diskuteras i Kapitel 4.

(iv) Regioners olikhet angående organisation och arbetssätt

Regioner kan vara olika varandra till exempel genom att vara organiserade och styrda på olika sätt; det kan också finnas skillnader avseende ansvarsområden och befogenheter gällande regional utveckling i regionen.²⁴

Att vid utvärderingar ta hänsyn till den här typen av regionala olikheter kräver först att regioner kan identifieras och särskiljas efter dessa dimensioner. Att sedan kontrollera för olikheterna i analysen kan göras på ett liknande sätt som antydde i punkten (iii) ovan med regressio-
ner där regioner skiljs åt med hjälp av dummyvariabler eller med flernivåanalys.

(v) Regioners olika storlek och påverkan på landet som helhet

En annan aspekt är att regioner är olika stora med avseende på till exempel yta, befolkning eller ekonomisk styrka. Om till exempel arbetslösheten har gått ner i sju av tio regioner, kan det vara grovt missvisande att tolka det som att arbetslösheten för riket som helhet har gått ner. Problemet med ovanstående tolkning är att varje region studeras som enskilda observationer, där varje observation indirekt antas ha samma storlek. I studier av regioners påverkan på landet som helhet, kan det därför vara motiverat att vikta de olika regionernas betydelse, till exempel med befolknings storlek, för att ge ett mer rättvisande resultat.

2.6 Sammanfattning

Syftet med det här kapitlet är att definiera och beskriva skillnaden mellan olika typer av utvärderingar, samt att beskriva vad som behövs för att utföra tillförlitliga och användbara utvärderingar.

Utvärderingar kan utföras i olika skeenden av en åtgärd: på förhand (ex ante) det vill säga innan åtgärden har startat, under tiden den pågår (on-going), eller efter att åtgärden avslutats (ex post). Syftet med en utvärdering avgör var i processen utvärderingen bör äga rum.

²³ ITPS (2007c).

²⁴ ITPS (2008g).

Utvärderingar kan delas in i genomförande- och utfallsutvärderingar. Den förstnämnda typen av utvärderingar studerar planeringen och/eller genomförandet av åtgärder, medan den andra typen studerar åtgärdens utfall. Utfallsutvärderingar delas in i mål-, effekt- och effektivitetsutvärderingar. En måltvärdering svarar på frågan om målen uppfyllts; en effektutvärdering om måluppfyllelsen är en effekt av åtgärden; en effektivitetsutvärdering om de uppnådda effekterna värderas högre än kostnaderna för åtgärden, medan slutligen en genomförandeutvärdering svarar på frågor om hur själva genomförandet kan ha påverkat måluppfyllelse, effekt och effektivitet. Utfallsutvärderingar kan med fördel kombineras med genomförandeutvärderingar.

En grundförutsättning för att genomföra en utvärdering är att en adekvat datainsamling och dokumentation har utförts. Rent generellt handlar en adekvat datainsamling om att de utfall som observeras kan hänföras till målgruppen antingen det rör sig om individer, företag eller andra organisationer, regioner eller kommuner så att utfallet kan studeras både före och efter åtgärdens implementering. Ur utvärderingssynpunkt är det också fördelaktigt att kunna studera det kontrafaktiska tillståndet.

Det första steget i en utvärderingsprocess är att fastställa utvärderingsbarheten. Först därefter är det möjligt att svara på om det är möjligt att utvärdera åtgärden, det vill säga bedöma om en utvärdering kan ge svar på relevanta frågor. Om utvärderingen inte kommer att kunna ge svar på det som efterfrågas, kan det bli aktuellt att ändra inriktning eller upplägg för utvärderingen. Ett sätt att minska risken för svårutvärderade åtgärder är att utvärderare kopplas in tidigt i processen för att garantera datainsamling och dokumentation och därmed även utvärderbarheten.

Vid utvärderingar av den regionala tillväxtpolitiken är det speciellt viktigt att ta hänsyn till den regionala dimensionen och heterogeniteten. Det innebär bland annat att beakta sådant som pendling över regiongränser, regioners påverkan på andra regioner, regioners olikhet angående organisation och arbetssätt samt regioners olika storlek.

3 Indikatorer

3.1 Inledning

Indikatorer är ett vanligt förekommande verktyg för att beskriva och följa upp den regionala utvecklingen. Indikatorer kan emellertid även användas som ett hjälpmedel vid utvärdering av åtgärder. Därför bör man vara medveten om indikatorers styrkor och svagheter, samt hur det väljs ut på bästa sätt.

Det här kapitlet beskriver vad en indikator är, samt dess för- och nackdelar. Kapitlet beskriver även hur lämpliga och tillförlitliga indikatorer tas fram. Kapitlet bygger vidare på ITPS indikatorrapport *Indikatorer för uppföljning av regional utveckling*²⁵ och innehåller erfarenheter och lärdomar från detta arbete. Kapitlet beskriver även hur indikatorerna i den rapporten kan kompletteras för att ge en mer heltäckande bild av den regionala utvecklingen. Då det många gånger är viktigt att välja endast ett fåtal indikatorer, är det av stor vikt att dessa enstaka indikatorer ger en så heltäckande bild av den regionala utvecklingen som möjligt.

3.2 Vad är en indikator?

Ett sätt att skilja på olika typer av indikatorer är dela upp indikatorer i dem som visar på ett tillstånd, så kallade ”uppföljningsindikatorer” och dem som visar på en måluppfyllelse, så kallade ”utvärderingsindikatorer”. En uppföljningsindikator ger en indikation eller ett tecken på att någonting har hänt. Hur eller varför ”någonting” har hänt, ger uppföljningsindikatorn dock vanligtvis inte svar på. En uppföljningsindikator visar vanligtvis ett tillstånd såsom BNP per capita eller arbetslöshet i landet som helhet eller för olika regioner.²⁶ Dessa indikatorer beskriver endast ett tillstånd och det är svårt att utan vidare studier hänföra vissa åtgärders effekter till dessa tillstånd. ITPS har tidigare tagit fram ett flertal lämpliga uppföljningsindikatorer för att följa den regionala utvecklingen i olika delar av landet.²⁷

En utvärderingsindikator, å andra sidan, är mer precis i den meningen att det på goda grunder kan antas att förändringar i indikatorn har orsakats av den åtgärd som studeras. En utvärderingsindikator används således ofta i målutvärderingar, där indikatorn ger en misstanke om eller indikation på att åtgärden har orsakat måluppfyllelsen. Denna ”misstanke” kan vara mer eller mindre stark beroende på indikatorns utformning. Om det mål som studerats inte kan ha påverkats av något annat än åtgärden, visar dock utvärderingsindikatorn på en effekt såsom i en effektutvärdering.

Uppdelningen i uppföljnings- och utvärderingsindikatorer beskriver även indikatorers två syften: En uppföljningsindikator ger information om ett tillstånd, vilket eventuellt efterföljs av ett visst agerande. En utvärderingsindikator ger däremot information om hur ett visst agerande har påverkat måluppfyllelsegraden. Indikatorer av det senare slaget kan därför vara mycket användbara för utvärderingar, givet att de används på rätt sätt och att de som tolkar indikatorerna är medvetna om deras styrkor och svagheter.

Fördelen med att använda uppföljningsindikatorer är att de oftast är lätta att ta fram och att de till exempel kan ge en adekvat beskrivning av den regionala variationen och

²⁵ ITPS (2008a).

²⁶ Se till exempel Nutek (2006), ITPS (2003d), ITPS (2008a).

²⁷ ITPS (2003d).

utvecklingen utifrån ett ekonomiskt, socialt och miljömässigt perspektiv. Uppföljningsindikatorer ger dock vanligtvis inte information om olika åtgärders påverkan.

Fördelen med att använda utvärderingsindikatorer är att de ger en bild över en åtgärds effekt, samt att de kan vara smidiga att arbeta med. Nackdelen är att de kan vara svåra att utforma så att åtgärdens effekt verkligen särskiljs från övriga faktorerers inverkan. Utvärderingsindikatorer fyller en viktig funktion i monitoring sammanhang, för att i on-going utvärderingar möjliggöra återkoppling tidigt i processen.

3.3 Ett lämpligt urvalsförfarande för indikatorer

Ett första steg vid valet av indikatorer bör vara att besvara frågan: Ska indikatorn användas vid uppföljning eller vid utvärdering? Ambitionsnivån för utvärderingsindikatorer är naturligtvis högre. Ett val av indikator föregås här naturligt av att klargöra projektets mål, där målformuleringen med fördel bör följa SMART-kriterierna, det vill säga att målen ska vara: specifika, mätbara, accepterade, realistiska, och tidsbestämda (se Kapitel 2). Vid ett projekt vars mål följer SMART-kriterierna underlättas betydligt utformningen av lämpliga indikatorer. För att indikatorn ska kunna användas som indikation på resultat av ett visst agerande, bör orsakssambandet vara så pass klart att påverkan av andra faktorer kan anses vara mycket ringa.

Enligt SCIP (Sustainable Community Indicators Program) bör indikatorer följa vissa kvalitetskriterier. SCIP:s urvalskriterier för indikatorer är:²⁸

- Vetenskaplig validitet/teoretisk förankring
- Förändringsbenägenhet
- Tydliga samband mellan orsak och verkan
- Representativitet för aktuella frågor
- Korrekta tidsseriedata är tillgängliga eller insamlingsbara
- Kostnadseffektivitet
- Relevant och begriplig för användare
- Jämförbar mellan olika jurisdiktioner
- Användbar i stor och liten geografisk skala
- Möjlighet till jämförelse med mål, tröskelvärden eller standarder

Att ta hänsyn till alla dessa urvalskriterier samtidigt kan dock vara svårt. I ITPS rapport ”Indikatorer för uppföljning av regional utveckling” lyftes följande kriterier fram som särskilt betydelsefulla.²⁹

- Vetenskaplig och teoretisk validitet

²⁸ *European Commission 2008-03-04, denna lista överlappar även med Environment Canada, 2008-03-03.*

²⁹ *ITPS (2008a)*

- Datatillgänglighet (data över tid och på kommunnivå)
- Benägenhet att reagera på förändring
- Jämförbarhet mellan olika kommuner
- Tydlighet och begriplighet

Arbetsgången i framtagandet av uppföljningsindikatorerna i den ovan nämnda ITPS rapporten var följande: För det första var fyra fokusområden givna av regeringsuppdraget. Dessa områden var: innovation och förnyelse; kompetensförsörjning och ökat arbetskraftsutbud; tillgänglighet; och miljö. För varje område gjordes en teoriöversikt för att säkerställa vetenskaplig och teoretisk validitet. Därefter valdes indikatorerna beroende på datatillgänglighet, benägenhet att reagera på förändring, jämförbarhet mellan olika kommuner, samt tydlighet och begriplighet (utan inbördes rangordning). En försvårande omständighet visade sig vara att ta fram indikatorer på kommunnivå.

För att skapa ordning bland, framför allt, uppföljningsindikatorer, kan det vara bra att dela upp indikatorerna efter deras plats i orsakskedjan. Europeiska miljöbyrån (EEA) har utvecklat den så kallade DPSIR-modellen som tillkom för att bringa ordning bland olika typer av miljöindikatorer. Enligt denna modell kan indikatorer klassificeras i enlighet med deras plats i en orsakskedja enligt följande:

- | | |
|------------------------|-----------------------------------------------------------------------------------------------------------|
| (D) Drivkrafter | Aktiviteter som ligger bakom ett miljöproblem, till exempel transporter |
| (P) Påverkan | Vad som direkt orsakat problemet, till exempel surt nedfall. |
| (S) Status | Tillståndet i miljön, till exempel pH i grundvattnet. |
| (I) Inverkan | Effekt och konsekvenser, till exempel antal sinade brunnar. |
| (R) Respons | Åtgärder som görs för att minska eller lösa miljöproblemet, till exempel skydd av geologiska formationer. |

Först kommer ”drivkrafter” (D) vilket är de aktiviteter som ligger bakom miljöproblemet, exempelvis transporter och energianvändning. Drivkrafterna får en ”påverkan” (P) vilket är det som direkt orsakar problemet, exempelvis surt nedfall eller skogsavverkning. Påverkan ger i sin tur upphov till ett tillstånd eller en ”status” (S) för miljön i form av pH i grundvattnet, radonhalten i bostäder, eller brunnar med ohälsosamt vatten osv. Efter ”status” finns en kategori kallad ”inverkan” (I) som visar på effekter och konsekvenser av status, exempelvis kan radonhalten i bostäder påverka antalet cancerfall. Den femte och sista kategorin är ”respons” (R) vilken visar vilka åtgärder som gjorts för att minska eller lösa miljöproblemet. Såsom i de flesta typer av klassificeringar är det inte alltid helt självklart i vilken kategori en enskild indikator hör hemma, men orsakskedjan underlättar ändå betydligt framtagningen av lämpliga indikatorer. I ITPS rapport om indikatorer ansågs det som mest intressant att ta fram miljöindikatorer som visar på en påverkan (P) av olika aktiviteter.³⁰

³⁰ ITPS (2008a).

DPSIR-modellen är framtagen för att klassificera miljöindikatorer, men skulle också kunna användas inom andra områden. En DPSIR-modell för andra områden än miljö skulle exempelvis kunna förenklas till att endast betrakta *input* och *output*, eller *aktivitet* och *resultat*. På liknande sätt kan indikatorer delas in utifrån ett användareperspektiv så att de indikerar om det är främst ett företags- eller ett individperspektiv som beaktas (indikatorer för tillgänglighet bör till exempel både innefatta person- och godstransport).

Baserat på ovanstående diskussion kan ett lämpligt urvalsförfarande för uppföljningsindikatorer sammanfattas i följande tre steg:

Steg 1: Vetenskaplig grund

Baserat på tidigare vetenskapliga studier, vilka indikatorer rekommenderas?

Steg 2: Helhetsbild

Då det är vanligt för uppföljningsindikatorer att endast ett fåtal indikatorer används, är det viktigt att de som används ger en så heltäckande bild av tillståndet som möjligt. Ett hjälpmedel att få en helhetsbild är att utgå från *orsakskedjan* (Är hela orsakskedjan representerad?) och att tillämpa ett *användareperspektiv* (Är både ett företags- och ett individperspektiv representerat?).

Steg 3: Begränsningar i datatillgång

Begränsningar i datatillgång sätter dock de yttre förutsättningar för vilka indikatorer som kommer att kunna användas. En indikator bör förslagsvis kunna beskrivas kvantitativt över både tid och rum, vara benägen att förändras över tid, kunna jämföras över olika enheter, samt vara tydlig och begriplig.

Ett lämpligt urvalsförfarande för utvärderingsindikatorer innefattar främst steg 1 och 3 ovan, samt tillägget att ta hjälp av SMART-kriterierna att vara specifika, mätbara, accepterade, realistiska, och tidsbestämda. Indikatorn bör även vara utformad så att påverkan av andra faktorer på utfallet kan anses vara mycket ringa.

3.4 Sammanfattning

Indikatorer är ett hjälpmedel för att beskriva och följa den regionala variationen och utvecklingen och är ett nödvändigt verktyg även i utvärderingar. Det är av betydelse att veta indikatorers styrkor och svagheter, samt hur det väljs ut på bästa sätt.

Indikatorer kan delas in i *uppföljningsindikatorer* vilka ger information om ett tillstånd vilket eventuellt efterföljs av ett agerande, och *utvärderingsindikatorer* vilka syftar till att ge information om ett visst agerande. En uppföljningsindikator visar på *att* något har hänt, men ger vanligtvis inte svar på *varför* detta har hänt. Fördelen med att använda uppföljningsindikatorer är att de kan ge en översiktlig och korrekt bild över till exempel den regionala utvecklingen, samt att de oftast är lätta att ta fram. Nackdelen är de att vanligtvis inte ger information om olika åtgärders påverkan.

Fördelen med att använda utvärderingsindikatorer är att de kan ge en snabb bild över en åtgärds effekt, samt vara smidiga att arbeta med speciellt i monitoring sammanhang.

Nackdelen är att de kan vara svåra att utforma så att åtgärdens effekt verkligen särskiljs från övriga faktorerers inverkan.

Ett urval av indikatorer bör grundas på vetenskaplig och teoretisk relevans. En indikator bör förslagsvis kunna beskrivas kvantitativt över både tid och rum, vara benägen att förändras över tid, kunna jämföras över olika enheter, samt vara tydlig och begriplig.

4 Flernivåanalys och det regionala utvecklingsbidraget

4.1 Inledning

Inom den regionala tillväxtpolitiken är det troligt att de olika regionernas karakteristika har en betydande inverkan för hur väl ett projekt eller program lyckas. Att ta hänsyn till denna regionala heterogenitet är således något som är av betydelse vid i princip alla utvärderingar där man jämför utfallet i olika regioner. Flernivåanalys är ett verktyg för att ta tillvara och kontrollera för den regionala heterogeniteten i analysen. Syftet med det här kapitlet är att beskriva vad flernivåanalys är för något, när det kan användas, samt dess fördelar och nackdelar. Kapitlet ger även en kort introduktion till metoden propensity score matching. Kapitlet bygger på lärdomar från en av ITPS utförda utvärderingar (en effektutvärdering av det regionala utvecklingsbidraget³¹). Då kapitlet beskriver en statistisk metod är det något mer krävande än övriga kapitel. Kapitlet har följande disposition: Efter den här inledningen följer i avsnitt 4.2 en kort beskrivning av ITPS utvärdering av det regionala utvecklingsbidraget och propensity score matching. Avsnitt 4.3 beskriver mera utförligt för flernivåanalys. Kapitlet avslutad med en sammanfattning i avsnitt 4.4.

4.2 Att utvärdera det regionala utvecklingsbidraget

Under perioden 2003–2005 investerades i Sverige ungefär 25–30 miljarder kronor årligen i olika typer av stöd till företag. Det regionala utvecklingsbidraget (numera det regionala investeringsstödet) har delats ut sedan 1960-talet och är ett av de största företagsstöden i Sverige med ungefär 9,2 miljarder kronor i stöd under perioden 1990 till 2005.³²

Det finns dock väldigt få utvärderingar av det regionala utvecklingsbidraget, vilket var orsaken till att ITPS 2007 initierade en effektutvärdering av stödet.³³ En kritik som riktades mot den utvärderingen var att det ej tillräckligt hade kontrollerats för den regionala heterogeniteten. År 2008 initierade därför ITPS ytterligare en effektutvärdering av det regionala utvecklingsbidraget vilken bättre skulle kontrollera för skillnader mellan regioner.³⁴ De analysverktyg som användes var flernivåanalys, för att kontrollera för den regionala heterogeniteten, och propensity score matching, för att matcha stödföretag med icke stödföretag. Skillnaderna mellan utvärderingarna 2007 och 2008, var att den senare använde flernivåanalys, täckte en annan tidsperiod, samt utnyttjade delvis andra data.

En utvärderare som får till uppgift att utvärdera effekten av företagsstöd ställs inför det klassiska utvärderingsproblemet att skatta det kontrafaktiska utfallet, det vill säga i det här fallet det utfall som företag hade fått om de inte erhållit stödet. Det är ju skillnaden mellan detta och utfallet när företag får stöd som är företagsstödet effekt. Ett framgångsrikt angreppssätt är att jämföra utfallet i stödföretag med det i icke-stödföretag som i utgångsläget är så lika stödföretagen som möjligt. Genom att jämföra företag med deras ”tvillingar”, som så nära som möjligt representerar en kontrafaktisk situation, kan effekterna av stödet utvärderas. Propensity score matching är en metod för att matcha individer, företag osv.

³¹ Gadd et al (2008).

³² Gadd et al (2008).

³³ ITPS (2007b).

³⁴ Gadd et al (2008).

med varandra.³⁵ Den metoden användes i ITPS utvärderingar av det regionala utvecklingsbidraget.³⁶ Matchning kan mycket kort sammanfattas i fyra punkter:

- Målet är att skapa en kontrafaktisk situation genom att hitta företag som är så lika stödföretagen som möjligt.
- I ett första steg estimeras sannolikheten för företag att få stöd.
- I ett andra steg används dessa sannolikheter till att para ihop stödföretag med de icke-stödföretag som har samma sannolikhet att få stöd. Stödföretagen blir på så sätt matchade med ”tvillingar” bland icke-stödföretagen.
- I ett tredje steg jämförs stödföretagens utveckling i avseenden som är relevanta för utvärderingen med utvecklingen hos de matchade icke-stödföretagen. Om stödföretagen har haft en bättre utveckling än sina ”tvillingar” tolkas detta som en effekt av stödet.

Det är i det första steget, det vill säga vid skattningen av sannolikheten att få stödet, som det kan vara motiverat att använda flernivåanalys.

4.3 Flernivåanalys

Flernivåanalys används när det finns ett samspel mellan data på olika nivåer. Det klassiska exemplet är studiet av uppnådda poäng på tester för skolelever, där eleverna utgör nivå 1, klassen nivå 2, och skolan nivå 3. Med hjälp av flernivåanalys är det möjligt att ta hänsyn till att elevers testresultat kan bero av vilken klass de går i och i vilken skola som klassen finns. I ITPS effektutvärdering av företagsstödet finns tre nivåer: företagen på nivå 1, kommunerna på nivå 2, och länen på nivå 3. Metoden gör det möjligt att identifiera effekter som relaterar sig till dessa olika nivåer, exempelvis en specifik Jämtlandseffekt (län) och en specifik Östersundseffekt (kommun). Denna egenskap hos metoden gör den speciellt lämpad vid studier med en regional dimension.³⁷

För att förklara metoden använder vi ett exempel. Anta att vi ska undersöka variationen i sannolikheten att företag får stöd. I en modell utan hänsyn till regional variation kan modellen formuleras som ekvation (1):

$$y_{ij} = \beta_0 + \beta_1 x_{ij} + \varepsilon_{ij}, \quad (1)$$

där y_{ij} är den beroende variabeln som är lika med 1 om företag i , i kommun j får stöd, och 0 om företaget inte får stöd. β_0 är ett intercept, β_1 beskriver effekten av en förklaringsvariabel som vi i exemplet antar är antalet anställda x_{ij} och ε_{ij} är en slumpterm som tar upp allt det brus som inte förklaras av modellen. I ekvation (1) antas för det första att alla kommuner ger samma effekt på sannolikheten att få stöd (vilket fångas av β_0) och för det andra att effekten av antalet anställda är densamma i alla kommuner (vilket fångas av β_1).

³⁵ Se Rosenbaum och Rubin (1983).

³⁶ Se ITPS (2007b), Gadd et al (2008).

³⁷ Se till exempel: Twisk (2006), Brickel (2007), och Gelman och Hill (2007).

Ett sätt att kontrollera för om olika kommuner har olika effekter skulle kunna vara att införa dummyvariabler för olika kommuner. Ekvation (1) kompletterad med dummyvariabler för Sveriges 290 kommuner blir då:

$$y_{ij} = \beta_1 x_{ij} + \delta_1 d_1 + \delta_2 d_2 + \dots + \delta_{290} d_{290} + \varepsilon_{ij}. \quad (2)$$

I denna ekvation är $d_j = 1$ om företaget finns i kommun j och annars 0. Koefficienten δ_j visar därför den effekt på sannolikheten att få stöd som förklaras av att företaget befinner sig i kommun j . Märk att det gemensamma interceptet β_0 i ekvation (1) är här ersatt med de kommunspecifika intercepten $\delta_1 d_1 + \delta_2 d_2 + \dots + \delta_{290} d_{290}$.³⁸ För en modell som omfattar hela Sverige innebär det således att 289 extra koefficienter skattas. Om detta angreppssätt används kommer det att minska antalet frihetsgrader och således minska precisionen i skattningen i den meningen att koefficienterna skattas med en större osäkerhet. Ju fler koefficienter som skattas desto större är osäkerheten i skattningen av dessa koefficienter.

Ett sätt att komma runt det problemet är att i stället för en modell med dummyvariabler använda flernivåanalys. Modellen i ekvation (2) kan då ersättas med den i ekvation (3):

$$y_{ij} = \beta_0 + \beta_1 x_{ij} + \nu_{0j} + \varepsilon_{ij}. \quad (3)$$

I ekvation (3) har dummyvariablerna blivit ersatta av ν_{0j} , där varje enskilt värde j representerar effekten av varje enskild kommun. Genom att formulera problemet på detta sätt skattas endast två koefficienter (β_0 och β_1) och är således ett mer effektivt sätt att estimeras effekten av regiontillhörighet då det finns många olika regioner.

Slumftermen (ν_{0j}) kallas för kommunspecifika intercept (*random intercepts*) medan koefficienter likt den för antalet anställda i vårt exempel (β_1) kallas för ”fasta koefficienter” (*fixed coefficients*). De fasta koefficienterna (β) tolkas som i en vanlig regression. I exemplet ovan så betyder det att i ett företag med x antal anställda påverkas sannolikheten att få stöd med $\beta_1 * x$. De kommunspecifika intercepten visar däremot på effekten av att företaget är lokaliserat till en viss kommun. Om den skattade variansen för det kommunspecifika interceptet ($\hat{\sigma}_{0v}^2$) är signifikant skild från noll, betyder det att variationen mellan kommunerna har betydelse för det utfall som studeras, det vill säga att det finns en regional heterogenitet på kommunnivå som i vårt exempel påverkar sannolikheten att få stöd.

Ett alternativt sätt att bedöma vikten av den kommunspecifika variationen är att beräkna en så kallad *Intraclass Correlation Coefficient* (ICC). Detta mått varierar mellan 0 och 1 där värden nära noll innebär att variationer mellan kommuner inte påverkar utfallet och ett värde lika med 1 betyder att all variation i utfallet beror på skillnader mellan kommuner. I utvärderingen av det regionala utvecklingsbidraget var $ICC = 0,1$ vilket tolkas som att tio procent av variationen i sannolikheten att få stöd förklaras av skillnader mellan kommuner.

³⁸ *Tas inte interceptet bort hamnar vi i den så kallade dummyvariabel fällan. Se till exempel Greene (2003) sid. 118.*

I ekvation (3) ovan så antas fortfarande att effekten av antalet anställda är likadan i alla kommuner (det vill säga β_j är lika stor i alla kommuner). Ett sätt att kontrollera för om effekten av antalet anställda är olika i olika kommuner fås genom att utöka modellen i ekvation (2) med en term för ”interaktion” mellan kommunintercepten och antalet anställda. Modellen i ekvation (3) skulle med interaktioner skrivas:

$$y_{ij} = \beta_0 + \beta_1 x_{ij} + v_{0j} + v_{1j} x_{ij} + \varepsilon_{ij}, \quad (4)$$

där $v_{1j} x_{ij}$ är de olika effekterna av antal anställda i olika kommuner. Ekvation (4) innehåller således både individuella intercept för varje kommun (v_{0j} , *random intercepts*) och tar hänsyn till att det kan finnas skillnader mellan kommuner vad gäller effekten av antalet anställda (v_{1j} , *random slopes*).

Eftersom ekvationen ovan ska skatta sannolikheten att erhålla regionalt utvecklingsbidrag har, som nämndes inledningsvis, den beroende variabeln bara två värden: 1 om företaget fått bidrag och 0 om företaget inte fått bidrag. Därför görs skattningen lämpligen med en probit- eller logitestimator, vilka är standard inom ekonometri och som därför inte kommer att behandlas närmare här³⁹. Ekvation (4) i logitform blir:

$$f(\pi_{ij}) = \beta_0 + \beta_1 x_{ij} + v_{0j} + v_{1j} x_{ij}, \quad (5)$$

där π_{ij} är sannolikheten att ett företag i i kommun j får stöd. $f(\pi_{ij})$ är den logit-link-funktion som garanterar att de skattade sannolikheterna ligger mellan 0 och 1.⁴⁰

Ekvation (5) kan generaliseras på många olika sätt. Exempelvis kan modellen utökas med ytterligare nivåer, exempelvis länsnivå. Nackdelar med att inkludera flera nivåer är att resultaten blir mer svårtolkade och att de beräkningar som behöver utföras blir mer tidskrävande.

I ITPS effektutvärdering av det regionala utvecklingsbidraget specificerades en modell med både kommun- och läns effekt, vilken tog flera timmar att estimeras. Resultaten visade på en signifikant kommuneffekt medan läns effekten inte var signifikant. Detta kan tolkas som att det fanns variation på kommunnivå som förklarade sannolikheten att erhålla bidrag, men att däremot inga sådana effekter gick att finna på länsnivå. Att läns effekten inte var signifikant ska dock tolkas med en viss försiktighet. När antalet grupper (till exempel antalet län) på en viss nivå (till exempel länsnivå) är få till antalet, finns det generellt inte tillräckligt med information för att skatta de olika gruppeffekterna⁴¹. I sådana fall kan det därför vara ändamålsenligt att inte använda flernivåanalys utan en ”klassisk” multipel regression som i ekvation (2) kan i stället vara att föredra. I ITPS effektutvärdering fanns

³⁹ Se till exempel: Gelman och Hill (2007).

⁴⁰ Att beräkna ICC för en logit modell kan approximeras av $\hat{\sigma}_{ov}^2 / (\hat{\sigma}_{ov}^2 + \pi^2/3)$. Det finns även en diskussion om ICC:s relevans i denna typ av modell i exempelvis Twisk (2006) sid. 46.

⁴¹ Gelman och Hill (2008) sid. 247.

endast tio län medtagna vilket skulle kunna vara en förklaring till att läns effekterna inte var signifikanta.

De finns således både för- och nackdelar med flernivåmodeller. En grundläggande fördel är att flernivåanalys är ett effektivt sätt att ta hänsyn till flera nivåer i analysen, genom att skatta individuella gruppeffekter (*random intercepts*) och att effekten från en variabel (till exempel antal anställda) kan tillåtas att vara olika för olika grupper (*random slopes*). Värt att notera är att modellerna endast används för att ta hänsyn till att det finns olika grupp-effekter, inte för att testa om dessa är statistiskt skilda åt⁴². Metoden tillåter även analyser med olika nivåer då antalet observationer inom en del grupper är få⁴³. Programvara för att genomföra flernivåanalys finns tillgänglig i flera statistikprogram till exempel STATA, R och MLWIN, men programmen är beräkningsintensiva och det krävs insikt i metoden för att tolka resultaten. Flernivåanalys används inte lika frekvent som andra metoder, vilket gör att det kan krävas rätt utförliga förklaringar när det gäller att förmedla resultat av sådana analyser till en publik.

4.4 Sammanfattning

Syftet med det här kapitlet är att beskriva flernivåanalys, när det kan användas, samt dess fördelar och nackdelar. De erfarenheter och lärdomar som beskrivs är hämtade från ITPS effektutvärdering av det regionala utvecklingsbidraget.⁴⁴

Flernivåanalys är en effektiv metod dels för att kontrollera för fasta effekter förknippade med olika regioner, dels för att kontrollera för skillnader mellan regioner vad gäller effekten av en viss variabel. Flernivåanalys ger således möjlighet att ta hänsyn till olika nivåer i analysen, det vill säga att kontrollera för till exempel en specifik Jämtlandseffekt (länsnivå), eller en specifik Östersundseffekt (kommunnivå), samt om effekten av en variabel är olika i olika regioner (interaktionseffekt). Flernivåanalys lämpar sig väl för olika typer av analyser där det finns anledning att ta hänsyn till att regioners specifika karaktäristika kan påverka analysen.

En grundläggande fördel är att flernivåanalys är ett effektivt sätt att ta hänsyn till flera nivåer och grupper i analysen. Metoden tillåter analyser med olika nivåer då antalet observationer inom en del grupper är få. Metoden fungerar dock sämre då antalet grupper inom en nivå är få. Värt att notera är även att metoden endast används för att ta hänsyn till att det finns olika grupp-effekter, inte för att testa om dessa är statistiskt skilda åt.

Programvara för att genomföra flernivåanalys finns tillgänglig i flera statistikprogram till exempel STATA, R och MLWIN, men programmen är beräkningsintensiva och det krävs insikt i metoden för att tolka resultaten. Flernivåanalys används inte lika frekvent som andra metoder, vilket gör att det kan krävas rätt utförliga förklaringar när det gäller att förmedla resultat av sådana analyser till en publik.

⁴² Gellman och Hill (2008) sid. 271.

⁴³ Gellman och Hill (2008) sid. 247.

⁴⁴ Gadd et al (2008).

5 Lärdomar från metautvärdering

5.1 Inledning

En metautvärdering är en utvärdering av en eller flera utvärderingar. Fokus är vanligen på att bedöma kvaliteten hos och nyttan med de underliggande utvärderingarna. ITPS genomförde 2008 en metautvärdering av sex strukturfondsprogram i Sverige under 2000–2006.⁴⁵ Den dåvarande utvärderingspolicyn innebar att de sex studerade programmen, två Mål 1- och fyra Mål 2-program, utvärderades på förhand, i halvtid och efter att programmen avslutats. 2003 utfördes halvtidsutvärderingar av vart och ett av dessa tillväxtorienterade program. Syftet var att undersöka programmets resultat i förhållande till det förväntade och ge rekommendationer till eventuella förändringar för att nå uppställda mål. Dessa sex halvtidsutvärderingar bildade underlag för ITPS metautvärdering. Tanken var att bedöma utvärderingarnas kvalitet. Ett ytterligare syfte var att undersöka strukturfondernas utfall. Detta utfördes med hjälp av en så kallad sekundäranalys.

Syftet med det här kapitlet är att sammanfatta erfarenheter och lärdomar från den utförda rapporten *Metautvärdering av strukturfondsprogrammen – nya ansatser och lärdomar*. Först diskuteras därför i avsnitt 5.2 metautvärdering som metod. Därefter beskrivs i avsnitt 5.3 erfarenheter av rapportens två angreppssätt: metautvärdering och sekundäranalys. Fokus ligger på tillämpning av metoderna och vilka lärdomar som kan dras av denna.

5.2 Metautvärdering

5.2.1 Tänkbara syften med metautvärdering

Begreppet metautvärdering introducerades 1969 av Michael Scriven och har sedan dess fått ökad betydelse, framför allt inom den amerikanska utvärderingstraditionen. Med tanke på den ständigt växande omfattningen av utvärderingar, i synnerhet sådana som i någon mening behandlar politiska insatser världen över, har även behoven av att granska utvärderarnas arbete ökat. Det är i detta sammanhang metautvärderingar har en roll att fylla.⁴⁶

Metautvärdering definieras i olika ordalag även om innebörden vanligen pekar i en liknande riktning. Stufflebeam definierade 1974 metautvärdering som: ”... ett förfaringsätt för att beskriva en utvärderingsaktivitet och bedöma den i förhållande till uppfattningar kring vad som utgör en god utvärdering.”⁴⁷ Vedung framhåller att metautvärdering grundas på föreställningen om att även utvärderingar kan utvärderas. På så sätt utgör metoden en efterhandskontroll.⁴⁸ Statskontoret skiljer på tre former av metautvärderingar, de som:

- bearbetar resultaten från två eller flera utvärderingar
- prövar kvaliteten och precisionen i en eller flera utvärderingar
- utvärderar en utvärderingsorganisation.⁴⁹

⁴⁵ ITPS (2008e).

⁴⁶ Scriven (1969); Vedung (1998).

⁴⁷ Stufflebeam (1974).

⁴⁸ Vedung (1998), sid. 39.

⁴⁹ Statskontoret (2001), sid. 4.

Metautvärderingar kan även ha en jämförande ambition. Det kräver att flera utvärderingar bildar underlag och kan exempelvis innebära att de underliggande utvärderingarna graderas utifrån ett antal kriterier.⁵⁰

Ett antal, mer utvecklade, skäl till att göra metautvärderingar kan återfinnas i litteraturen kring sådana. Flera teoretiker framhåller vikten av och nyttan med att studera och bedöma utvärderingars utformning. Denna bör ha utgått från målbeskrivningen för den studerade aktiviteten. Till bedömningen av utvärderingars utformning kan det även höra att granska eventuella hot mot utvärderingars validitet. En övergripande anledning att göra metautvärderingar är att studera utvärderingars syften och avgöra om dessa nåts. Det innefattar även att avgöra rimligheten i undersökta utvärderingars resultat och om möjligt bedöma om det existerar andra variabler som kan förklara det studerade utfallet än de som inkluderats i utvärderingen. Det sistnämnda är något som man främst har riktat in sig på vid metautvärderingar inom områdena psykologi och pedagogik.⁵¹

Sammanfattningsvis innebär detta att på en övergripande nivå syftar en metautvärdering till att bedöma kvaliteten hos och värdet av en eller flera utvärderingar. Vad gäller inriktningen på en mer detaljerad nivå så ger den aktuella forskningen uppslag till olika infallsvinklar.

5.2.2 Tillvägagångssätt

En ökad förståelse för metautvärdering kan fås genom att studera vilka tillvägagångssätt som rekommenderas. Förutom vad som förordas i litteraturen finns det även råd och riktlinjer utformade av organisationer. Exempel på två sådana organisationer är Joint Committee in Standards for Educational Evaluation och American Evaluation Association vilka utarbetar riktlinjer för olika typer av utvärderingar, däribland metautvärderingar.⁵²

Stufflebeam ger en konkret beskrivning av hur en metautvärdering kan genomföras, vilken här sammanfattats i Tabell 5-1. Några av punkterna i tabellen är av särskild vikt. En av dessa är att bestämma metautvärderingens frågor. De bör formuleras så att de båda övergripande syftena med metautvärderingar – att bedöma kvaliteten i utvärderingar och berörda aktörers grad av nytta av utvärderingarna – kan uppfyllas. Vidare är det betydelsefullt att metautvärderingen ställer särskilda, mer specifika, frågor som kan bedömas vara relevanta för exempelvis beslutsfattare och initiativtagare till ett program eller för dess aktörer.

Beträffande den faktiska datainsamlingen kan olika metoder användas beroende på metautvärderingens utformning. Den eller de utvärderingar som ligger till grund för arbetet utgör naturligtvis den primära källan. Det kan dock finnas skäl som gör det motiverat att gå vidare med datainsamlingen och inhämta ny information. Detta beror vanligen på att mer material krävs för att besvara metautvärderingens frågor och kan exempelvis innebära att intervjuer, enkäter eller andra datainsamlingsmetoder får tillgripas.

Med tanke på en metautvärderings möjliga variation i handhavandet av empiriskt material kan inte tänkbara analystekniker begränsas på förhand. Att upprepa analyser av den eller de underliggande utvärderingens material kan ge upphov till andra resultat än de som ursprungligen presenterades. Detta tydliggör vikten av metautvärdering som ett viktigt moment i en efterhandskontroll.

⁵⁰ För en utvecklad beskrivning, se: *Stufflebeam (2001), sid. 202–204.*

⁵¹ *Stufflebeam (1974); Bamberger et al (2006).*

⁵² Se <http://www.wmich.edu/evalctr/jc/> och <http://www.eval.org/>

Av stor vikt vid en metautvärdering är att bedöma hur väl de studerade utvärderingarna överensstämmer med gällande standarder, principer eller kriterier. Som nämnts tidigare finns det då ett antal olika etablerade riktlinjer från organisationer att följa. I de fall det är möjligt bör man naturligtvis även beakta hur väl utvärderarna följt befintliga policy-dokument eller avtal.⁵³

Tabell 5-1 Struktur för tillvägagångssätt vid metautvärdering.

- Bestäm och arrangera interaktionen mellan metautvärderare och personer som på något sett är berörda av program eller reformer.
- Besätt projektet med en eller flera kvalificerade utvärderare.
- Definiera metautvärderingens frågor.
- Bestäm standarder, principer och/eller kriterier för att kunna bedöma utvärderingarna.
- Utveckla en överenskommelse eller ett kontrakt som ligger till grund för styrningen av metautvärderingen.
- Samla in och bedöm relevant tillgänglig information.
- Samla om nödvändigt in ny information, exempelvis intervjuer, observationer och intervjuundersökningar.
- Analysera den kvalitativa och den kvantitativa informationen.
- Bedöm utvärderingarnas anknytning till gällande standarder, principer och/eller kriterier.
- Meddela metautvärderingens resultat via rapporter, korrespondens, muntlig framställning och liknande.
- Hjälプ vid behov, och om möjligt, klienter och programdeltagare med tolkning av resultatet.

Källa: *Stufflebeam (2001)*.

5.3 Erfarenheter från ITPS metautvärdering av strukturfondsprogrammen

ITPS rapport *Metautvärdering av strukturfondsprogrammen – nya ansatser och lärdomar*⁵⁴ består av två empiriska delar. Den första, som utgörs av en traditionell metautvärdering, behandlas nedan i avsnitt 5.3.1. Därefter beskrivs i avsnitt 5.3.2 metodik och erfarenheter från den andra empiriska delen som är en sekundäranalys av strukturfondernas utfall.

5.3.1 Metautvärdering

I den del av rapporten som utgjorde en metautvärdering stod de sex halvtidsutvärderingarna av strukturfondsprogrammen Mål 1 respektive Mål 2 för det huvudsakliga materialet. Två, mer specifika, frågeställningar bedömdes vara särskilt intressanta för denna del:

- Hur svarar utvärderingarna mot allmänna krav samt riktlinjer från EU och upphandlande myndighet?
- Hur väl tillgodoser halvtidsutvärderingarna de behov av uppföljning och utvärdering som aktörer inom programmen kan tänkas ha?

⁵³ *Stufflebeam (2001)*.

⁵⁴ *ITPS (2008e)*.

Då det var flera utvärderingar som studerades utformades en checklista baserad på riktlinjer och bestämmelser från Europeiska kommissionen och de två upphandlingsunderlagen. Genom att applicera samma standarder och principer för bedömning av samtliga sex utvärderingar gavs möjlighet till en systematisk bedömning och utrymme för jämförelser.

För att besvara den andra frågan studerades dokument från programmen där huvudrekommendationerna i utvärderingarna bemötts och kommenterats.

Erfarenheter

Metautvärderingen utgick från sex utvärderingar vars innehåll och fokus var tämligen strikt styrt av den regelapparat som omger strukturfonderna. De riktlinjer som uttryckts av EU operationaliserades i upphandlingsunderlaget. På så sätt har en tydlig kravprofil styrt arbetet på en relativt detaljerad nivå. I det här fallet har det medfört att det vid metautvärderingen har varit grundläggande att jämföra utvärderingarna i förhållande till ställda krav. Det arbetet har underlättats av att dessa innehållsmässiga krav i de flesta fall varit konkreta och väl formulerade.

Metautvärderingar med flera underliggande utvärderingar kan per definition bedömas ha en fördel eftersom det finns möjlighet till jämförelser som kan ligga till grund för bedömning av olika tillvägagångssätt. Förutom att utvärderingar kan vägas mot lämpliga standarder och kriterier kan en metautvärdering som handhar flera utvärderingar relatera dessa till varandra. Då kan särskilt framgångsrika förfaranden lyftas fram, vilket kan vara till nytta för aktörer inom de utvärderade programmen, ansvariga utvärderare och utvärderare av kommande program. I det aktuella fallet underlättades jämförelser av att de granskade utvärderingarna avsåg program med ett likartat innehåll. Variationen mellan de sex utvärderingarna var emellertid liten vad gäller struktur och inriktning. Det har sin bakgrund i att en och samma aktör erhöll uppdraget att utföra samtliga utvärderingar. En optimal utgångspunkt för jämförelse är däremot att utvärderingar som studerar ett och samma eller likartade samhällsfenomen har tillräckligt stor variation sinsemellan för att fördelar och nackdelar med särskilda tillvägagångssätt ska kunna belysas.

Litteraturen om metautvärderingar rekommenderar att utvärderingar jämförs med rådande standarder, principer eller kriterier. I det här fallet med strukturfondsprogrammen har bedömningen varit att förutom att relatera utvärderingarna till riktlinjer och upphandlingsunderlag samt till varandra har det vid metautvärderingen även varit viktigt att försöka anlägga ett perspektiv som fångar vad aktörer inom programmet kan tänkas efterfråga. I mångt och mycket överensstämmer detta rimligtvis med de allmänna krav man kan ställa på en skickligt utförd utvärdering. Den aktuella metautvärderingen gör dock inte anspråk på att alla efterfrågade aspekter inkluderats. En ansats för att säkerställa detta, som emellertid i det här fallet låg utanför arbetet, kan vara att intervjua aktörer inom och i anslutning till programmen om deras uppfattning av nyttan med utvärderingarna. I det här fallet studerades i stället hur förvaltningsmyndigheterna för respektive program ställt sig till utvärderingarnas huvudrekommendationer.

På basis av beskrivningen ovan kan tre ansatser för att bedöma kvaliteten i utvärderingar lyftas fram: För det första framhåller forskningen att utvärderingar bör relateras till rådande standarder, principer eller kriterier. För det andra kan det även vara viktigt att bedöma utvärderingar i förhållande till politiskt beslutade riktlinjer och upphandlingsunderlag. För det tredje kan utvärderingars kvalitet bedömas genom jämförelser med andra

likartade utvärderingar. Den sistnämnda ansatsen får dock rimligtvis fungera mer som ett komplement till övriga. Hos en ideal metautvärdering bör all tre ansatser finnas.

Precis som i all forskning kan inte en metautvärderings metoder bestämmas på förhand. När det gäller valet av metoder för en metautvärdering så får detta naturligtvis baseras på aktuell frågeställning och det empiriska materialet. Den rapport som det här avsnittet utgår från är ett exempel på hur kvalitativa och kvantitativa metoder kan kombineras i en metautvärdering och en sekundäranalys. På förhand ligger häri ingen begränsning, i stället bör vikten av stor öppenhet när det kommer till metodval understrykas.

Metautvärdering framstår som en bra metod för att kvalitetssäkra arbetet med utvärderingar av program. Naturligtvis blir detta extra motiverat när programmen handhar stora ekonomiska medel. Även i fall där utvärderare ställts inför stora utmaningar, exempelvis vad gäller utvärdering av särskilt komplexa program, kan uppföljningar i form av metautvärderingar vara motiverade. I många fall kan metautvärderingar vara befogade och i vissa fall är alltså angreppssättet särskilt angeläget för att bedöma kvaliteten hos tematiskt likartade utvärderingar och dra slutsatser om hur utvärderingar kan förbättras.

5.3.2 Sekundäranalys

Sekundäranalys som ansats

Sekundäranalys betyder i allmänhet att data som redan samlats in av andra och gjorts tillgänglig analyseras på nytt. Tillvägagångssättet kan således inte sägas vara någon speciell metod eller teknik. Enligt Glass separeras på initial analys och sekundäranalys på följande vis: *”Primary analysis is the original analysis of data in a research study... Secondary analysis is the re-analysis of data for the purpose of answering the original research question with better statistical techniques, or answering new question with old data...”*⁵⁵

Fördelen med sekundäranalys består i att stora delar av arbetsprocessen redan är avklarade när sekundäranalysen påbörjas. Det gör att nya resultat relativt snabbt kan produceras. Denna fördel är dock förknippad med nackdelen att sekundäranalysen inte har något inflytande över vilken information som samlats in och hur inhämtandet har gått till. Det insamlade materialet kan mycket möjligt ha varit helt tillräckligt för den primära analysen men vara otillräckligt för sekundäranalysen, då denna vanligen är metodologiskt eller statistiskt annorlunda.

En sekundäranalys kan använda data från olika typer av källor. Ofta kan det röra sig om data som samlats in av andra utvärderare eller forskare med en viss typ av analyser i åtanke. Men en sekundäranalys kan också baseras på data som enligt påbud är insamlad av organisationer och vars egentliga syfte är något annat än att utgöra underlag till analyser eller utvärderingar.⁵⁶

Tillvägagångssätt för sekundäranalys

Syftet med sekundäranalysen i ITPS rapport om strukturfondsprogram var att studera utfallet av Mål 1- och Mål 2-program för perioden 2000–2006.

⁵⁵ Glass i: Wolf (1986), sid. 11.

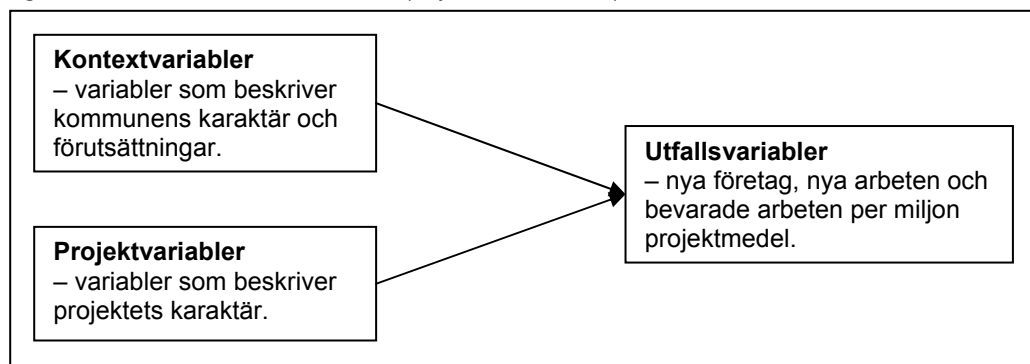
⁵⁶ Bryman (2004).

Sekundäranalysen byggde på data som projektägare rapporterat in till uppföljningsdatabasen Stins.⁵⁷ I databasen återfinns uppgifter om beslutsdatum, utbetalningsdatum, budget etc. Stins innehåller därutöver ett förhållandevis välutbyggt indikatorsystem.

Till sekundäranalysen hämtades information från Stins om det enskilda projektets ort, typ och budget. De indikatorer avseende projektens utfall som i detta fall var särskilt intressanta var de tre kärnindikatorerna: nya företag, nya arbetstillfällen och bevarade arbetstillfällen. Dessa var gemensamma för samtliga program. Kärnindikatorerna relaterades till projektens totala budget vilket skapade de tre relativa utfallsvariablerna: *nya företag per miljon*, *nya arbetstillfällen per miljon* och *bevarade arbetstillfällen per miljon*.

Den teoretiska utgångspunkten var att dessa relativa utfallsvariabler kan påverkas av både projekt- och kontextspecifika variabler, se Figur 5-1. Genom att studera båda variabeltyperna gavs möjlighet att identifiera policyrelevanta interaktionseffekter. En sådan analys skulle till exempel kunna se om vissa typer av projekt fungerar bättre under särskilda strukturella omständigheter. På så sätt skulle man idealt kunna ge rekommendationer om bästa typ av projekt för respektive regiontyp.

Figur 5-1 Utfallsvariabler i förhållande till projekt- och kontextspecifika variabler.



De projektspecifika variablerna utgjordes till exempel av projekttyp, budget och andel privat medfinansiering. Urvalet av dessa variabler begränsades av vad som fanns tillgängligt i Stins. De kontextspecifika variablerna karaktäriserade projektets omgivning. I detta fall matchades variabler på kommunnivå till varje projekt. Varje projekt tilldelades sålunda kontextvariabelvärden beroende på i vilken ort projektet hade sitt säte. Dessa variabler hämtades från en kommundatabas⁵⁸ och kopplades till respektive projekt med kommunkod för projektets huvudsakliga säte (projektägarens adress). De projekt som bedrivit verksamhet i flera kommuner gavs kontextvariabler med utgångspunkt i projektets adress. Ett alternativt förfarande skulle vara att ge dessa projekt genomsnittliga kontextvariabelvärden (baserat på alla de kommuner där projektet i fråga haft verksamhet) eller att helt exkludera projekt med verksamhet i flera kommuner från analysen. Alla tre alternativen har sina styrkor och svagheter. I detta fall bedömdes alternativet med att ge projekten kontextvariabelvärden på basis av projektets adress som mest lämpligt.

⁵⁷ Förkortning för stödinstrument.

⁵⁸ Vad som avses är databasen KFAKTA07, Lunds universitet, statsvetenskapliga institutionen, lektor Leif Johansson.

Konkret innefattade avsnittet med sekundäranalysen både rena beskrivningar och analyser. Inledningsvis redogjordes för utfallsvariablerna utifrån geografiska och institutionella indelningar. Därefter presenterades särskild deskriptiv statistik för respektive utfallsvariabel. Avslutningsvis analyserades variationen i utfallsvariablerna med hjälp av både projekt- och kontextspecifika variabler.

Erfarenheter av sekundäranalys

De erfarenheter och slutsatser som kan dras från den aktuella sekundäranalysen kan kopplas till de allmänna för- och nackdelarna med sekundäranalyser. Ett antal omständigheter är värda att betona. Ett exempel på sekundäranalyserns negativa egenskaper är, som redan nämnts, avsaknaden av inflytande över vilken information som är insamlad och kunskap om hur insamlingen har gått till. Den stora variationen i utfallsvariablerna som erhöles vid sekundäranalysen i det här fallet tyder på, men bevisar inte, att det kan föreligga olikheter mellan de handläggande myndigheternas sätt att bedöma de indikatorer som rapporteras in. Ur ett utvärderingsperspektiv bedöms att uppföljning av programarbetet hade underlättats om deltagande individer och företag gått att spåra i Stins. Detta kunde lämpligen ha skett genom att koppla på personnummer och organisationsnummer.⁵⁹ Sådana rapporteringskrav skulle göra det mycket påtagligare för projektägarna vad som egentligen ligger bakom de siffror de rapporterar. Bedömningen är att möjligheter att identifiera företag och individer skulle ha haft flera positiva konsekvenser för utvärderbarheten av strukturfonderna såsom:

1. Mer realistisk resultatredovisning från projektägare.
2. Möjliggöra identifiering av ”dubletter”.
3. Möjliggöra långsiktigare uppföljning av individer och företag.
4. Möjliggöra utvärdering av ”verkliga” effekter för regioner, företag och individer.

Erfarenheterna från sekundäranalysen av strukturfondernas uppföljningsdatabas, Stins, tyder i stort på att det fanns en bristande överensstämmelse mellan initiala ambitioner och tänkt utvärderingsdesign, samt vad data i Stins kan användas till. Detta innebar förstås begränsningar vid utvärderingsarbetet. Erfarenheterna stämmer mycket väl överens med de allmänna nackdelar som brukar diskuteras i litteraturen. Det bästa vore naturligtvis att möjligheterna att genomföra tillförlitliga utvärderingar prioriterades högre vid planeringen och utformningen av uppföljningsdatabaser. Att till exempel kunna identifiera och följa individer och företag med hjälp av person- och organisationsnummer, skulle således kunna bidra till att framtida utvärderingar underlättas och förmår svara på de mest relevanta frågorna. Erfarenheterna i det aktuella fallet visar att varje sekundäranalys bör inledas med en mycket grundlig genomgång av datakvalitet med särskild uppmärksamhet på brister. Detta kan ge viktig kunskap om hur betydelsefullt det är att tidigt i utvärderingsprocesser göra klart vilken typ av frågor som de data man planerar att samla in kan användas till att besvara. Ambitiösa forskningsdesigner riskerar annars att utlova svar på frågor som det sedan visar sig inte gå att besvara på grund av brister i datamaterialet.

⁵⁹ Enligt personlig kommunikation med Lena Engwall-Skeri på Nutek så finns det möjligheter att inkludera denna typ av data i det nya uppföljningssystemet Nyps.

5.4 Sammanfattning

Under året 2008 genomförde ITPS en så kallad metautvärdering av strukturfondernas Mål 1- och Mål 2-program i Sverige för perioden 2000–2006. Rapporten gavs titeln: *Metautvärdering av strukturfondsprogrammen – nya ansatser och lärdomar*. De empiriska delarna i arbetet utgjordes dels av en metautvärdering, dels av en sekundäranalys. Det här kapitlet har diskuterat dessa båda tillvägagångssätt och beskrivit erfarenheter och lärdomar från arbetet.

Förenklat innebär en metautvärdering en utvärdering av en eller flera utvärderingar. Ambitionen är ofta att på så vis bedöma kvaliteten hos och värdet av redan genomförda utvärderingar. Övergripande skäl att utföra metautvärderingar är enligt litteraturen att undersöka validiteten i utvärderingar, bedöma lämpligheten i utvärderingars utformning och att ta reda på om och hur utvärderingarna har nått sina mål.

Syftet med rapportens metautvärdering var både att bedöma kvaliteten i de studerade utvärderingarna och att försöka få en uppfattning om hur förvaltare av granskade program bedömt nyttan med utvärderingarna. Metautvärdering betraktas i allmänhet som en adekvat metod för att kvalitetssäkra arbetet med utvärderingar av program. Att bedöma kvaliteten i de utvärderingar som ska granskas kan göras på olika sätt. I det här fallet jämfördes utvärderingarna med riktlinjer och upphandlingsunderlag och med allmänna krav som kan ställas på en väl utförd utvärdering och därutöver bedömdes även utvärderingarna mot varandra.

Rapportens andra empiriska del, vilken utgjordes av en sekundäranalys, studerade utfallet av strukturfondsprogrammets kärnindikatorer i relation till satsade medel. Sekundäranalys innebär kortfattat att redan insamlade data analyseras på nytt. I rapporten visas deskriptiv statistik som beskriver situationen för de studerade programmen utifrån geografiska och institutionella indelningar. Variationer i programmets utfall analyserades med både projekt- och kontextspecifika variabler. Erfarenheter från sekundäranalysen tydliggör de nackdelar det innebär att sakna kontroll över vilken data som samlas in och hur datainsamlingen ska gå till. Sekundäranalys skulle generellt underlättas som metod i utvärderingssammanhang om möjligheterna att utvärdera program i större utsträckning prioriterades vid utformning av uppföljningssystem.

6 Metoder för att utvärdera komplexa program

6.1 Vad utmärker komplexa program?

Åtgärdsprogram går mot att bli allt större och mer komplexa, med många olika aktörer och målgrupper. Det här kapitlet behandlar hur komplexa program kan utvärderas. Syftet är att konkretisera vad det är som gör att tillväxtpolitiska program kan vara komplexa och belysa begränsningarna och möjligheterna för utvärdering av dessa program.

Ur utvärderingssynpunkt skapar ett komplext program speciella utmaningar. Det som kännetecknar komplexa program är att det är svårt att entydigt spegla programmets resultat.⁶⁰ En av orsakerna till detta kan vara att det inte är uppenbart hur de kausala sambanden mellan insatser och utfall är utformade.⁶¹ En annan orsak kan vara att programmet innefattar samverkan mellan flera organ och funktioner i samhället. Politiker och beslutsfattare uppmärksammar att verkligheten är komplex och utformar därför program som bygger på partnerskap och nätverk för att kunna hantera flera aspekter av en och samma problematik. Dessa komplexa program syftar till att genom samverkan bättre kunna hantera problematiken än vad enskilda samhällssektorer kan göra. Metoder för att utvärdera en sådan horisontell samverkan behandlas närmare i Kapitel 7.

Utgångspunkten i det här kapitlet är att programmets komplexitet kan härledas från att de olika organisationerna som är involverade i programmet befinner sig på olika geografiska eller organisatoriska beslutsnivåer. De lägre beslutsnivåerna utför ofta programmen på uppdrag av de högre beslutsnivåerna, eftersom de lägre nivåerna oftast har mer information om hur programmet kan genomföras. Programstrukturen följer således subsidiaritetsprincipen: högre beslutsnivåer utför inte det som lägre beslutsnivåer kan göra. Det betyder att programmen decentraliseras och integreras i samhällsstrukturen och då kan både offentliga och privata aktörer med olika perspektiv och agenda involveras. I nästa avsnitt beskrivs dessa omständigheter närmare, varefter det efterföljande avsnittet diskuterar olika angreppssätt och strategier för utvärderingen av komplexa program.

6.2 Komplexa aspekter på den regionala tillväxtpolitiken

De regionala tillväxtprogrammen (RTP) är exempel på program som har komplexa egenskaper. RTP anpassas till regionala förhållanden och syftar till att samordna offentliga och privata aktörer i arbetet med att utveckla lokala arbetsmarknader ute i landets regioner.⁶²

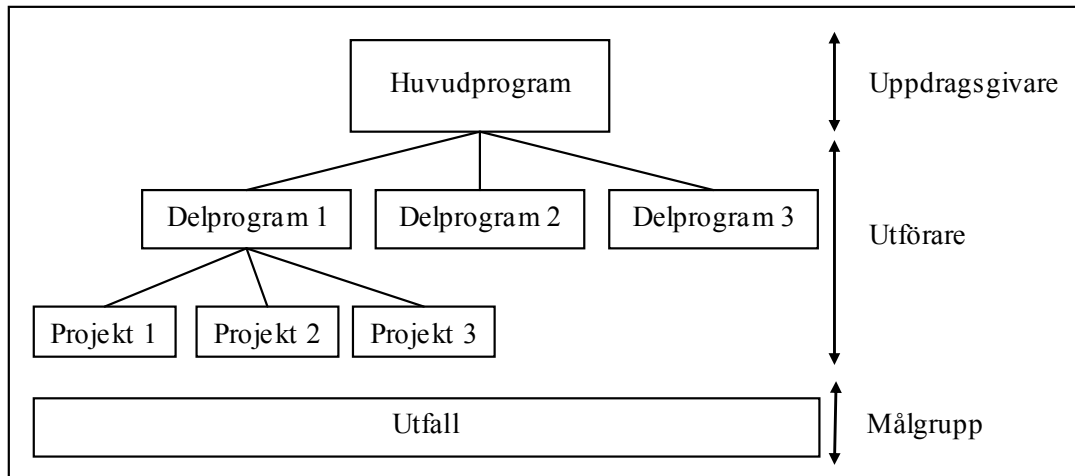
För att åskådliggöra vad det kan innebära att ett program är komplext i den meningen att det har en hierarkisk struktur, utgår vi från Figur 6-1 nedan. Organisationen av ett program kan bestå av ett huvudprogram med de huvudsakliga målsättningarna, planeringen och organiseringen av genomförandet. De regionala programmets utformning och genomförande bestäms av de olika delprogrammen under huvudprogrammet. Delprogrammen genomförs på många olika platser och under olika ledningsformer. Varje delprogram kan i sin tur resultera i ett eller flera lokala eller regionala projekt. Projekten skapar slutligen ett utfall för programmets målgrupp. För enkelhetens skull delar vi in hela programstrukturens aktörer (*stakeholders*) i uppdragsgivare, utförare och målgrupp.

⁶⁰ Se exempelvis Forss (2007).

⁶¹ I Rogers (2008) görs en distinktion mellan komplicerade program med många kausala samband och komplexa program med oklar kausalitet mellan insatser och utfall.

⁶² Nutek (2007).

Figur 6-1 Ett exempel på den hierarkiska strukturen för program.



En utvärdering av ett program med en hierarkisk struktur som den i Figur 6-1 har att hantera flera aspekter som var och en är förknippad med sitt specifika problem:

- *Kausalitet*: Det finns flera möjliga kausala samband mellan utförarnas genomförande av delprogrammen/projekten och utfallet för målgruppen. Flera kausala samband kan vara verksamma samtidigt. Åtgärdernas utfall kan dessutom återverka på åtgärderna och skapa ett cirkulärt samband; en ömsesidig påverkan mellan utförare och målgruppen uppstår. De kausala sambanden i programmet blir därför oklara.
- *Kontext*: Det kausala sambandet kan även bero på den regionala kontexten. Det betyder att det kan finnas alternativa kausala samband som det krävs en förståelse för om man ska kunna uttala sig om hur programmets kontext påverkar programmets funktionssätt. Yttre omständigheter och oförutsägbara faktorer kan påverka programmets utfall på olika sätt i olika regioner.
- *Aktörer*: Gränsdragningen mellan utförare och målgrupp kan vara otydlig, där utförare med olika ekonomiska, sociala, kulturella och fysiska resurser samverkar.
- *Utvecklingsprojekt*: Syftet med ett projekt är ofta att finna nya vägar för att utveckla en samverkan. Arbetssättet och samverkan i regionerna kan inte förutsägas på förhand utan är under utveckling, vilket kan försvåra utvärderingen.

Programmets komplexitet kan ofta härledas från mångfalden av utförare med olika dagordningar och kontexter, samt från att de heterogena målgrupper sammansatt av näringsliv, offentliga organisationer, arbetstagare och egenföretagare leder till att varje aspekt av programmet kan belysas från olika perspektiv.

6.3 Utvärdering av komplexa program

Ett program kan således vara komplext av flera olika anledningar. Oftast har det som sagt sin grund i att många olika aktörer är inblandade. Att utvärdera programmet i sin helhet kan därför visa sig vara mycket svårt. I det här avsnittet beskriver vi tre strategier för att utvärdera komplexa program med en hierarkisk struktur.

6.3.1 En övergripande målutvärdering

Ett sätt att utvärdera hela programmet är att fokusera intresset till relationen mellan de två aktörsgrupperna *uppdragsgivare* och *målgrupp* i Figur 6-1 ovan.⁶³ Utvärderingen får därmed ofta en naturlig avgränsning mot utfallsutvärdering. Att utföra en genomförandeutvärdering av programmet kräver insamling av data från de aktörer som står för dess genomförande. Om utvärderaren bedömer att programmets komplexitet hindrar utvärderingen att med befintliga budget- och tidsramar samla in tillräcklig information om sambanden mellan utförande och utfall, kan det vara bättre att fokusera intresset kring vad uppdragsgivarna har velat uppnå med programmet och hur målgruppen verkar ha påverkats.

Utvärderingens syfte blir att bedöma huvudprogrammets totala måluppfyllelse. Denna utvärdering försöker således inte kartlägga de komplexa aspekterna i programmet, utan söker i stället indicier och belägg för att uppdragsgivarnas målsättning med hela programmet har uppfyllts. För komplexa program kan detta vara en nog så krävande uppgift.

6.3.2 En sammansatt utvärdering av flera nivåer

I en del fall utgör den övergripande målutvärderingen inte ett tillräckligt beslutsunderlag för uppdragsgivarna. Det kan då vara lämpligt att utföra en sammansatt utvärdering av det hierarkiska programmets olika delar och nivåer.

Det kan till exempel vara nödvändigt att komplettera med att utvärdera delegeringen av programmets genomförande från *uppdragsgivare* till *utförare*. Delegeringen kan gå till på olika sätt (frivilligt deltagande, ansökningsprocess, lagstiftning) och motivationen att genomföra aktiviteter enligt huvudprogrammets intentioner kan variera (det kan påverkas av ekonomiska incitament och eventuella målkonflikter mellan nationell och regional nivå m.m.). Förehållandet mellan uppdragsgivare och utförare beskrivs i litteraturen som *principal-agent* problematiken och kan kännetecknas av en hög grad av asymmetrisk information.⁶⁴ Utvärderingar med sådant syfte som det här beskrivna kan klassificeras som genomförandeutvärderingar.

Vidare kan relationen mellan aktörsgrupperna *utförare* och *målgrupp* utvärderas för enskilda projekt eller delprogram. På dessa nivåer kan det ofta vara möjligt att utföra inte endast genomförande- och målutvärderingar, utan även effekt- och effektivitetsutvärderingar. Dessa utvärderingar av enskilda projekt blir fallstudier vars resultat är mer eller mindre generaliserbara för flera projekt. Det är därför viktigt att bedöma hur generaliserbara dessa olika fallstudier är.

Den sammansatta utvärderingen består alltså av tre delar: övergripande målutvärdering, utvärdering av delegeringen från uppdragsgivare till utförare och utvärderingar av enskilda projekt eller delprogram. När de tre utvärderingarna har utförts kan de tillsammans ge en sammansatt bild över hela programmet. Denna sammansatta utvärdering kommer troligtvis

⁶³ Tuynman et al (2008).

⁶⁴ Se till exempel Goldmann et al (1997).

att bestå av målutvärderingar, genomförandevärderingar, samt effektutvärderingar av enstaka projekt. Det blir på så sätt en kombination av kvantitativa och kvalitativa metoder (*mixed methods*) som får forma den övergripande utvärderingen av hela programmet.

6.3.3 Behovet av programteori

Programteori har det övergripande syftet att verifiera en kausalitet inom programmet/projektet och att dra generella slutsatser om hur programmet/projektet fungerar.⁶⁵ Teorin strävade ursprungligen efter att vara fri från överväganden om hur aktörernas värderingar påverkade programmet och betydelsen av politiska överväganden. Därmed uppstod kritik mot programteorin för att den inte kunde belysa programmets hela komplexitet, utan låste in utvärderaren i förutfattade meningar om hur verkligheten borde se ut.⁶⁶ Under senare år har programteori utvecklats till att i större utsträckning beakta värderingar och politiska processer. Flera kvantitativa och kvalitativa metoder kan användas för att försöka verifiera teorin om hur programmet fungerar. Samtidigt utbildas aktörerna i att bättre förstå och hantera hur värderingar påverkar programmet och hur politiska processer påverkar programmet.⁶⁷

Modern programteori kan vara ett bra verktyg för utvärderingens kartläggning av hur lokala eller regionala projekt fungerar i praktiken. Denna kartläggning kan vara nödvändig för att förstå hela programmets komplexitet. Utvärderingen av projekt längst ned i programhierarkin kan således tjäna som en introduktion till utvärderingen av hela programmet. Den vertikala kommunikationen mellan utvärderingar av lokala projekt och mer övergripande programutvärderingar för regioner eller nationer blir av central betydelse. Vad har utvärderingen på lokal nivå kommit fram till? Överensstämmer resultaten med målsättningen på en regional/nationell nivå? Den institutionella strukturen för utvärderingsorganen är därför avgörande för att det av genomförda utvärderingar ska kunna utläsas om ett program verkligen påverkar målgruppen. Om det lokala utvärderingsorganet inte kan kommunicera sina utvärderingsresultat till de regionala eller nationella utvärderingsorganen, minskar avsevärt möjligheterna att nå en gedigen utvärdering av hela det komplexa programmet. Programteori som belyser de lokala delprogrammets eller projektens funktionalitet och hur de påverkas av den lokala kontexten ger kunskaper som är nödvändiga för förståelsen för hur hela det komplexa programmet fungerar. Utan sådana kunskaper är utvärderingen högre upp i den hierarkiska strukturen hänvisad till indikatorer och statistiska metoder som har svårt att belysa det komplexa programmets resultat.⁶⁸

Ett exempel som visar på vikten av lokala utvärderingar och deras betydelse för den centrala utvärderingen, är att metoderna för de centrala utvärderingarna av strukturfonderna kommer att ändras.⁶⁹ Från att ha tillämpat en utvärderingsorganisation med fristående utvärderingar som vägs samman till en övergripande utvärdering kommer den nya utvärderingarna (från oktober 2007 till december 2009) i stället att inriktas mot programområden, med en förbättrad vertikal kommunikation. Detta visar att insikten om vikten av kontakt mellan lokala utvärderingar som kan kartlägga programmets komplexitet och centrala utvärderingar har fått genomslag.

⁶⁵ Se till exempel Vedung (1998) och Mayne (2001). Andra namn för programteori är *contribution analysis*, *verksamhet logik*, eller *händelsekedja*.

⁶⁶ Stufflebeam (2001b).

⁶⁷ Chen och Rossi (1989), Weiss (1995) och Pawson och Tilly (1997).

⁶⁸ Stame (2004).

⁶⁹ Gaffey (2008).

6.4 Sammanfattning

Åtgärdsprogram går mot att bli allt större och allt mer komplexa, med många olika aktörer och målgrupper. Att utvärdera komplexa program är en stor utmaning och kräver ett öppet synsätt inför nya metoder och angreppssätt. Syftet med det här kapitlet är att konkretisera vad det är som gör att tillväxtpolitiska program kan vara komplexa och att belysa begränsningar och möjligheter för utvärderingen av dessa program.

Ett komplext program kan beskrivas som en hierarkisk struktur, med ett huvudprogram som delegerar utformning och genomförande till delprogram och lokala projekt. Programmets komplexitet kan bero på att flera olika aktörer inom denna hierarkiska struktur samverkar, samtidigt som kausaliteten mellan aktiviteter och utfall är oklar, samt av att det finns en tydlig påverkan av kontextuella faktorer.

Vi föreslår tre strategier för att utvärdera komplexa program. Den första strategin är att göra en övergripande målutvärdering. Utvärderingens syfte blir att bedöma huvudprogrammets totala måluppfyllelse. Denna utvärdering försöker således inte kartlägga de komplexa aspekterna i programmet, utan söker i stället indicier och belägg för att *uppdragsgivarnas* målsättning med hela programmet har uppfyllts för *målgruppen*.

Den andra strategin försöker nå längre än till en övergripande målutvärdering, genom att utvärdera olika delar och nivåer av programmets hierarkiska struktur. En sådan utvärdering blir en sammansatt utvärdering av olika nivåer, t.ex. genomförandevärderingar av relationen mellan uppdragsgivare och utförare, samt effektutvärderingar av enskilda projekt.

Den tredje strategin använder sig av modern programteori som i vårt fall appliceras på lokala utvärderingar av enskilda projekt, där det beaktas att programmets kontext påverkar aktörernas förutsättningar i programmet. Utvärderingen av projekt längst ned i programhierarkin tjänar som en introduktion till utvärderingen av hela programmet. Den vertikala kommunikationen mellan utvärderingar av lokala projekt och mer övergripande programutvärderingar för regioner eller nationer blir av central betydelse. Denna tredje strategi är den som befinner sig i utveckling och kan vara den mest krävande av de tre föreslagna strategierna.

7 Metoder för att utvärdera horisontell samverkan

7.1 Inledning

Det är ingen tvekan kring att termen ”governance” i dag är flitigt använd i diskussioner som förs ur ett samhällsvetenskapligt perspektiv. Enligt Pierre och Peters betyder ”governance” dels hur samhället och ekonomin styrs och dels hur kollektiva mål kan uppnås.⁷⁰ Om man tar begreppet governance som utgångspunkt för diskussioner om hur samhället i dag styrs och organiseras blir det tydligt att denna process sker på olika nivåer (så kallad multilevel governance) och innefattar en stor flora av aktörer där staten endast är en. Den moderna förvaltningens mål- och resultatstyrning samt ett uppifrån-och-ned-perspektiv har i vissa fall skapat en rigid och vertikal organisering av verksamheter med barriärer mellan olika sektorer i samhället.⁷¹ En stor del av dagens forskning på det här området handlar om hur denna sektorisering ska kunna brytas och hur samverkan mellan olika aktörer effektivt ska kunna ske. En ofta framförd lösning på en alltför stark sektorisering är vad som brukar benämnas som horisontell samverkan. Det här kapitlet har till syfte att översiktligt diskutera hur olika former av horisontell samverkan inom offentlig verksamhet kan ta sig uttryck och hur dessa kan utvärderas.

Samverkan delas ofta upp i vertikal och horisontell sådan, där som sagt detta kapitel är avgränsat till det sistnämnda slaget. Många samverkansprocesser inkluderar dock båda dimensionerna då de vanligen är interdependenta. Den svenska regionala tillväxtpolitiken innehåller många exempel på samverkan, både vertikal och horisontell samt kombinationer av de båda. De tematiska myndighetsgrupperna grundade på prioriteringar i den nationella strategin för det regionala utvecklingsarbetet respektive det nationella forumet för regional konkurrenskraft, entreprenörskap och sysselsättning är två exempel på samverkan inom politikområdet. Ett annat är de regionala tillväxtprogrammen. Dessa, vilka bedrivits i samtliga svenska län, är utformade utifrån värdet av att samla och samordna insatser som en grund för att skapa hållbara tillväxtförutsättningar. De regionala tillväxtprogrammen är således ett gott exempel på hur samverkan kan användas som en ansats inom regional tillväxtpolitik.⁷²

Framställningen inleds med att i 7.2 diskutera särskilda begrepp som är centrala i forskningen kring horisontell samverkan. I samma avsnitt nämns också olika typer av sådan samverkan och på vilka nivåer detta kan utspelas. I avsnitt 7.3 diskuteras vilka särskilda konsekvenser som horisontell samverkan ger ur ett utvärderingsperspektiv. I avsnitt 7.4 ges exempel både på hur horisontell samverkan i praktiken kan gå till väga och hur uppföljning och utvärdering kan ske. Kapitlet avslutas med en sammanfattning.

7.2 Relevanta begrepp och former av horisontell samverkan

Svårigheter med samverkan och koordinering inom staten framstår som ett evigt dilemma. Tidigare forskning har även noterat att trenden är negativ. Många genomförda förändringar

⁷⁰ Pierre och Peters (2000), sid. 1.

⁷¹ I Sammanhanget brukar synsättet *new public management* framhållas. Vad gäller den vertikala strukturen kan begrepp som ”stuprör” och ”silos” noteras. Se: ITPS (2008c); ITPS (2008g); Statskontoret (2007).

⁷² Finansdepartementet (2008).

inom den moderna förvaltningen har nämligen ytterligare försvårat målet med en sammanhållen och koordinerad förvaltning.⁷³

Internationellt används ett stort antal termer inom forskningsfältet för att beskriva samverkan. Det kan delvis förklaras med respektive lands ambition att lansera egna lösningar på problemet med bristande samverkan. Inledningsvis bör man skilja de engelska uttrycken *cooperation* och *collaboration* från varandra. Med *cooperation* menas ömsesidigt formaliserat samarbete mellan parter utifrån gemensamma intressen. *Collaboration* är däremot ett vidare begrepp och innefattar hela spannet från löst knutna samarbeten till officiella partnerskap. Av de mer precisa begrepp som används i forskningsfältet framstår *joined-up government* respektive *horizontal government* som två av de mest centrala. Det förstnämnda har sin bakgrund i den brittiska statsförvaltningen men har sedermera spritts till många andra länder. *Joined-up government* definieras som konsistens mellan organisatoriska arrangemang av program, policy eller aktörer vilket möjliggör samarbete.⁷⁴ *Horizontal government* är en term som bland annat återfinns i den kanadensiska statsförvaltningen. Det kan definieras som koordinering och ledning av aktiviteter mellan två eller flera aktör mellan vilka det inte råder någon tydlig hierarkisk ordning och vilka har ett mål att prestera som inte vore möjligt att nå utan samarbete.⁷⁵ Detta går i linje med Peters resonemang, det vill säga att en framgångsrik koordinering i förvaltningen lyckas minimera omfattningen av dubbelarbete utan att för den sakens skull riskera att ingen aktör tar ansvar för arbetsprocessen och ett tomrum uppstår eller att brist på samstämmighet råder mellan deltagarnas målsättningar och utförande.⁷⁶

Den svenske forskaren Quist har utkristalliserat fyra former av horisontell samverkan. Den första formen – *samverkan utifrån sammanhållen process* – kännetecknas av att två eller flera aktörer tillsammans levererar en tjänst till medborgare. Den andra varianten benämns *samverkan utifrån målgrupp* och utgörs av att flera aktörer samverkar för att förbättra för en identifierad målgrupp. Till skillnad från den första formen av samverkan karaktäriseras inte arbetet i den andra formen av en tydlig värdeskapande kedja utan grundas i stället på förståelse hos de olika aktörerna för varandras arbete, något som kan minska förekomsten av överlappande arbete. Den tredje varianten – *samverkan utifrån gemensam vision* – saknar en gemensam målgrupp men kan exempelvis komma till uttryck i att insatser syftar till att i visst eller vissa avseenden förbättra eller vårda ett geografiskt område. Den minst avgränsade formen betecknas *horisontella frågor* och är inriktad mot stora frågor såsom jämställdhet och demokrati vilka många aktörer bär ett ansvar för. Termen för när vissa, mer övergripande frågor, ska innefattas i övrig politik är policyintegrering.⁷⁷

Flera författare diskuterar hur samverkan sker på olika nivåer i förvaltningen. I en indelning används de tre nivåerna *politisk*, *policy-* och *administrativ nivå*. Studier på den förstnämnda nivån gäller hur de övergripande ramarna skapas för de lägre nivåerna. Undersökningar avseende den mellanliggande nivån rör hur samverkan sker bland de policyskapande aktörerna inom och mellan departement. När den administrativa nivån studeras så behandlas var den beslutade politiken i själva verket genomförs och hur olika aktörer samverkar för att åstadkomma fattade beslut. Somliga menar att samverkan är lättare på

⁷³ Peters (1998).

⁷⁴ 6 (2004), sid. 106.

⁷⁵ Bakvis och Juliatt i: Quist (2005).

⁷⁶ Peters (1998).

⁷⁷ Quist (2005), sid. 13–15; Peters (1998). För policyintegrering och hur detta kan avgränsas och utvärderas se: Mickwitz och Kivimaa (2007).

denna nivå då implementering till sin natur kräver ett mer pragmatiskt förhållningssätt. Därutöver väljer vissa teoretiker att skilja på sådana horisontella relationer inom förvaltning som kan benämnas som intern samverkan och sådana former där offentliga aktörer samverkar med andra, extern samverkan. Vid extern horisontell samverkan blir exempelvis frivilliga organisationer och andra privata aktörer centrala deltagare.⁷⁸

7.3 Konsekvenser för utvärdering

7.3.1 Svårigheter för utvärdering

Utvärderingar av horisontell samverkan präglas ofta av ett antal svårigheter som kan härledas till de fenomen som studeras. Att utvärdering av samverkan skiljer sig från annan utvärdering beror på att samverkansprocesser har sin utgångspunkt i eller förutsätter en helhetssyn. En utmaning för utvärderingar av samverkan kan, beroende på utvärderingens syfte, vara att bedöma det reella mervärde som samverkan haft för utfallet. Metodologiskt är detta svårt eftersom det förutsätter en jämförelse med det kontrafaktiska utfallet. Samverkan upptar naturligtvis arbetstid och ger även upphov till andra kostnader. Forss menar att den nytta, i förhållande till kostnaden, som uppnås vid samverkan av en viss omfattning är snabbt avtagande. Till slut nås en punkt där en ökad grad av samverkan kostar mer än vad den ger tillbaka. Forss riktar en skepsis mot utvärderingar som inte gör någon grundlig genomgång av om samverkansprocesser har nått en optimal omfattning men som trots det ger patentsvaret att ”samverkan bör stärkas”.⁷⁹

I praktiken innebär de metodologiska svårigheterna att utvärdering av horisontell samverkan vanligen inriktar sig på genomförande och inte utfall. En ytterligare faktor som förklarar inriktningen mot genomförande och inte utfall är att sätta målformuleringar sällan är tillräckligt konkreta.⁸⁰

Ett vanligen framfört dilemma i diskussioner som anknyter till begreppet governance handlar om att ansvarsutkrävande försvåras i och med etablerandet av nya arbetssätt i den offentliga sfären. Oavsett formerna för horisontell samverkan kan det bli problematiskt att redogöra för vilka aktörer som gjort vad, något som kan ge svårigheter för staten som ytterst ansvarig för hur offentliga medel används. Michels och Meijer noterar emellertid att det har utvecklats mekanismer för att tillgodose att ansvarsutkrävande kan ske. Det gäller i synnerhet samverkan på mellanliggande och lägre nivåer. Konkret kan det innebära att administrativa organ skapas för att tydliggöra att en enskild aktör tar ansvar inför bland annat medborgare och media.⁸¹ Även ur ett utvärderingsperspektiv framhålls lösningar på problemet. Det sägs kräva en pluralistisk utvärderingsansats med insikter om att relationen mellan insats och utfall även bestäms av interaktioner och förhandlingar mellan olika aktörer.⁸²

Utformningen av samverkan med avseende på den distinktion som diskuterades i 7.2, collaboration kontra cooperation, kan ha betydelse för möjligheten att utvärdera horisontell samverkan. En samverkansprocess som tar formen av cooperation innebär en mer avgränsad samarbetsform där deltagande aktörer har en hjälpsam inställning till varandra och

⁷⁸ Quist (2005), sid. 15–16; Michels och Meijer (2008); Agranoff och McGuire (2003); Kooiman (2003).

⁷⁹ Forss (2007), sid. 85–87; Quist (2007).

⁸⁰ Stame (2006).

⁸¹ Michels och Meijer (2008).

⁸² van der Meer och Edelenbos (2006).

samarbetet. Detta kan underlätta utvärdering. Samverkan genom collaboration kan innefatta hela spannet från löst knutna samarbeten till högt formaliserade. Ur ett utvärderingsperspektiv blir då förutsättningarna helt olika från fall till fall.

7.3.2 I forskningen diskuterade angreppssätt

De speciella hänsynstaganden som bör tas vid utvärdering av samverkan av horisontell natur får sökas på olika håll både bland forskning och hos genomförda utvärderingar. Det existerar nämligen ingen beprövad och enhetlig metod för att utvärdera samverkan. Detta beror rimligtvis på att horisontell samverkan kan ske på många olika sätt. Nedan följer exempel både på hur man i utvärderingsteori samt i faktiska utvärderingar hanterat den problematiken. van der Meer och Edelenbos presenterar i en artikel rekommendationer för utvärdering av policyområden med många deltagande aktörer. Ett antal av dessa rekommendationer är särskilt värda att uppmärksamma. Författarna föreslår framför allt att olika tillvägagångssätt bör kombineras. Det kan innebära att kvantitativa och kvalitativa metoder används parallellt eller att olika typer av utvärderingar genomförs. Det sistnämnda kan innebära en sådan kombination av utvärderingar att både utfallet av en åtgärd och det genomförande som resulterat i utfallet studeras. Ett annat, mer nydanande förslag är att i fall där resultat inte tydligt kan tillskrivas en enskild aktör låta utvärderingen vila på ett antal samarbetande aktörer. På så sätt kan även utvärderande aktörer nå samverkansvinster. Det innebär att de deltagande aktörerna gemensamt ansvarar för utvärderingen, något som även kan stärka trovärdigheten av resultaten.⁸³

I ett aktuellt forskningsbidrag diskuterar Jobin utvärderingsstrategier för partnerskap. Efter att ha definierat partnerskap som ”*a joint effort toward common public policy objective(s), by means of shared governance based in incomplete written agreement*”⁸⁴ argumenterar författaren för att vid utvärdering av partnerskap använda tillvägagångssätt hämtade från institutionell ekonomi. Mer detaljerat framhålls att variationer inom institutioner kan förklaras av ekonomiska beräkningar av transaktionskostnader. Baserat på detta menar Jobin att två perspektiv bör inkluderas vid utvärdering av partnerskaps resultat. Till att börja med är det centralt att koordineringen bland partnerskapets olika aktörer studeras. Ambitionen är att minska sådana transaktionskostnader som kan hänföras till aktiviteter som gäller koordinering. Genom att mäta hur intensivt och effektivt denna koordinering sker kan transaktionskostnader uppskattas. Den andra aspekten som Jobin understryker rör hur förekomsten av socialt kapital påverkar transaktionskostnader. En hög grad av socialt kapital minskar ett partnerskaps kostnader för koordinering. Avslutningsvis noteras att beräkning av transaktionskostnader antingen kan bli ett kompletterande moment i utvärderingar eller bli den centrala aspekten vid utformning av utvärderingar. Ansatsen kan även användas för att utforma en effektivitetsutvärdering, då transaktionskostnaderna relateras till ett partnerskaps utfall.⁸⁵

Så kallade teoriorienterade ansatser inom utvärdering har i hög grad anammats i särskilt komplexa sammanhang där exempelvis partnerskap och nätverk representerar de organisatoriska formerna. Ett övergripande motiv för användningen av sådana ansatser är att de är redskap för att kunna öppna och beskriva innehållet i programmets ”svarta låda”. Teoriorienterade utvärderingar vill således, genom att titta in i lådan, ge kunskap om omvandlingen av input till output och utgår från en på förhand formulerad modell som försöker

⁸³ van der Meer och Edelenbos (2006).

⁸⁴ Jobin (2008).

⁸⁵ Jobin (2008).

redogöra för kausaliteten. På så sätt tillåts teori bilda plattform för en utvärderings beskrivande och förklarande ansatser, något som styr datainsamling och tolkning av resultat. Stame finner ett antal gemensamma drag för teoriorienterade ansatser. I linje med de uppslag som ges av teoretiska utgångspunkter låter man utvärderingen vila på antaganden om vad som kan komma att ske. Vidare betonas vikten av ett programs kontext och dess betydelse för utförandet. Därutöver finns en öppenhet inför alla lämpliga metoder då man låter teorin föranleda och styra metodvalet. Avslutningsvis har, som sagt, teoriorienterade utvärderingar ett särskilt fokus på kausalitet vilket är skälet till att man behöver ta reda på innehållet i ett programs ”svarta låda” – det gäller ju att utröna orsakssamband mellan vad som stoppas in i och vad som kommer ut ur lådan. Stame noterar att i dagsläget är det framför allt genomförandeutvärderingar som är frekventa när det kommer till granskning av program med många inblandade aktörer. Det behövs dock en större vilja att kombinera olika ansatser, vilket skulle göra det möjligt att samtidigt utvärdera både samarbetsformerna i sig och utfallet av dem.⁸⁶

7.4 Två exempel på horisontell samverkan och uppföljning och utvärdering av dessa

I följande avsnitt presenteras två exempel på horisontell samverkan. Det första, de finska *politikprogrammen*, utgörs av samverkan på departementsnivå medan det andra, utvecklingsprojektet *VISAM*, rör samverkan mellan svenska myndigheter. Både utformningen och innehållet av samverkansprocesserna respektive arbetet med att följa upp och utvärdera desamma diskuteras.

7.4.1 De finska politikprogrammen

De finska *politikprogrammen* lanserades 2003 efter flera år av förberedelse och analysarbete. De hade sin bakgrund i en uppfattning om att regeringen hade brister i sin horisontella samverkan och blivit alltför sektoriserad i sitt arbete. Politikprogrammen är tväradministrativa och består av åtgärder, projekt och anslag och leds av en utsedd minister. Målsättningen med arbetssättet är tudelad: att förbättra politisk styrning och stärka det horisontella perspektivet. Under perioden 2003–2007 var politikprogrammen fyra till antalet: informationssamhällsprogrammet, sysselsättningsprogrammet, politikprogrammet för företagsamhet och politikprogrammet för medborgarinflytande. Specificering av politikprogrammets genomförande och uppföljning framläggs årligen i ett så kallat strategidokument där även andra tväradministrativa politikområden specificeras. Vad gäller utvärdering av programmen följs utfall vad gäller måluppfyllelse upp med hjälp av indikatorer. På så sätt ämnar man bedöma de samhällsliga effekterna av programmet.⁸⁷

I en artikel från 2004 görs några reflektioner kring utvärdering av dessa program. Man använder där termen effektivitet med hänsyftning på i vilken omfattning politikprogrammen nått specificerade mål. (Begreppet har alltså här en annan innebörd än den i utvärderingssammanhang gängse.) Utfallet vad gäller effektiviteten hävdas ha en central betydelse för utformning av framtida riktlinjer för programmen. Artikelförfattarna framhåller att de till programmen kopplade målen är så pass konkreta och tydliga att de med hjälp av indikatorer på ett adekvat sätt kan följas upp. Detta är nödvändigt för framgångsrikt framtida policyskapande och möjliggör att de mest brådskande åtgärderna snabbt kan vidtas av beslutsfattare. Förutom uppföljning av de uttryckta målen, med hjälp av indikatorsystemet,

⁸⁶ Stame (2004); Stame (2006); Langerth Zetterman och Strömdahl (2001).

⁸⁷ Statsrådets kansli (2004).

efterfrågar artikelförfattarna kvalitativa ansatser, framför allt för att skapa kunskap kring relationen mellan åtgärdens direkta utfall vad gäller realiserandet av målen och den önskade samhällsförändringen. Artikelförfattarna avslutar med att lyfta fram ett antal frågor. Dessa rör exempelvis hur man ska identifiera de mest lämpliga måtten för måluppfyllelse inom respektive politikområde samt hur informationen från utvärderingar ska återföras till politiker.⁸⁸

I en senare publikation diskuteras erfarenheterna av politikprogrammen i ljuset av genomförda utvärderingar. En genomförandeutvärdering samlade information om deltagande parter uppfattning av programmet, både genom intervjuer och enkäter. Resultaten visade att det horisontella arbets sättet med politikprogrammen i stora drag bedöms som positivt. Den allmänna uppfattningen var att politikprogrammen har förbättrat effektiviteten vad gäller hanterandet av horisontella frågor. Samtidigt betonades att det är svårt att mäta det eventuella mervärdet av ett dylikt tillvägagångssätt. Det kan också finnas en inneboende problematik när det gäller att bestämma vilka positiva utfall som kan tillskrivas respektive deltagande aktör.⁸⁹

Från och med 2007 arbetar regeringen med de tre politikprogrammen för arbete, företagande och arbetsliv; för hälsofrämjande; samt för barns, ungas, och familjers välfärd. Precis som tillvägagångssättet under 2003–2007 används distinkta mål som mäts med indikatorer. Vissa av indikatorerna är så kallade processindikatorer som avser att mäta hur politikåtgärderna framskrider. För uppföljning används både statistiska uppgifter och enkäter. Uppföljningsarbetet bedrivs regelbundet inom regeringskansliet och initieras formellt av statsministern.⁹⁰

7.4.2 VISAM – Vi samverkar

”VISAM – vi samverkar” var ett av den svenska regeringen initierat projekt för att utveckla, pröva och etablera gemensamma kanaler för den offentliga förvaltningens möte med medborgare och företag. Projektet som bedrevs 2003–2005 skulle även utveckla, belysa och pröva samverkan över myndighetsgränser för att utveckla service-, organisations- och arbetsformer.⁹¹ Medverkande myndigheter i projektet var Arbetsmarknadsstyrelsen, Centrala Studiestödsnämnden, Försäkringskassan, Skatteverket, Tullverket och Migrationsverket. I projektet deltog även Statskontoret genom att tillhandahålla sekretariatsresurser och sprida erfarenheter från projektet.

Projektet ledde till samverkansarbete i 15 olika delprojekt. I en rapport från statskontoret beskrivs delprojekten och deras resultat.⁹² Initialt var tanken att till VISAM koppla ett uppföljningssystem med indikatorer på måluppfyllelse och nytta. Detta visade sig vara svårt då delprojekten hade mycket olika mål, löpte under olika tider och inte hade en likartad organisatorisk ram. Uppföljningssystemet har i stället inkluderat löpande rapportering från respektive delprojekt om hur arbetet fortskred och hur man närmade sig respektive mål. En särskild del av detta arbete har varit delprojektledarnas synpunkter på hur samordningen fungerat, med avsikten att göra en processbedömning. När delprojekten avslutats bedömdes delprojektens måluppfyllelse utifrån hur man uppnått uppdragets tre kärnsyften – utveckla, pröva och etablera nya former för servicesamverkan. Den regel-

⁸⁸ *Harrinvirta och Kekkonen (2004).*

⁸⁹ *Kekkonen (2006).*

⁹⁰ *Statsrådets kansli (2007); Kekkonen (2006).*

⁹¹ *Finansdepartementet (2003).*

⁹² *Se: Statskontoret (2005).*

bundet insamlade informationen från delprojekten samlades i tertialrapporter i vilka ett genomgående inslag har varit en successiv utvärdering av samverkansprocesserna, mer exakt vad gäller styrning, samordning och engagemang.⁹³

I sin rapport sammanställer Statskontoret lärdomar och resultat av VISAM baserat på uppföljningssystemet och delprojektens egna rapporteringar. Sett i relation till det övergripande syftet kan man notera att utvecklandet av former för samverkan inte varit speciellt problematiskt. Svårigheterna har i stället legat i att pröva och etablera motsvarande samverkan, det vill säga att gå från tanke till handling. Studerar man de 15 delprojektens målbeskrivningar bedöms 8 ha nått egna formulerade mål. Det påpekas emellertid i rapporten att man vid målformuleringarna inte tillräckligt har beaktat den begränsade tidsramen och satsade medel. En slutsats är också att måluppfyllelse och resultat bör ställas i relation till varje delprojekts kostnader. En bedömning av VISAM:s kostnadseffektivitet kräver dock att varje delprojekt granskas utifrån egna förutsättningar. Sammantaget menar emellertid Statskontoret att nyttan av VISAM har gått att påvisa med relativt enkla medel. Som underlag till detta har man använt dels VISAM:s regelbundna rapporteringssystem, dels modeller för att exempelvis beräkna tidsvinster och antalet nådda medborgare. På så vis har man kombinerat olika metoder för att ge en bild av utfallet.

Ett antal slutsatser kan dras på basis av VISAM. Förutom att mål och visioner ska vara klara och tydliga måste de även delas av de ingående parterna. Vad gäller möjligheten till ansvarsutkrävande har delprojekten olika lösningar. Summerat kan man notera både positiva och negativa arrangemang. Slutsatsen blir att lösningen är mycket beroende av situation och sammanhang. En annan betonad erfarenhet är att projektstyrningen varit oklar, något som i mångt och mycket bedöms botten i avsaknaden av en modell för projektstyrning. Även särskild lagstiftning har försvårat samverkan; det gäller när myndigheter samarbetat i hanteringen av ärenden. Ytterligare en viktig slutsats är att resultatstyrningen av myndigheter har försvårat arbetet. Som det sägs i statskontorets rapport: *Myndighets-specifika mål leder till att myndigheterna strävar åt olika håll.*⁹⁴

7.5 Sammanfattning

När det har diskuterats, föreslagits och prövats lösningar på problem förknippade med den alltmer tilltagande sektoriseringen i många av världens statsförvaltningar har uttryck som *collaboration* och *cooperation* getts allt större utrymme. På svenska används ofta begreppet samverkan för att täcka båda dessa termer och beteckna motsvarande arbetssätt det vill säga formaliserat samarbete (*cooperation*) eller mera löst samarbete (*collaboration*). Samverkan kan avse både vertikala och horisontella relationer men i praktiken rör det sig vanligen en kombination av båda. Det här kapitlet har behandlat utvärdering av horisontell samverkan.

Ett antal faktorer, av vilka flera ligger i sakens natur, försvårar utvärdering av horisontell samverkan. Det rör sig exempelvis om sådant som otydliga målformuleringar och svårigheter att utkräva ansvar men också om metodologiska problem, framför allt när det gäller att bedöma det eventuella mervärde som horisontell samverkan kan tänkas ge upphov till. Som förslag på lösningar föreslår aktuell forskning bland annat en större teorimedvetenhet och öppenhet vad gäller val av metoder och utnyttjandet av kombinationer av metoder.

⁹³ Forss (2007); Statskontoret (2005).

⁹⁴ Statskontoret (2005), sid. 241.

De tväradministrativa finska politikprogrammen presenterades som en lösning på den alltmer tilltagande sektoriseringsen i landets förvaltning. Med politiska mål som inbegriper flera departement krävdes samverkan. För utvärderingen av politikprogrammen utarbetades indikatorer för att mäta måluppfyllelse och en genomförandevärdering utfördes för att samla deltagande aktörers erfarenheter av det horisontella arbetssättet. Uppföljningen av de nuvarande politikprogrammen innefattar även särskilda processindikatorer för att kunna mäta hur politikåtgärderna framskrider. Ett annat exempel på horisontell samverkan är det svenska projektet ”VISAM – vi samverkar”. Uppgiften för detta var att utveckla, belysa och pröva samverkan i mötet med medborgare och företag. Erfarenheterna från VISAM har tillvaratagits genom regelbunden rapportering från programmet och delprojekten. Särskilt diskuterat är vikten av tydliga och rimliga målformuleringar. Nyttan av att kombinera metoder har framgått i utvärderings- och uppföljningsarbetet.

Detta kapitel har tydliggjort att det inte finns någon universallösning för hur horisontell samverkan ska utvärderas. Det beror på det faktumet att samverkansprocesser har en hög grad av komplexitet. Ett antal omständigheter, av vilka samtliga har berörts i avsnittet, bedöms vara försvårande för utvärderingar. Förutom svårigheter med att avgöra i vilken utsträckning som olika aktörer har deltagit och att skapa system för att möjliggöra ansvarsutkrävande, är en avgörande fråga hur värdet av samverkansprocesser påverkar resultat av program eller åtgärder.

Enligt aktuell forskning är ett antal faktorer särskilt värda att betona. I forskningen framhålls att teoriorienterade utvärderingar blir särskilt påkallade i svårare utvärderingssammanhang. Oavsett i vilken utsträckning man anammar användandet av en teoriorienterad ansats kan överväganden som härrör från denna inriktning vara relevanta. Det gäller framför allt vikten av att beakta samverkansprocessers kontext och låta teori styra metodval. Forskningen belyser vikten av att kombinera både olika typer av utvärderingar och kvantitativa ansatser med kvalitativa. På så sätt ges ökade möjligheter att besvara frågor rörande samverkans betydelse för att nå uppställda mål och ibland även rörande effekter och effektivitet.

8 Sammanfattande slutsatser och rekommendationer

Den regionala tillväxtpolitiken består av många olika typer av åtgärder och är av stor betydelse för den nationella tillväxten och den regionala utvecklingen. Att adekvat kunna utvärdera den regionala tillväxtpolitiken är naturligtvis av stor betydelse och ITPS har därför fått i uppdrag att utveckla metoder för att utvärdera program och projekt inom den regionala tillväxtpolitiken.

Vetenskapligt arbete utgår i regel från en specifik frågeställning varefter denna får styra valet av lämplig metod. Att föreslå att en viss specifik metod alltid bör användas är därför inte lämpligt eller ens möjligt. Det finns däremot allmänna utgångspunkter som bör beaktas för att utvärderingar ska bli tillförlitliga och användbara. Utvärderingar är i allmänhet inriktade mot att studera antingen *utfallet* eller *genomförandet* av en åtgärd. Utvärderingar som studerar utfallet delas in i mål-, effekt- och effektivitetsutvärderingar. En målvärdering svarar på frågan om uppställda mål har uppfyllts; en effektutvärdering om måluppfyllelsen är en effekt av åtgärden; en effektivitetsutvärdering om de uppnådda effekterna värderas högre än kostnaderna för åtgärden; medan slutligen en genomförandevärdering avser att besvara frågan om hur själva genomförandet kan ha påverkat måluppfyllelse, effekt och effektivitet. Det är som regel en mycket stor fördel om en utfallsutvärdering kombineras med en genomförandevärdering eftersom den senare kan ge information som kan ha betydelse för tolkningen av utfallsutvärderingens resultat.

Tillförlitliga och användbara utvärderingar förutsätter en noggrann insamling av pålitliga data avseende den studerade åtgärdens utfall samt en adekvat dokumentation av dess genomförande. Utan en välgjord datainsamling är det många gånger helt omöjligt att utföra en utvärdering. I ITPS sekundäranalys av strukturfonderna visade det sig till exempel att utvärderbarheten avsevärt försvårades av begränsningar i datainsamlingen. Rent generellt handlar en adekvat datainsamling om att de åtgärder som implementeras kan hänföras till individer, företag eller kommuner, så att utfallet kan studeras både före och efter åtgärdens implementering. För i synnerhet effekt- och effektivitetsutvärderingar är det dessutom nödvändigt att ha tillgång till data som gör det möjligt att studera det kontrafaktiska tillståndet, det vill säga hur individer, företag eller kommuner utvecklas över tiden om de inte exponeras av åtgärden.

Ett första steg i en utvärdering är att fastställa utvärderbarheten av den åtgärd, det program eller det projekt det gäller. Det är i det skedet oerhört viktigt med en god kommunikation mellan utvärderare och uppdragsgivare så att utvärderingen kommer att svara mot uppdragsgivarens förväntningar. Ett sätt att minska risken för svårutvärderade åtgärder är att utvärderaren kopplas in tidigt i processen för att exempelvis garantera datainsamling och dokumentation och därmed även utvärderbarheten.

Att utvärdera den regionala tillväxtpolitiken innebär speciella utmaningar och svårigheter. Utvärderingar av åtgärder inom området bör ta särskild hänsyn till den regionala dimensionen och heterogeniteten. Det handlar bland annat om att ta hänsyn till sådant som arbetspendling mellan regioner, regioners påverkan på andra regioner, regionernas olika storlek, samt regioners olikhet angående organisation och förutsättningar i olika avseenden. Att beakta en mångfald av sådana aspekter är en stor utmaning och kräver att utvärderare har kontakter med såväl utvärderingsforskning som forskning rörande regional utveckling samt har en ständig bevakning av nya metoder och angreppssätt. Två metoder

som tas upp i den här rapporten för att beskriva den regionala heterogeniteten är indikatorer och flernivåanalys.

Indikatorer är ett vanligt förekommande verktyg för utvärderingar med så kallade utvärderingsindikatorer, och för att beskriva den regionala utvecklingen med så kallade uppföljningsindikatorer. En uppföljningsindikator visar på *att* något har hänt, men ger vanligtvis inte svar på *varför* detta har hänt. Uppföljningsindikatorer ska därför användas med en viss försiktighet och ska användas främst för att ge en snabb bild över ett tillstånd. En utvärderingsindikator, å andra sidan, har en högre ambitionsnivå i att indikatorn ska kunna visa på att åtgärden har orsakat en måluppfyllelse. Utvärderingsindikatorer kan vara speciellt användbara i monitoring sammanhang. Nackdelen är dock att de kan vara svåra att utforma så att åtgärdens effekt verkligen särskils från övriga faktors inverkan.

En användbar metod för att ta hänsyn till den regionala heterogeniteten är flernivåanalys. Flernivåanalys är en effektiv metod för att kontrollera för dels om regioner har en viss fast effekt av att vara en viss region och att kontrollera för att effekten av en viss variabel är olika i olika regioner. Flernivåanalys är speciellt användbart när data finns för flera olika nivåer, till exempel kommuner och län. Flernivåanalys lämpar sig således väl för olika typer av analyser där det finns anledning att ta hänsyn till att regioners specifika karaktäristika kan påverka analysen.

En viktig aspekt i utvärderingssammanhang, är att kvalitetssäkra arbetet med utvärderingar. En metautvärdering är en utvärdering av andra utvärderingar. Syftet är att bedöma gjorda utvärderingars kvalitet och användbarhet för att på så sätt skaffa kunskaper om hur utvärderingar av projektet kan förbättras. ITPS anser att metautvärdering är en lämplig och användbar metod för att kvalitetssäkra arbetet med utvärderingar.

Att utvärdera komplexa program, med många olika aktörer och som består av ett antal projekt eller delprogram, kan vara speciellt utmanande. Att utvärdera komplexa program i sin helhet kan visa sig vara väldigt svårt. ITPS identifierar tre lämpliga angreppssätt för att utvärdera sådana program. Det första innebär att utföra en övergripande målutvärdering av hela programmet. En sådan ger dock inte svar om eventuella effekter. Det kan emellertid ofta vara möjligt att göra effektutvärderingar av de enskilda projekten eller delprogrammen. Ett andra angreppssätt är därför att utföra en sammansatt utvärdering baserad på resultaten från såväl utvärdering på en övergripande nivå som från utvärderingar avseende olika delar av programmet. Ett tredje alternativ är att använda programteori som, kort uttryckt, innebär att man beaktar det förhållandet att ett programs kontext påverkar förutsättningarna för programmets olika aktörer vilket i sin tur påverkar hur programmet fungerar och därmed även dess resultat.

Det har blivit allt vanligare med program som kännetecknas av en betydande samverkan mellan olika aktörer. Det är därför viktigt för utvärderare att ha kunskap om olika angreppssätt för att utvärdera samverkan. Tidigare utvärderingar av horisontell samverkan har visat att det inte finns någon universallösning när det gäller tillvägagångssätt utan att det i stället är viktigt att kombinera flera olika metoder. Ett fruktbart angreppssätt kan då vara att kombinera mera traditionella genomförande- och målutvärderingar med utvärderingar baserade på modern programteori.

I den här rapporten har olika metoder och angreppssätt för att utvärdera den regionala tillväxtpolitiken beskrivits och diskuterats. Att på förhand förutse alla de utmaningar och problem som utvärderare av regional tillväxtpolitik i praktiken kan ställas inför är naturligtvis inte möjligt. Syftet med rapporten har därför inte varit att åstadkomma en manual

för utvärderingar utan att presentera grundläggande synsätt som kan tjäna som utgångspunkter när det gäller att ta ställning till hur man på bästa sätt kan genomföra en systematiskt genomförd bedömning av en regional tillväxtpolitisk åtgärd. Det bästa sättet att slutligen skaffa sig och utveckla kunskaper om utvärderingar, är att utföra och medverka i sådana och att vara öppen för att de kvalitetsgranskas.

Referenser

- 6, P. (2004). "Joined-Up Government in the Western World in Comparative Perspective: A Preliminary Literature Review and Exploration", *Journal of Public Administration Research and Theory*, 14: 103–138.
- Agranoff, R. och McGuire M. (2003). *Collaborative Public Management – New Strategies for Local Governments*. Washington D.C.: Georgetown University Press.
- Bamberger M. J., Rush, J. och Mabry, L. (2006). *Real World Evaluation: Working Under Budget, Time, Data, and Political Constraints*. Thousand Oaks: Sage.
- Bickel, R. (2007). *Multilevel Analysis for Applied Research*, New York: The Guildford Press.
- Bryman, A. (2004). *Social research methods*, andra upplagan. Oxford: Oxford University Press.
- Chen, H. och Rossi, P. (1989). Issues in the Theory-driven Perspective, *Evaluation and Program Planning*, 12:299–306.
- Ds (2000:63). *Ekonomisk styrning för effektivitet och transparens*. Stockholm: Finansdepartementet.
- Engwall-Skeri, L. Nutek, personlig kommunikation, 2008-10-28.
- European Commission. (2006). The New Programming Period 2007–2013. Indicative guidelines on evaluation methods: Evaluation during the programming period. Working Document No. 5. Bryssel: European Commission, Directorate General Regional Policy, Directorate General Employment, Social Affairs and Equal Opportunities.
- European Commission, (2008). http://ec.europa.eu/regional_policy/sources/docgener/evaluation/evalsed/index_en.htm, 2008-03-04.
- Environment Canada, (2008). http://www.ec.gc.ca/soer-ree/English/scip/ind_selection.cfm, 2008-11-18.
- Finansdepartementet (2003). Regeringsbeslut FI 2003/3123. Stockholm.
- Finansdepartementet (2008). Budgetpropositionen för 2009, Prop. 2008/09:1. Utgiftsområde 19 Regional tillväxt. <http://www.regeringen.se/sb/d/10942/a/111679>
- Forss K. (2007). *Utvärdering som handverk – Bortom mallar och manualer*. Lund: Studentlitteratur.
- Gadd, H., Hansson, G., och Månsson, J. (2008). "The Relation between Firm Subsidy and Success – Does Regional Context Matters?" ITPS mimeo.
- Gaffey, V. (2008). "The European Regional Development Fund or How to evaluate the billions?" Paper presented at the EES Conference, Lisbon 1–3 October 2008.
- Gelman, A., och Hill. J. (2007) *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge: Cambridge University Press.

- Goldmann K., Pedersen M. N., och Østerud. Ø. (1997). *Statsvetenskapligt Lexikon*. Stockholm: Universitetsförlaget.
- Greene, W. H. (2003). *Econometric Analysis – Fifth edition*. Upper Saddle River, N.J.: Prentice Hall.
- Harrinvirta M. och Kekkonen, S. (2004). Evaluating Effectiveness of the Horizontal Policy Programs of the Finnish Government, EGPA Study Group on Productivity and Quality in the Public Sector, Ljubljana 1–4 September 2004. Prime Minister's Office, Finland.
- IFAU (2000). Dataproblem vid utvärdering av arbetsmarknadsproblem. Stencilserie 2000:5.
- ITPS (2003a). Ex ante-bedömning av de regionala tillväxtprogrammen – ITPS sammanfattade rapport till Näringsdepartementet, Regleringsbrevsuppdrag nr 5, 2003.
- ITPS (2003b). Sysselsättningsbidrag – utvärdering av regional politiskt stöd, A2003:006, Östersund.
- ITPS (2003c). 500 miljoner för regional sysselsättning och tillväxt – Utvärdering av programmet för regional näringspolitik och särskilda regionalpolitiska frågor. A2003:007, Östersund.
- ITPS (2003d). Regionala tillväxtindikatorer – teoretiska aspekter, begrepp och empiriska illustrationer. A2003:004. Östersund
- ITPS (2004). Omställningsarbete efter försvarsnedläggningar – Utvärdering av resultat av genomförda utvecklingsprogram i nio kommuner, Dnr: N2002/409/AE och N2002/6954/AE.
- ITPS (2005). Utvärdering av expertskatten, Regleringsbrevsuppdrag nr. 4, 2005. Östersund.
- ITPS (2007a). Förmånsrättsreformen, A2007:014, Östersund.
- ITPS (2007b). Effektutvärdering av det regionala utvecklingsbidraget: en studie av effekter på svenska aktiebolag. A2007:016. Östersund.
- ITPS (2007c). Utvärdera effekter av offentligt stöd till näringslivet – Förslag till analysverktyg. A2007:017. Östersund.
- ITPS (2008a). Indikatorer för uppföljning av regional utveckling, Regleringsbrevsuppdrag nr 6, 2008, Dnr 1-010-2008/0018.
- ITPS (2008b). Hållbar regional tillväxt – en kunskapsöversikt, Diarienummer 1-010-2008/0018.
- ITPS (2008c). Politisk styrning och regional utveckling – En litteraturoversikt, R2008:011. Östersund.
- ITPS (2008d). Policyspaning, regionala trender och utvärdering i EU – Erfarenheter från ESPON, Dnr.: 1-010-2008/0018.
- ITPS (2008e). Metautvärdering av strukturfondsprogrammen – nya ansatser och lärdomar, A2008:015. Östersund.

- ITPS (2008f). Regionernas tillstånd 2007 – En rapport om tillväxtens förutsättningar i svenska regioner. A2008:001. Östersund.
- ITPS (2008g). Tillväxt i relation till regionalt självstyre och regional storlek. R2008:004. Östersund.
- ITPS (2009). Hållbar regional tillväxt – en kunskapsöversikt om ekonomisk tillväxt, kluster och storstädernas betydelse, R2009:003, Östersund.
- Jobin, D. (2008). "A Transaction Cost-Based Approach to Partnership Performance Evaluation," *Evaluation*, 14: 437-465.
- Kekkonen, S. (2006). "Horizontal Government Policy Programmes and Programme Management Reform in Finland", i: *Finnish Public Management Building Sustainable Quality*. Helsingfors: Ministry of Finance.
- Kooiman, J. (2003). *Governing as Governance*. London: Sage.
- Layard, R. och Glaister, S. (eds). (1994). *Cost-Benefit Analysis*. Cambridge: Cambridge University Press.
- Langert Zetterman, M och Strömdahl, H. (2001). Teoriförankrad utvärdering – ett integrerat perspektiv. Uppsats publicerad i Kvalitetsbanken vid Kvalitetsmässan 2001, Göteborg.
- Mayne, J. (2001). "Addressing attribution through contribution analysis: using performance measure sensibly", *The Canadian Journal of Program Evaluation*, 16(1):1-24.
- Michels, A. och Meijer A. (2008). "Safeguarding public accountability in horizontal government", *Public Management Review*, 10: 165-173.
- Mickwitz P. och Kivimaa P. (2007). "Evaluating Policy Integration: The Case of Policies for Environmental Friendlier Technological Innovations", *Evaluation*, 13: 68-86.
- Nationalencyklopedin (2008). <http://www.ne.se/>
- Nutek (2006) Indikatorer för hållbar regional utveckling. PM 2006-09-11. Handlingsgruppen för hållbar utveckling. <http://www.nutek.se/content/1/c4/41/60/Indikatorerforhllbarregionalutvslutlig060911.pdf>
- Nutek (2007). Regioner i arbete för tillväxt 2007, En lägesrapport från det regionala tillväxtarbetet. Nutek
- Näringsdepartementet (2006). En nationell strategi för regional konkurrenskraft, entreprenörskap och sysselsättning 2007–2013 – Analysunderlag, Stockholm: Näringsdepartementet.
- Näringsdepartementet (2007). En nationell strategi för regional konkurrenskraft, entreprenörskap och sysselsättning 2007–2013. Stockholm: Näringsdepartementet.
- Pawson, R. och Tilly, N. (1997). *Realistic Evaluation*. London: Sage.
- Peters, B. G. (1998). "Managing Horizontal Government: The Politics of Co-ordination", *Public Administration*, 76: 295–311.
- Pierre, J. och Peters B. G. (2000). *Governance, Politics and the State*. Houndmills: Macmillan.

- Porter, M. E. (1990). *The competitive advantage of nations*, London: The Macmillan Press Ltd.
- Quist, J. (2005). *Sammanhållna processer – en studie om horisontell samverkan*. CTF.
- Quist, J. (2007). *Den vertikala dimensionen av en sammanhållen förvaltning*. CTF.
- Rogers, P. J. (2008). "Using Programme Theory to Evaluate Complicated and Complex Aspects of Interventions", *Evaluation*. 14: 29–48.
- Rosenbaum, P.R., och Rubin, D.B. (1983). "The Central Role of the Propensity Score in Observational Studies for Causal Effects", *Biometrika*, 70: 41–55.
- Scriven, M. S. (1969). "An introduction to meta-evaluation", *Educational Products Report*, 2: 36–38.
- Stame, N. (2004). "Theory-Based Evaluation and Types of Complexity", *Evaluation*. 10: 58–76.
- Stame, N. (2006). "Governance, Democracy and Evaluation", *Evaluation*. 12: 7–16.
- Statskontoret (2001). *Utvärdera för bättre beslut – Att beställa utvärderingar som är till nytta i beslutsfattandet*. 2001:22.
- Statskontoret (2005). *Mot en modern och sammanhållen förvaltning - Erfarenheter från sex myndigheters samverkan i 15 projekt*. Stockholm.
- Statskontoret (2007). *Joining-up for regional development*. Stockholm.
- Statsrådets kansli (2004). *Regeringens strategidokument 2004*. Helsingfors.
- Statsrådets kansli (2007). *Regeringens strategidokument 2007*. Helsingfors.
- Stufflebeam D. (1974). Meta-evaluation, Paper #3 Occasional Paper Series.
- Stufflebeam D. L. (2001). "The Metaevaluation Imperative", *American Journal of Evaluation*, 22: 183–209.
- Stufflebeam, D. L. (2001b) "Evaluation Models", *New Directions for Evaluation*, 89: 7–98.
- Tuynman, M., Gilsing. R., och van Houten. G. (2008) Centralised Evaluation of Decentralised Policy Programme: Methodological Choices. Paper presented at the EES Conference, Lisbon 1–3 October 2008.
- Twisk, Jos W.R. (2006) *Applied Multilevel Analysis*. Cambridge: Cambridge University Press.
- van der Meer, F.-B. och Edelenbos, J. (2006). "Evaluation in Multi-Actor Policy Processes: Accountability, Learning and Co-operation", *Evaluation*. 12: 201–218.
- Vedung, E. (1998). *Utvärdering i politik och förvaltning*. Lund: Studentlitteratur
- Weiss, C. (1987). "Where Politics and Evaluation Research Meet", I Palumbo. D (red.) *The Politics of Program Evaluation*. Newbury Park, CA: Sage
- Wolf F. M. (1986). *Meta-Analysis – Quantitative Methods for Research Synthesis, Quantitative Applications in the Social Sciences*. Newbury Park: Sage.

Bilaga 1: Regleringsbrevsuppdraget

Denna rapport utgör slutrapport för ITPS regleringsbrevsuppdrag nummer 6, 2008. Regleringsbrevsuppdraget är formulerat enligt följande:

6. Metodutveckling för utvärdering av regional tillväxtpolitik

ITPS ska utveckla metoder för att utvärdera den regionala tillväxtpolitiken utifrån ett ekonomiskt, miljömässigt och socialt perspektiv, med fokus på effekter och måluppfyllelse. Arbetet ska genomföras i dialog med Regeringskansliet (Näringsdepartementet). ITPS ska i relevanta delar av uppdraget samråda med Naturvårdsverket och Glesbygdsverket. En delrapport ska redovisas senast den 15 september 2008 och en slutrapport senast den 28 februari 2009.

Inom ramen för detta metodutvecklingsarbete ska ITPS följa upp och uppdatera de övergripande indikatorer som används i budgetpropositionen för 2008, utgiftsområde 19 Regional utveckling, exklusive servicetillgänglighet, och påbörja ett arbete med att utveckla miljöindikatorer. Denna del av uppdraget ska redovisas till Regeringskansliet (Näringsdepartementet) senast den 30 april 2008.

ITPS ska särskilt redovisa en sammanställning av aktuell forskning, inom olika discipliner, av betydelse för hållbar regional tillväxt samt forskning om politikutveckling inom området. Arbetet ska ta tillvara den kunskap som genererats genom ESPON. Denna del av uppdraget ska redovisas till Regeringskansliet (Näringsdepartementet) senast den 15 september 2008.

Bilaga 2: Utvärderingar utgivna av ITPS

Ex-ante

ITPS (2003). Ex-ante bedömning av de regionala tillväxtprogrammen, Regleringsbrevsuppdrag nr. 5, 2003, Östersund.

ITPS (2005). Samhällsekonomisk analys av effekterna av liberalisering av postmarknaden – Underlag och överväganden för Post- och kassaserviceutredningen. A2005:013, Östersund.

ITPS (2007). Förhandsbedömning av de operativa regionala strukturfondsprogrammen, A2007:003, Östersund.

Mid-term/on-going

Genomförandevärdering

ITPS (2003). Bredbandspolitiken – en utvärdering i halvtid: Delrapport till ITPS utvärdering av den svenska IT-politiken. A2003:015, Östersund.

ITPS (2006). Halvtidsutvärdering av de regionala tillväxtprogrammen, A2006:013, Östersund.

ITPS (2008). Metautvärdering av strukturfonderna – nya ansatser och lärdomar, A2008:015, Östersund.

Målvärdering

ITPS (2004). Omställningsarbete efter försvarsnedläggningar – Utvärdering av resultat av genomförda utvecklingsprogram i nio kommuner, Dnr: N2002/409/AE och N2002/6954/AE.

Ex-post

Målvärdering

ITPS (2003). 500 miljoner för regional sysselsättning och tillväxt – Utvärdering av programmet för regional näringspolitik och särskilda regionalpolitiska frågor. A2003:007, Östersund.

ITPS (2003). En lärande IT-politik för tillväxt och välfärd, A2003:015, Östersund.

ITPS (2005). Utvärdering av expertskatten, Regleringsbrevsuppdrag nr. 4, 2005. Östersund.

ITPS (2005). Utvärdering av delegationen för regional samverkan om högre utbildning, A2005:014, Östersund.⁹⁵

⁹⁵ Med inslag av att vara en genomförandevärdering.

ITPS (2006). Tack för oss! En utvärdering av Inlands- Tillväxtdelegationernas verksamhet, A2006:001, Östersund.

ITPS (2007). Förmånsrättsreformen, A2007:014, Östersund.

ITPS (2008). Metautvärdering av strukturfonderna – nya ansatser och lärdomar, A2008:015, Östersund.

Effektutvärdering

ITPS (2003). Sysselsättningsbidrag – utvärdering av regional politiskt stöd, A2003:006, Östersund.

ITPS (2003). Växtkraft Mål 4 – en effektstudie, A2003:001, Östersund.

ITPS (2003). Utvärdering av de näringspolitiska effekterna av stöd till start av näringsverksamhet, A2003:003, Östersund.

ITPS (2004). The EC Regional Structural fund impact in Sweden 1995-1999, A2004:028 & A2004:009, Östersund

ITPS (2004). Effektutvärdering av de geografiska målprogrammen inom EG:s strukturfonder, A2004:009, Östersund.

ITPS (2007). Effektutvärdering av det regionala utvecklingsbidraget, A2007:016, Östersund.

ITPS, Institutet för tillväxtpolitiska studier, är en statlig myndighet med uppdrag att bidra med kunskapsunderlag för tillväxtpolitiken.

ITPS har definierat fyra tillväxtpolitiska utmaningar för Sverige:

- Globalisering och internationalisering
- Regional tillväxt
- Näringslivets dynamik
- Teknisk utveckling

Verksamhet och arbetsuppgifter är fördelade inom tre avdelningar och tre kunskapsområden:

- Avdelningen för tillväxtanalys och statistik beskriver och analyserar tillväxten och dess drivkrafter i Sverige.
- Avdelningen för utvärdering redovisar resultatet av den politik som genomförs.
- Avdelningen för omvärldsanalys bevakar och analyserar händelser i omvärlden som kan komma att få betydelse i Sverige.

Uppdragsgivare är regeringen. I uppdraget ingår att sprida kunskapen till nationella, regionala och lokala tillväxtpolitiska aktörer.

ITPS har sitt huvudkontor i Östersund och verksamhet i Stockholm, Peking, Tokyo, New Delhi, San Francisco, Washington och Bryssel.