



MISSIV

Datum
2000-04-04
Ert datum

Diariernr
2000/163-5
Er beteckning

Regeringen
Justitiedepartementet

103 33 Stockholm

Uppdrag att genomföra en kartläggning och bedömning av det samlade utvärderingsarbetet inom de olika utgiftsområdena.

I regleringsbrevet för år 1999 uppdrog regeringen åt Statskontoret att *göra en översiktlig kartläggning och bedömning av det samlade utvärderingsarbete som bedrivs inom de olika utgiftsområdena.*

Statskontoret avrapporterar härmed uppdraget genom rapporten *Utvärdering och politik - en kartläggning av utvärderingarna i budgetpropositionen (2000:17)*. I kartläggningen ingår samtliga utredningar, rapporter m.m. som benämndes "utvärderingar" i budgetpropositionen för år 2000. Dessa utvärderingar har analyserats bl.a. vad gäller omfattning, kvalitet och hur regeringen använt dem i budgetpropositionen.

Tillförordnade generaldirektören Svante Holgersson har beslutat i detta ärende. Direktör Lars Dahlberg, avdelningsdirektör Magnus Medelberg, föredragande, och förste byråsekreterare Mathias Fredriksson var närvarande vid den slutgiltiga handläggningen.

Enligt Statskontorets beslut

Magnus Medelberg

STATSKONTORET

Postadress
Box 2280
103 17 STOCKHOLM
Besök Norra Riddarholmshamnen 1

Telefon (växel)
08 - 454 46 00
Telefon (direkt)
08 - 454

Fax (registrator)
08 - 791 89 72
Fax (direkt)

Internetadresser
www.statskontoret.se
statskontoret@statskontoret.se

Sammanfattning och förslag	7
<i>Uppdrag, syfte och metod</i>	7
<i>Resultat</i>	7
Utvärderingarnas fördelning på utgiftsområden	7
Beställare och producenter	9
Utvärderingarna i beslutsfattandet	10
Utvärderingarnas kvalitet	12
<i>Statskontorets förslag</i>	15
1. Inledning	18
1.1 <i>Uppdraget</i>	18
1.2 <i>Bakgrund och syfte</i>	18
1.3 <i>Metod</i>	19
1.4 <i>Disposition</i>	21
2. Utvärderingarnas fördelning på utgiftsområden	22
3 Beställare och producenter	32
3.1 Beställarna	32
3.2 Producenterna	34
4. Utvärderingarna i beslutsfattandet	40
5 Utvärderingarnas kvalitet	44
5.1 Hur utvärderingarna presenteras	45
5.2 Två betydelsefulla metodfaktorer	46
Bilaga, tabeller	54

Sammanfattning och förslag

Uppdrag, syfte och metod

Detta är en avrapportering av uppdraget i Statskontorets regeringsbrev för år 1999 att *Statskontoret skall göra en översiktlig kartläggning och bedömning av det samlade utvärderingsarbete som bedrivs inom de olika utgiftsområdena*. Kartläggningens syfte är att få en överblick över det samlade utbudet av utvärderingar som regeringen använder oavsett vilka organisationer som gjort utvärderingarna.

I kartläggningen ingår samtliga utredningar, rapporter m.m, som benämndes ”utvärderingar” i budgetpropositionen för år 2000. Dessa utvärderingar har analyserats bl.a. vad gäller omfattning, kvalitet och hur regeringen använt dem i budgetpropositionen.

Resultat

Utvärderingarnas fördelning på utgiftsområden

Ojämn fördelning på utgiftsområden

De 114 utvärderingar vi funnit i budgetpropositionen fördelar sig mycket ojämnt på de olika utgiftsområdena.

Ett utgiftsområde står i en klass för sig vad det gäller antal utvärderingar. Det är *Utbildning och universitetsforskning* som har hela 30 utvärderingar nämnda. Därefter kommer fem utgiftsområden med mellan tretton och nio utvärderingar nämnda. Dessa är *Hälsovård, sjukvård och social omsorg, Kultur, medier, trossamfund och fritid, Arbetsmarknad och arbetsliv, Regional utjämning och utveckling* samt *Samhällsekonomi och finansförvaltning*.

Det finns hela 15 utgiftsområden som endast har en eller ingen utvärdering nämnd. I många fall kan det finnas godtagbara förklaringar till att ett utgiftsområde har få utvärderingar:

- det kan vara litet utgiftsmässigt sett
- det kan ha en ensartad verksamhet som inte behöver utvärderas varje år
- det kan råda politisk enighet om att området fungerar bra

Efter att vi beaktat dessa förklaringar kvarstår några utgiftsområden (och ett verksamhetsområde) som förefaller ha mycket få utvärderingar i förhållande till utvärderingsbehovet:

- *Totalförsvaret* som både har höga utgifter och en mycket komplex verksamhet
- De fyra utgiftsområden som handlar om *Ekonomisk trygghet* som har höga utgifter¹
- *Utrikesförvaltning och internationell samverkan* som har varit i fokus för den politiska debatten under senare år
- *Allmän miljö- och naturvård* som innehåller frågor som ofta framhålls som centrala för vår framtida välfärd
- Dessutom finner vi det märkligt att det i *Arbetsmarknad och arbetsliv* inte finns någon utvärdering under verksamhetsområdet *Arbetsliv*

Få sektorsövergripande utvärderingar

Vi har vidare funnit att antalet sektorsövergripande utvärderingar är mycket lågt. Detta trots att många effektivitetsproblem inom den offentliga sektorn är just av sektorsövergripande karaktär och att det därför borde finnas ett behov av sådana utvärderingar.

¹ Ekonomisk trygghet vid sjukdom och handikapp, Ekonomisk trygghet vid ålderdom, Ekonomisk trygghet för familjer och barn, Ekonomisk trygghet vid arbetslöshet.

Beställare och producenter

Inga beställningar från "Försvaret" och "Miljö"

Av de utvärderingar som omnämns i budgetpropositionen är 58 procent beställda av regeringen. De flesta av de övriga har initierats av myndigheter. I några fall är det något EU-organ som beställt utvärderingen.

En analys av vilka departement som är de flitigaste beställarna överensstämmer med vad man kan vänta sig efter analysen av under vilka utgiftsområden utvärderingarna presenteras. Utbildningsdepartementet är således den flitigaste beställaren, medan Försvarsdepartementet och Miljödepartementet inte har beställt någon av de utvärderingar som finns redovisad i budgetpropositionen.

Få utvärderingar görs av utvärderingsorganisationer

En intressant iakttagelse när det gäller *producenterna* är att de organisationer inom statsförvaltningen som har utvärdering som en central uppgift svarar för en ganska liten del av utvärderingarna. Vi tänker då på *kommittéväsendet*, de *allmänna utvärderingsmyndigheterna*, Statskontoret och RRV, och *de sektoriella utvärderingsmyndigheterna* som Brottsförebyggande rådet (BRÅ), Statens institut för kommunikationsanalys (SIKA), Statens institut för regionalforskning (SIR) och Institutet för arbetsmarknadspolitisk utvärdering (IFAU). Dessa tre organisationstyper svarar bara för 18 procent av utvärderingarna.

Istället är det kategorierna *Fem stora sektorsmyndigheter* och *Övriga myndigheter*, som inte har utvärderingar som en lika framträdande uppgift, som är de största utvärderingsproducenterna. De fem stora sektorsmyndigheterna; Socialstyrelsen, RFV, NUTEK, Skolverket och Högskoleverket, svarar för 26 procent av alla utvärderingar.² Den kategori som producerat flest utvärderingar är *Övriga myndigheter*, där vi även räknat in statliga stiftelser och bolag. Denna kategori har producerat 29 procent av utvärderingarna.

² Dessa fem stora myndigheter är av olika karaktär med mer eller mindre starka inslag av utvärdering i sina uppgifter.

Tillsammans producerar således *Stora sektorsmyndigheter* och *Övriga myndigheter* mer än tre gånger så många utvärderingar som de tidigare nämnda utvärderingsorganisationerna.

En avvägning mellan sakkunskap och oberoende

En central fråga för departementen vid valet av utvärderingsorganisation är avvägningen mellan *sakkunskap* och *oberoende*. Att göra utvärderingen på departementet eller att lägga ut den på den av departementets myndigheter som har hand om själva saksfrågan ger en hög sakkunskap i utvärderingen. Samtidigt finns risken att utvärderingen inte kommer att betraktas som oberoende.

Motsatsen gäller när utvärderingarna istället läggs ut på från departementet fristående organisationer. Då minskar risken att utvärderingarna kommer att betraktas med skepsis ur objektivitetssynpunkt men å andra sidan ökar risken att sakkunskapen ifrågasätts.

Det finns ingen generell lösning på hur denna avvägning bör hanteras. Däremot är det viktigt att departementen behandlar frågan i sina utvärderingsstrategier.

Utvärderingarna i beslutsfattandet

Utvärderingarna leder sällan till ändrad politik

För att få en uppfattning om hur regeringen använder utvärderingarna har vi analyserat hur de kommenteras i budgetpropositionen.

I 42 procent av fallen *kommenterar* regeringen bara att utvärderingen har genomförts utan några värderingar. I 26 procent av fallen används utvärderingen och dess resultat för att *bekräfta* att den politik som bedrivs fungerar bra. I 19 procent av fallen lovar regeringen att man ska *återkomma* senare med något slags ställningstagande. Här ingår även de fall då regeringen avvaktar ett remissförfarande. Endast 12 procent av utvärderingarna i budget-

propositionen nämns i samband med att regeringen fattat något beslut om en *förändring* av politiken.

Förklaringarna till detta kan vara flera:

- budgetpropositionen har delvis karaktären av resultatredovisning, d.v.s. här redovisas vad myndigheterna har åstadkommit under året. Det betyder att många utvärderingar omnämns snarare som exempel på myndighetsresultat än för att de har relevans för regeringens beslutsfattande. Detta kan förklara den höga andel utvärderingar där regeringen nöjer sig med allmänna kommentarer.
- utvärderingar utgör ofta en del i ett långsiktigt kunskapsuppbyggande snarare än ett direkt beslutsunderlag. Mot detta talar att vi endast funnit enstaka exempel på att regeringen i budgetpropositionen nämner utvärderingar som är mer än några år gamla.
- de utvärderingar som ingår i denna kartläggning kan leda till beslut som fattas i annat sammanhang än i budgetpropositionen. I ungefär en femtedel av de fall vi analyserat lovar regeringen att återkomma till frågan, eller att tillsätta en ny utredning.
- I ungefär en fjärdedel av de fall vi analyserat använder regeringen utvärderingens resultat för att bekräfta att den förda politiken fungerar bra. Vill man vara positiv kan även dessa fall tolkas som att regeringen använt utvärderingen i beslutsfattandet. Om utvärderingens resultat varit annorlunda hade den kanske lett till ett beslut om en förändring av politiken...

Men det kan också vara så att regeringens användning av utvärderingarna inte är större än vad våra siffror indikerar. Det kan i såfall bero på att utvärderingarna håller för låg kvalitet. Men det kan också bero på att regeringens beställningar inte är tillräckligt genomtänkta så att utvärderingen svarar på andra frågor än de som är relevanta för beslutsfattandet. Vår översiktliga kartläggning är inte tillräcklig för att kunna säga något om varför så få utvärderingar används som beslutsunderlag, eller ens om detta verkligen är fallet. Det behövs mer analys kring var i kedjan, beställning – utvärdering – mottagande – beslut, som bristerna finns.

Utvärderingarnas kvalitet

Många utvärderingar otydligt presenterade

Kvaliteten i hur utvärderingarna presenteras i skrift är mycket skiftande. Det finns goda exempel där det är lätt att följa hur utvärderingen gått till och varför man drar de slutsatser man drar. Men det finns samtidigt alltför många exempel där det som läsare är mycket svårt att se kopplingen mellan de slutsatser som dras och utvärderingens resultat eller vilken metod som använts för att få fram resultaten. Det gör att utvärderingarnas användbarhet som beslutsunderlag försämras.

Två tredjedelar av utvärderingarna har en otydlig måttstock.

Med måttstock avses det ”nollalternativ” som den utvärderade verksamheten eller åtgärden jämförs med. Vår kartläggning visar att nästan två tredjedelar av utvärderingarna inte har någon tydlig måttstock. En tydlig måttstock underlättar bedömningen av om den utvärderade verksamheten är bra eller dålig. Saknas däremot en tydlig måttstock, d.v.s. något att jämföra med, blir denna bedömning mycket svårare.

De flesta utvärderingar som innehåller tydliga måttstockar har gjort jämförelser med antingen tiden före den utvärderade åtgärden infördes eller jämfört med någon ”kontrollgrupp” som inte omfattats av åtgärden.³

I den majoritet av utvärderingar som har otydliga måttstockar lyser oftast före-efter jämförelser eller jämförelser med kontrollgrupper med sin frånvaro. Ofta förekommande exempel på otydliga måttstockar är utvärderingar som grundas på attitydundersökningar, gjorda vid ett enda tillfälle och utan kontrollgrupper, där respondenterna visserligen fått uttala sig om den utvärderade verksamheten fungerar ”bra” eller ”dåligt” men där det är oklart vad de som svarat egentligen jämfört med.

³ Om utvärderingen inte gäller en specifik åtgärd utan snarare en organisations löpande verksamhet passar inte begreppen ”före-efter” och ”kontrollgrupp”. Motsvarande måttstockar finns då i form av ”tidsserier” och jämförelser med ”liknande verksamheter”.

Det är också vanligt med utvärderingar som bygger på att "utomstående bedömare" gör en, mer eller mindre ingående, granskning av en verksamhet ofta genom att läsa den dokumentation som finns om verksamheten samt samtala med personal och en del avnämare. Detta leder till kvalitativa värdebedömningar som t.ex. att "verksamheten fungerar bra" men det är oklart vad de har jämfört med när de gjort bedömningen.

I båda dessa fall är det troligt att det finns måttstockar (i respondenternas eller bedömarnas huvuden) men det är mycket svårt för den som läser utvärderingen att avgöra vad bedömningen grundar sig på.

Mindre än hälften mäter effekter

Vi har också bedömt utvärderingarna utifrån Finansutskottets önskemål om att den resultatinformation som lämnas i budgetpropositionen i högre grad ska vara inriktad på de statliga insatsernas *effekter på samhällsnivå* och att resultaten i högre grad ska redovisas i kvantitativa termer.

Med en sträng definition av *effekter på samhällsnivå* är det mellan 20 och 25 procent av utvärderingarna som bygger på sådana mått. Med en vidare definition ökar andelen effektmätningar till mellan 40 och 50 procent.

Mellan 35 och 40 procent av utvärderingarna avser inte alls, eller bara till en liten del, effekter på samhällsnivå. Dessa kan klassificeras som kvalitetsstudier där fokus ligger på kvaliteten i myndigheternas prestationer snarare än på vilka effekter dessa prestationer åstadkommer. De flesta utvärderingarna i denna kategori bygger antingen på resultat från attitydundersökningar till kunder eller på utomstående bedömares granskning av verksamheten. I det senare fallet är redovisningen av resultat i kvantitativa termer ofta obetydlig.

Slutligen har vi klassificerat ungefär 20 procent som *genomförandevärderingar*. Dessa bedömer i första hand om verksamheten genomförs enligt uppdragsgivarens riktlinjer och om processen fungerar som det var tänkt. Många av genomförandevärderingarna

värderingarna avser introduktionsfasen av nya åtgärder eller verksamheter. I flera fall tänker man göra effektutvärderingar när verksamheterna kommit så långt att det blir möjligt att avläsa effekter. Genom att kombinera genomförande- och effektutvärderingar ökar möjligheterna att förklara vad resultaten i effektutvärderingen beror på.

Mindre än var femte utvärdering uppfyller båda kriterierna

När vi väger samman bedömningarna i de föregående avsnitten finner vi att mindre än 20 procent av de utvärderingar vi granskat klarar den stränga definitionen på att mäta effekter samtidigt som de innehåller en tydlig måttstock.

Vad är förklaringen till att inte fler utvärderingar klarar båda kriterierna? Till en del handlar säkert förklaringarna om korta tidsramar, knappa resurser eller bristande metodkunskap hos dem som utfört eller beställt utvärderingen. En annan förklaring är att vissa politikområden till sin karaktär är sådana att det är mycket svårt att finna bra effektmått.

Vi är dock övertygade om att det går att göra betydligt fler högkvalitativa utvärderingar av effekter på samhällsnivå än det görs idag. Dessutom kan kvaliteten i de utvärderingar som inte avser effekter höjas betydligt. Vi vill här lyfta fram två nyckelfaktorer som på ett väsentligt sätt skulle öka möjligheterna att åstadkomma utvärderingar som ger regeringen ett bättre beslutsunderlag.

Tänk efter före

Den ena nyckelfaktorn handlar om att tänka i utvärderingstermer redan i samband med att nya åtgärder eller verksamheter planeras. För att säkra tydliga måttstockar och relevanta resultatmått krävs ofta att utvärderingarna planeras redan i samband med att den åtgärd som ska utvärderas introduceras. För att kunna göra en bra före - efter studie behöver man säkra data om tillståndet före åtgärden kommit igång. För att kunna göra en bra jämförelse med en kontrollgrupp krävs ofta att man, i detta tidiga skede, utser försöksgrupper och kontrollgrupper. Sätter man däremot

igång åtgärden utan att tänka på detta blir det betydligt svårare att göra en bra utvärdering när detta blir aktuellt.

Bra uppföljning underlättar bra utvärdering

Den allra viktigaste faktorn för att öka möjligheterna att producera utvärderingar av hög kvalitet är att den löpande uppföljningen inom det område som ska utvärderas är anpassad till att göra utvärderingar. Med detta avses bl.a. att det, på ett så smidigt sätt som möjligt, skall gå att koppla ihop kostnader, insatser, prestationer och effekter för den verksamhet eller åtgärd som ska utvärderas. Ett, på detta sätt, flexibelt uppföljningssystem ökar t.ex. möjligheten att göra bra före-efter studier även om man "glömt" att planera för utvärderingen innan åtgärden infördes. I de fall man planerar för utvärderingen i god tid så blir datainsamlingen betydligt enklare och billigare än om den löpande uppföljningen på området inte är anpassad för utvärderingar.

Statskontorets förslag

Departementen behöver utveckla strategier för utvärdering

1. En grundläggande förutsättning för att uppnå en hög kvalitet i utvärderingarna är en bra löpande uppföljning som är anpassad till att göra utvärderingar.
2. För att uppnå hög kvalitet i utvärderingarna är det också viktigt att man planerar för framtida utvärderingar redan i samband med att nya åtgärder, verksamheter, o.s.v. introduceras.
3. Beställningen av utvärderingen måste läggas upp på ett sådant sätt att resultaten av utvärderingen svarar på de frågor som är relevanta för regeringens beslutsfattande.
4. Vid valet av vilka organisationer som ska utföra utvärderingarna är det viktigt att göra en genomtänkt avvägning mellan sakkunskap och oberoende.

5. Vissa grundläggande riktlinjer för hur en utvärdering ska presenteras i skrift är önskvärda så att den i dag mycket skiftande kvaliteten härvidlag minskas. Det måste finnas en tydlighet i såväl vilka resultat utvärderingens slutsatser grundas på som vilken metod som använts för att få fram resultaten.

För att verkligen betona vikten av att utvärderingar kommer in tidigt i beslutsprocessen bör regeringen dessutom överväga att införa ett generellt krav på att framtida utvärderingar finns med i beredningen redan i samband med att nya åtgärder och verksamheter beslutas.

Utvärderingarnas roll i beslutsfattandet behöver stärkas

Vår kartläggning visar att det görs mycket få sektorsövergripande utvärderingar. Det finns anledning för Regeringskansliet att överväga hur man kan förstärka de mekanismer som gör att sektorsövergripande utvärderingar kommer till stånd.

Utvärderingarna borde kunna spela en större roll än idag för att fördela om resurser mellan olika offentliga ändamål. Vi föreslår att regeringen för detta syfte inhämtar erfarenheter från andra länder där utvärderingar på ett mer aktivt sätt används vid beslut om hur de offentliga resurserna ska allokeras.

Återkommande översikter behöver göras

Vi anser vidare att översikter över utvärderingsläget, liknande den vi gjort här, behöver göras då och då. En sådan mer återkommande uppföljning av utvärderingsläget blir en värdemätare på hur departementens arbete med att utveckla strategier för uppföljning och utvärdering lyckas. Med en återkommande uppföljning av utvärderingsläget kan regeringen bl.a. få information om:

- vilka utgifts- och verksamhetsområden som gör alltför få utvärderingar
- hur kvaliteten på utvärderingarna utvecklas totalt sett och om det finns stora kvalitetsskillnader mellan olika utgiftsområden

- vilka utgiftsområden som lyckas utveckla användningen av utvärderingarna så att de blir ett mer slagkraftigt instrument i beslutsfattandet

Vi föreslår att Statskontoret får som uppgift att löpande följa upp utvärderingsläget på det sätt vi beskrivit.

Särskild utredning om utvärderingar som beslutsunderlag

Förutom den löpande uppföljningen behövs vissa fördjupningar. En viktig fråga som vi tvingats lämna obesvarad är varför inte utvärderingarna i högre grad används i regeringens beslutsfattande. Vi föreslår därför att regeringen beslutar om en särskild utredning med uppgift att belysa denna fråga. Vårt förslag är att denna utredning använder två olika angreppssätt som tillsammans bör utgöra ett bra underlag för förslag som ökar regeringens användande av utvärderingar.

- Det ena angreppssättet är att utgå från ett antal av regeringen fattade beslut och undersöka i vilken mån utvärderingar ingått i beslutsunderlaget, samt vilken typ av utvärderingar som "borde" ha ingått i underlaget.
- Det andra angreppssättet är att utgå från ett antal utvärderingar gjorda för några år sedan och undersöka hur dessa har använts, och varför de inte använts i ännu högre grad.

1. Inledning

1.1 Uppdraget

Detta är en avrapportering av uppdraget i Statskontorets regleringsbrev för år 1999 att *Statskontoret skall göra en översiktlig kartläggning och bedömning av det samlade utvärderingsarbete som bedrivs inom de olika utgiftsområdena.*

1.2 Bakgrund och syfte

Bakgrunden till vårt uppdrag är att det inte finns någon översikt över det samlade utvärderingsarbete som bedrivs inom den offentliga sektorn och som regeringen använder som beslutsunderlag eller för kunskapsuppbyggnad.

Det har gjorts flera studier där man t.ex. bedömt kvaliteten i ett urval av utvärderingar eller bedömt hur mycket av "utvärdering" det finns i kommittéväsendets utredningar.⁴ Vår kartläggning är ett första försök att få en överblick över det samlade utbudet av utvärderingar som regeringen använder oavsett vilka organisationer som gjort utvärderingarna.

Syftet med kartläggningen är att svara på frågor som

- hur stor är tillgången till utvärderingar?
- vilka är det som beställer och producerar utvärderingarna?
- används utvärderingarna i beslutsfattandet?
- vilken kvalitet håller utvärderingarna?

⁴ Se t.ex. *Att mäta resultatanalysen. Vem analyserar vad, hur mycket och på vilket sätt?* RRV, Dnr 23-91-1736, 1991, *Kan myndigheter utvärdera sig själva*, Ds 1996:36, samt *Användningen av utvärderingar i svensk statsförvaltning*. Murray, Richard, PM Statskontoret.

Frågorna ska besvaras både ur ett helhetsperspektiv och med en analys av om det finns stora skillnader mellan olika utgiftsområden eller departement.

Vi vill betona att det handlar om en översiktlig och bred kartläggning. I vissa fall blir svaren på de frågor som vi ställt bara indikationer på hur det verkligen ser ut. För att kunna ge mer definitiva svar på frågorna krävs en djupare analys än den vi gjort i denna kartläggning.

1.3 Metod

I detta avsnitt redovisas hur vi gjort avgränsningen av vilka utvärderingar som ingår i kartläggningen. Hur vi sedan bedömt utvärderingarna framgår under respektive kapitel.

Analysen omfattar utvärderingar omnämnda i budgetpropositionen

Eftersom ingen tidigare gjort en översikt över det samlade utvärderingsarbetet och det heller inte finns något register över gjorda utvärderingar blev själva ”datainsamlingen” en viktig metodfråga.

Vi valde att avgränsa *det samlade utvärderingsarbetet* till att gälla samtliga utvärderingar som omnämns i budgetpropositionen för år 2000. Dessa identifierades genom att med hjälp av sökordsfunktionen i Acrobat Reader söka igenom budgetpropositionen efter ordet ”utvärd”. På detta sätt kom vi åt alla former och böjningar av ordet *utvärdering*. Om ordet användes i något sammanhang där man hänvisade till någon *specifik* utredning eller liknande så beställde vi denna skrift och den ingår därmed i vår kartläggning. Däremot har vi inte tagit med utvärderingar som omnämns gruppvis, t.ex. när sökordsfunktionen fastnade på *Statskontoret har till regeringen redovisat ca 40 utrednings- och utvärderingsuppdrag*.

Därefter har vi gjort motsvarande sökning på ordet ”effekt”. I de fall vi då fann hänvisningar till specifika effektanalyser, har även dessa definierats som utvärderingar, beställts in och ingår därmed i vår kartläggning. De ytterligare utvärderingar vi fann på detta

sätt utgör en mindre del, ca 10 procent, av det totala antalet utvärderingar i kartläggningen.⁵

I kartläggningen ingår endast de utvärderingar som var avslutade när budgetpropositionen publicerades. De utvärderingar som i budgetpropositionen omnämns som ”pågående”, ”på väg att tillsättas” eller liknande ingår inte i vår analys även om de hunnit publiceras när detta skrivs.⁶

Fördelar och nackdelar

Vårt val att använda budgetpropositionen som ”ram” för vilka utvärderingar som ingår i kartläggningen bygger på att vi ville finna en metod som

- koncentreras till de utvärderingar som är av intresse för regeringen
- är så ”rättvis” som möjlig mellan olika utgiftsområden
- är enkel att genomföra

Vi är medvetna om att metoden har brister. Budgetpropositionen är inte något fullständigt ”rättvist” dokument där alla utgiftsområden beskrivs med samma djup och samma språkbruk. Såväl skillnader i hur orden ”utvärdering” och ”effekt” används, som skillnader i vilken grad beslutsunderlag kommenteras, kan påverka vår analys.

Vi fann dock inget bättre alternativ. Det som låg närmast till hands var att kontakta *alla* departement och be att få de utvärderingar de använt som beslutsunderlag eller för kunskapsuppbyggnad under det senaste året. Vi bedömde dock att ett sådant tillvägagångssätt snarare riskerade att leda till ännu större skillnader mellan olika områden när det gällde vilket

⁵ Vi testade även några andra sökord, t.ex. ”resultat”, men fann då inga ytterligare utredningar som kunde klassificeras som utvärderingar. Dessa tester gjordes bara på vissa delar av budgetpropositionen.

⁶ Vi har dock använt statistik över de utvärderingar som i budgetpropositionen för år 2000 var ”pågående”, ”på väg att tillsättas” eller liknande för att kontrollera att våra resultat inte beror på tillfälligheter. Av samma skäl har vi sökt igenom budgetpropositionen för år 1999 på samma sätt som årets. Denna statistik kommenteras på några ställen i rapporten och finns dessutom redovisad i tabellbilagan.

underlag som skulle betraktas som utvärderingar. Dessutom fanns risken att vår förfrågan skulle expedieras av personer som inte hade överblick över alla departementets områden och, kanske också, att man främst skickat med de ”goda exemplen”.

Det bästa tillvägagångssättet hade kanske varit att kombinera de båda metoderna. Om vi haft mer tid på oss hade vi kunnat skicka en lista på de utvärderingar vi funnit till departementen och bett dem lägga till sådant vi missat. Det hade möjligtvis kunnat ge en mer ”rättvis” beskrivning av skillnaderna mellan utgiftsområden. Nu får vi acceptera den metod vi valt och istället göra en reservation för att jämförelserna, i något fall, kan bli missvisande.

1.4 Disposition

Rapporten är disponerad på följande sätt. I kapitel 2 redovisar vi hur utvärderingarna fördelar sig på de olika utgiftsområdena. Kapitel 3 handlar om vilka som beställt respektive producerat utvärderingarna. Kapitel 4 analyserar i vilken mån utvärderingarna använts för att fatta beslut och i kapitel 5 diskuteras utvärderingarnas kvalitet.

2. Utvärderingarnas fördelning på utgiftsområden

I följande tabell redovisas hur utvärderingarna fördelar sig på de utgiftsområden där de finns omnämnda. Antalet utvärderingar i tabellen är något högre än de 114 olika utvärderingar vi granskat. Detta beror på att vissa utvärderingar är omnämnda i flera utgiftsområden och därför räknats mer än en gång i tabellen.

Tabell 2.1 Utvärderingarna fördelade på de olika utgiftsområdena och finansplanen

Utgiftsområde (Nr)	Antal
Utbildning och universitetsforskning (16)	30
Hälsovård, sjukvård och social omsorg (9)	13
Kultur, medier, trossamfund och fritid (17)	13
Arbetsmarknad och arbetsliv (14)	12
Regional utjämning och utveckling (19)	11
Samhällsekonomi och finansförvaltning (2)	9
Näringsliv (24)	5
Jord- och skogsbruk, fiske med anslutande näringar (23)	4
Rättsväsendet (4)	4
Internationellt bistånd (7)	3
Totalförsvaret (6)	3
Allmänna bidrag till kommunerna (25)	2
Skatteförvaltning och uppbörd (3)	1
Ekonomisk trygghet vid ålderdom (11)	1
Ekonomisk trygghet för familjer och barn (12)	1
Samhällsplanering, bostadsförsörjning och byggande (18)	1
Allmän miljö- och naturvård (20)	1
Energi (21)	1
Kommunikationer (22)	1
Rikets styrelse (1)	1
Finansplan	1
Ekonomisk trygghet vid sjukdom och handikapp (10)	1
Utrikesförvaltning och internationell samverkan (5)	0
Invandrare och flyktingar (8)	0
Ekonomisk trygghet vid arbetslöshet (13)	0
Studiestöd (15)	0
Statsskuldräntor m.m. (26)	0
Avgiften till europeiska gemenskapen (27)	0
Totalt antal utvärderingar	119

Ett utgiftsområde står i särklass vad det gäller antalet utvärderingar. Det är *Utbildning och universitetsforskning* som har hela 30 utvärderingar nämnda. Därefter kommer fem utgiftsområden med mellan tretton och nio utvärderingar nämnda. Dessa är *Hälsovård, sjukvård och social omsorg, Kultur, medier, trossamfund och fritid, Arbetsmarknad och arbetsliv, Regional utjämnning och utveckling* samt *Samhällsekonomi och finansförvaltning*.⁷

Det är anmärkningsvärt att hela 15 utgiftsområden endast har en eller ingen utvärdering nämnd. Att ett utgiftsområde inte kommenterar några avslutade utvärderingar i budgetpropositionen för år 2000 kan dock bero på rena tillfälligheter. Utvärderingar görs med ojämna mellanrum och kanske är det så att några av dessa utgiftsområden alldeles nyligen varit föremål för en omfattande utvärderingsverksamhet eller att en sådan är planerad inom en snar framtid.

Därför har vi även studerat i vilken mån pågående och kommande utvärderingar omnämnts i budgetpropositionen för år 2000. Dessutom har vi sökt igenom budgetpropositionen för år 1999 med samma metod som årets (se tabellerna B.1 och B.2 i bilagan). Tas dessa uppgifter med i beräkningen finner vi sex utgiftsområden med sammanlagt högst tre avslutade, pågående eller kommande utvärderingar omnämnda i budgetpropositionerna för åren 1999 och 2000. Av dessa är det två utgiftsområden som inte har någon utvärdering nämnd i någon av budgetpropositionerna. Detta gäller områdena *Studiestöd* och *Avgiften till Europeiska gemenskapen*. De utgiftsområden som har mellan en och tre utvärderingar nämnda är *Utrikesförvaltning och internationell samverkan, Ekonomisk trygghet vid ålderdom, Ekonomisk trygghet vid arbetslöshet* samt *Statskuldräntor m.m.* Alla fyra utgiftsområden som rör "Ekonomisk trygghet" har ett lågt antal utvärderingar även när man summerar antalet utvärderingar i budgetpropositionerna för åren 1999 och 2000.

⁷ Av dessa utgiftsområden har *Samhällsekonomi och finansförvaltning* och *Kultur, medier, trossamfund och fritid* nämnt en betydligt högre andel utvärderingar som är tunna sett till antalet sidor och som rör förhållandevis små frågor.

En ytterligare fördjupning som kan göras är att dela upp utvärderingarna på *verksamhetsområden* som är en underavdelning inom utgiftsområdena. Detta har vi endast gjort i begränsad utsträckning. Vi fann då ett anmärkningsvärt förhållande. I utgiftsområdet *Arbetsmarknad och arbetsliv* återfinns alla utvärderingar utom en under verksamhetsområdet *Arbetsmarknad*, medan det omfattande verksamhetsområdet *Arbetsliv* inte har någon utvärdering nämnd.

Att enbart studera antalet utvärderingar per utgiftsområde räcker dock inte för att bedöma inom vilka områden behovet av ytterligare utvärderingar är störst. Utgiftsområdena är ju väldigt olika både till storlek och innehåll. För att ytterligare analysera materialet har vi därför jämfört antalet utvärderingar per utgiftsområde med några faktorer som kan påverka utvärderingsbehovet. Vi har särskilt riktat in analysen på att finna områden där det görs få utvärderingar.

Antal utvärderingar jämfört med utgifter

I detta avsnitt har vi jämfört antalet utvärderingar med utgiftsområdenas utgifter enligt statsbudgeten (se tabellerna B.3 och B.4 i bilagan). Utgifternas storlek borde vara en faktor som påverkar behovet av utvärderingar.

Det utgiftsmässigt sett största området, *Allmänna bidrag till kommunerna*, med ca 14 procent av utgifterna för statsbudgeten, har bara två utvärderingar nämnda. Detta kan förklaras med att den verksamhet bidragen finansierar utvärderas inom det område som bidragen går till. Detta bekräftas av att två utgiftsområden med mycket kommunal verksamhet, *Hälsovård, sjukvård och social omsorg* samt *Utbildning och universitetsforskning* också är de områden som har flest utvärderingar. Inom utgiftsområdet *Allmänna bidrag till kommunerna* är det främst effektiviteten i bidragens konstruktion som ska utvärderas. Sambandet mellan dessa utgiftsområden visas av att en av de utvärderingarna som nämns i *Allmänna bidrag till kommunerna*, utvärderingen av hur de ökade bidragen till kommunerna använts, också finns nämnd i just *Hälsovård, sjukvård och social omsorg* samt *Utbildning och universitetsforskning*.

Utgiftsområdet *Totalförsvaret* har de tredje största utgifterna av alla utgiftsområden men endast tre utvärderingar nämnda och dessa tre utvärderas samma myndighet, FOA. Det rör sig dessutom om en årligen återkommande utvärdering av ett av FOA:s många forskningsområden där regeringen i budgetpropositionen för år 2000 även nämner de som gjorts under två tidigare år. Utöver dessa "löpande" utvärderingar finns det alltså inte en enda utvärdering nämnd. Ser man till utgifterna borde det finnas betydligt fler utvärderingar inom detta område. En möjlig förklaring är att Försvarsmakten har en annan planeringsprocess än övriga myndigheter med stora försvarsbeslut vart femte år. Inför dessa beslut används olika "hotbilder" för att dimensionera framtida behov. Man blickar således snarare framåt än bakåt i planeringsprocessen.

Även de utgiftsområden som handlar om "ekonomisk trygghet" har få utvärderingar i förhållande till utgifterna. Dessa fyra utgiftsområden står för 28 procent av statsbudgetens utgifter men deras andel av utvärderingarna är endast 3 procent. En anledning till att dessa områden har få utvärderingar nämnda under utgiftsområdet kan vara att den fördelningspolitiska analysen istället görs i en bilaga till budgetpropositionen, den s.k. fördelningspolitiska redogörelsen, och inte i särskilda utvärderingar. Överhuvudtaget är det tänkbart att mycket av utvärderandet inom dessa områden görs inom Finans- och Socialdepartementen utan att resultera i rapporter som omnämns inom respektive utgiftsområde.

Ser vi till utgifternas storlek borde även utgiftsområdet *Studiestöd* innehålla fler utvärderingar. Detta område har relativt stora utgifter men inga utvärderingar nämnda vare sig i 1999 års eller 2000 års budgetproposition. Detta kan möjligen förklaras av att utgiftsområdet endast rör *en* relativt väl avgränsad fråga. Självklart behöver även studiestödssystemet utvärderas ibland, men inte varje år, utan när behovet uppkommer vid mer ojämna tillfällen.

Statsskuldräntor m.m. och *Avgiften till EU* har också relativt stora utgifter men få utvärderingar. Även i dessa fall kan förklaringen vara att det är sakpolitiskt begränsade områden där det inte behövs utvärderingar varje år.

I följande avsnitt ska vi analysera just sambanden mellan utgiftsområdenas komplexitet och antalet omnämnda utvärderingar.

Antal utvärderingar jämfört med antal sidor i statsliggaren

Vi har i detta avsnitt använt antalet sidor i statsliggaren som ett mått på utgiftsområdenas komplexitet (se tabellerna B.3 och B.4 i bilagan). Vi har redan tidigare i kapitlet väckt tanken att utvärderingsbehovet, räknat i antal utvärderingar, borde vara större i väldigt komplexa utgiftsområden än i mer avgränsade.

Ställer vi antalet utvärderingar per utgiftsområde mot antalet sidor i statsliggaren visar det sig att det finns ett visst samband. Utgiftsområden med få sidor i statsliggaren har färre utvärderingar. Det finns tio områden med högst 26 sidor i statsliggaren. Alla dessa har relativt få utvärderingar nämnda i budgetpropositionerna. Här ingår bl.a. samtliga områden som handlar om ”Ekonomisk trygghet” och utgiftsområdena *Studiestöd; Statskuldräntor* samt *Avgiften till Europeiska gemenskapen*.

Tre utgiftsområden har över 200 sidor i statsliggaren. Två av dessa; *Utbildning och universitetsforskning* samt *Kultur, medier, trossamfund och fritid* återfinns högt upp vad gäller antal utvärderingar. Det tredje området är *Totalförsvar* som behandlas på 238 sidor i statsliggaren men endast har tre utvärderingar omnämnda i budgetpropositionen. Som vi redovisade i föregående avsnitt är det dessutom det tredje största sett till utgifterna.

Antal utvärderingar jämfört med antalet debattartiklar

En tredje faktor som torde påverka utvärderingsbehovet är i vilken mån området är i fokus för den politiska debatten eller ej. Ett område där det råder enighet kring politiken är troligen inte lika aktuellt för utvärdering som ett område som står i centrum för den politiska debatten.

Denna faktor har vi fångat upp genom att gå igenom DN-debatt under perioden juni 1997 till juni 1999 och undersöka vilka utgiftsområden som debatterats (se tabellerna B.3 och B.4 i bilagan).

Det finns ett visst samband mellan förekomst på DN-debatt och antalet utvärderingar. *Hälsovård, sjukvård och social omsorg, Arbetsmarknad och arbetsliv* samt *Utbildning och universitetsforskning* återfinns mycket högt upp på listan, både vad gäller debattartiklar och utvärderingar.

Även *Utrikesförvaltning och internationell samverkan* har många debattartiklar, de flesta med anknytning till EU. Ändå har vi inte funnit någon utvärdering under detta utgiftsområde. Detta är ett förhållande vi finner märkligt.

Ett utgiftsområde som visserligen bara ligger ”en bit ovanför mitten” vad gäller antalet debattartiklar, men som innehåller frågor som ofta framhålls som centrala för vår framtid, är *Allmän miljö- och naturvård*. Även i detta fall är det märkligt att antalet utvärderingar är så få

Var behövs det fler utvärderingar?

Sammanfattningsvis indikerar vår analys att det inom några områden förefaller som att antalet utvärderingar är lågt i förhållande till behovet. Dessa områden har inga eller mycket få utvärderingar trots att det finns faktorer som pekar på ett stort utvärderingsbehov:

- *Totalförsvaret* som både har höga utgifter och en mycket komplex verksamhet
- De fyra utgiftsområden som handlar om ”Ekonomisk trygghet” som har höga utgifter
- *Utrikesförvaltning och internationell samverkan* som har varit i fokus för den politiska debatten under senare år
- *Allmän miljö- och naturvård* som innehåller frågor som ofta framhålls som centrala för vår framtid
- Dessutom finner vi det märkligt att det i *Arbetsmarknad och arbetsliv* inte finns någon utvärdering under verksamhetsområdet *Arbetsliv*

Vi vill betona att vår översiktliga analys inte *bevisar* att de nämnda områdena är eftersatta vad det gäller utvärderingar. Det kan finnas skäl till att så få utvärderingar gjorts de senaste åren som vi inte lyckats fånga med vår metod. Det kan också vara så att det gjorts betydligt fler utvärderingar än vad våra siffror visar men att dessa inte alls omnämns i budgetpropositionen eller där kallas för något annat än ”utvärderingar”.

Varför så få sektorsövergripande utvärderingar?

Av alla de utvärderingar vi kartlagt är det ingen som utgör ett riktigt bra exempel på en sektorsövergripande utvärdering. Det finns några som berör de flesta områden som t.ex. den om effekterna av lagen om offentlig upphandling eller den om myndigheternas beredskap inför millennieskiftet. Dessutom finns det några få utvärderingar som nämns i två eller tre olika utgiftsområden, till exempel utvärderingen av de ökade statsbidragen till kommunerna som finns omnämnd i *Hälsovård, sjukvård och social omsorg, Utbildning och universitetsforskning* samt *Allmänna bidrag till kommunerna*. I samtliga dessa fall handlar det om utvärderingar som inte behandlar det vi menar med sektorsövergripande frågor utan där det snarare är utgiftsområdesindelningen som inte riktigt ”stämmer”.

Vi hittar inte någon utvärdering som nämns under flera utgiftsområden därför att den behandlar något effektivitetsproblem som uppstår p.g.a. att den offentliga sektorn är indelad i olika sektorer.

Ändå vet vi att många effektivitetsproblem inom den offentliga sektorn är just av sektorsövergripande karaktär och att det därför också borde finnas ett behov av sektorsövergripande utvärderingar. En förklaring till varför det görs så få sektorsövergripande utvärderingar kan vara organisationsproblem inom regeringskansliet. Beställarna av utvärderingar är kanske alltför inriktade på sitt eget område och förmår därför inte se över departements- och utgiftsområdesgränserna.

Statskontorets förslag

Vår kartläggning visar att det görs mycket få sektorsövergripande utvärderingar. Det finns anledning för Regeringskansliet att överväga hur man kan förstärka de mekanismer som gör att sektorsövergripande utvärderingar kommer till stånd.

3 Beställare och producenter

Detta kapitel handlar om vilka som beställer och vilka som producerar de utvärderingar som omnämns i budgetpropositionen. Vi inleder kapitlet med en redovisning av beställarna, därefter följer en analys av producenterna.

3.1 Beställarna

Regeringen mesta beställare

Av de utvärderingar som omnämns i budgetpropositionen är 58 procent beställda av regeringen. I tabell 3.1 redovisas hur dessa utvärderingar fördelas på olika departement.

Tabell 3.1 De utvärderingar som beställts av regeringen fördelade på olika departement, antal

Departement	Antal utvärderingar
Utbildning	14
Näring	10
Social	10
Arbetsmarknad	8
Finans	8
Kultur	7
Inrikes	3
Justitie	3
Utrikes	2
Jordbruk	1
Försvar	0
Miljö	0

En analys av vilka departement som är de flitigaste beställarna överensstämmer med vad man kan vänta sig efter redovisningen i kapitel 2 av under vilka utgiftsområden utvärderingarna presenteras. Utgiftsområdet *Utbildning och universitetsforskning* har de flesta utvärderingarna och följaktligen är också *Utbildningsdepartementet* den flitigaste beställaren.

Inga beställningar från "Försvaret" och "Miljö"

I den andra änden av tabellen finner vi *Försvarsdepartementet* och *Miljödepartementet* som inte har beställt någon utvärdering som finns redovisad i budgetpropositionen. Vi uttryckte ju i kapitel 2 vår förvåning över att antalet utvärderingar är lågt inom utgiftsområdena *Totalförsvaret* och *Allmän miljö- och naturvård*. Här får vi en indikation på att de departement som har hand om nämnda områden inte är några särskilt flitiga beställare av utvärderingar.

Det är också intressant att notera att Näringsdepartementet som är den näst största beställaren har sina beställningar spridda över fem olika utgiftsområden nämligen *Arbetsmarknad och arbetsliv*, *Regional utjämning och utveckling*, *Energi*, *Kommunikationer* och *Näringsliv*. Om man i Näringsdepartementets siffror även räknar in de numera införlivade Arbetsmarknads- och Inrikesdepartementen blir departementet den allra största beställaren.⁸

Drygt 40% är inte regeringsbeställningar

Av utvärderingarna i budgetpropositionen är det 42 procent som inte är beställda av regeringen. Det rör sig mestadels om utvärderingar som initierats och producerats av myndigheter. Indirekt kan säkert även många av dessa sägas vara beställda av regeringen. Många myndigheter har ju som en uppgift att kontinuerligt följa upp och utvärdera det som händer inom deras sakområden. Dessutom ingår det i samtliga myndigheters uppgift att redovisa hur effektivt verksamheten bedrivs.

En förklaring till att så många myndighetsinitierade utvärderingar kommenteras i budgetpropositionen kan vara att det i många fall snarare handlar om en del i regeringens rapporteringen av vad myndigheterna åstadkommit under året, än att regeringen använt utvärderingens resultat som underlag i sitt beslutsfattande (se vidare kapitel 4).

⁸ Det nya Näringsdepartementet bildades den 1 januari 1999 genom en sammanslagning av Arbetsmarknads-, Kommunikations-, Inrikes- samt Närings- och Handelsdepartementet.

Ett ökat EU-inflytande

Även EU:s ökad inflytande märks bland de utvärderingar som inte är beställda av regeringen. I denna grupp finns nämligen ungefär ett dussin utvärderingar som på ett eller annat sätt initierats av EU. Bakom de flesta av dessa står de s.k. övervakningskommittéerna. Genom dessa granskar EU-kommissionen att EU-anslagen till de olika ”målprogrammen” används som det är avsett. I kommittéerna ingår förutom representanter för EU-kommissionen även berörda departement, myndigheter och organisationer samt arbetsmarknadens parter.

3.2 Producenterna

I tabell 3.2 har vi gjort en grov kategorisering över vilka typer av organisationer som producerat de utvärderingar som omnämns i budgetpropositionen. Med producent menas i detta fall, enkelt uttryckt, den organisation som givit ut den rapport i vilken utvärderingen presenteras.

Tabell 3.2 Utvärderingarna fördelade på olika typer av producenter, procent

<i>Producent</i>	<i>Andel av samtliga utvärderingar</i>
Departementen själva	6%
Sektorsanknutna utvärderingsmyndigheter (BRÅ, IFAU, SIR, SIKA)	5%
Allmänna utvärderingsmyndigheter (RRV, Statskontoret)	7%
Fem stora sektorsmyndigheter (Socialstyrelsen, RFV, NUTEK, Skolverket, Högskoleverket).	26%
Övriga myndigheter, inklusive statliga stiftelser, bolag, m.m.	29%
Kommittéväsendet	6%
Konsulter, fristående experter m.m.	18%
EU-organ	2%
Totalt	100%

Få utvärderingar görs av utvärderingsorganisationer

En intressant iakttagelse är att de organisationer inom statsförvaltningen som har utvärdering som en huvuduppgift, eller åtminstone en mycket central uppgift, svarar för en ganska liten del av utvärderingarna. Vi tänker då på *kommittéväsendet*, *de allmänna utvärderingsmyndigheterna*; Statskontoret och RRV, och *de sektorsanknutna utvärderingsmyndigheterna* som Brottsförebyggande rådet (BRÅ), Statens institut för kommunikationsanalys (SIKA), Statens institut för regionalforskning (SIR) och Institutet för arbetsmarknadspolitisk utvärdering (IFAU). Dessa tre organisationstyper svarar bara för 18 procent av utvärderingarna.

Istället är det kategorierna *Fem stora sektorsmyndigheter* och *Övriga myndigheter*, som inte har utvärderingar som en lika framträdande uppgift, som är de största utvärderingsproducenterna. De fem stora sektorsmyndigheterna; Socialstyrelsen, RFV, NUTEK, Skolverket och Högskoleverket, svarar för 26 procent av alla utvärderingar.⁹ Den kategori som producerat flest utvärderingar är *Övriga myndigheter*, där vi även räknat in statliga stiftelser och bolag. Denna kategori har producerat 29 procent av utvärderingarna.

Tillsammans producerar således *Fem stora sektorsmyndigheter* och *Övriga myndigheter* mer än tre gånger så många utvärderingar som de tidigare nämnda utvärderingsorganisationerna.

Konsulter och fristående experter anlitas ofta

Kategorin *Konsulter, fristående experter m.m.* svarar för nästan en femtedel av utvärderingarna. Här ingår förutom privata konsultbolag, även en relativt stor andel "peer-reviews" där främst forskningsmyndigheter, universitet och liknande låtit fristående experter granska verksamheten.

Den andel av utvärderingarna som gjorts av *Konsulter, fristående experter m.m.* blir ännu högre om man som producent i stället räknar den "underproducent" som gjort själva utvärderingsarbetet i de fall då den organisation som producerat den skriftliga rapporten använt en sådan. Med en sådan definition svarar *Konsulter, fristående experter m.m.* för 37 procent av utvärderingarna. I många av dessa fall är det någon myndighet, ofta med forskningsinslag, som för objektivitetens skull, låtit någon extern utvärderare genomföra en granskning av myndigheten. Dessutom ingår här flera rapporter som givits ut av NUTEK men där själva utvärderingen utförts av något konsultföretag.

Endast 6 procent av utvärderingarna har gjorts inom departementen. Det kan vara en låg siffra i förhållande till det utvärderingsarbete som faktiskt bedrivs inom departementen. En förkla-

⁹ Dessa fem stora myndigheter är av olika karaktär med mer eller mindre starka inslag av utvärdering i sina uppgifter.

ring kan vara att mycket av regeringskansliets eget utvärderingsarbete inte resulterar i officiella rapporter som omnämns i budgetpropositionen.

Det mesta görs av departementens egna myndigheter

Vi har konstaterat att en mycket liten del av utvärderingarna görs av departementen själva. Men vad händer om man vidgar definitionen och jämför hur mycket av utvärderingarna som görs av departementet och dess egna myndigheter med hur mycket som läggs ut på från departementet mer oberoende organisationer? I denna analys inkluderar vi bara de utvärderingar som har regeringen som beställare.

Vi finner då att 67 procent av de regeringsbeställda utvärderingarna görs av departementet eller dess egna myndigheter, medan 34 procent görs av andra organisationer. När det gäller de utvärderingar som departementen beställer från andra organisationer är fördelningen jämn mellan *allmänna utvärderingsmyndigheter* (Statskontoret och RRV), *kommittéväsendet* och *konsulter, fristående experter m.m.*

Stora skillnader i organisation

Det finns stora skillnader i hur departement/utgiftsområden organiserar utvärderingsarbetet. En viktig skillnad är på vilket avstånd från departementet som utvärderingsresurserna finns. Vissa departement, framförallt Finansdepartementet, har betydande utvärderingsresurser inom departementet. Andra departement har särskilda utvärderingsmyndigheter eller stora, mer eller mindre utvärderingsinriktade, sektorsmyndigheter att vända sig till. Medan en tredje grupp departement knappast har några sådana ”fasta” utvärderingsresurser överhuvudtaget.

Det går inte att utifrån vår statistik visa att tillgång till särskilda utvärderingsorgan eller utvärderingsinriktade sektorsmyndigheter leder till att fler utvärderingar görs. Sannolikt är sambandet starkare åt andra hållet, dvs. att det främst är på områden där det finns stora utvärderingsbehov som de fasta utvärderingsresurserna inrättats.

Egna myndigheter eller oberoende organisationer?

Hur ser det då ut på de områden där vi hittat tillräckligt många utvärderingar för att man ska kunna dra några slutsatser? Är det stora skillnader i vilken typ av organisationer som departementen anlitar? Det finns exempel på departement/utgiftsområden där nästan alla utvärderingar som nämns i budgetpropositionen produceras av departementet och dess egna myndigheter. Det finns också exempel på en större spridning av producenter.

Av de tio utvärderingar som är beställda av *Socialdepartementet* är en producerad av departementet självt, fem av Socialstyrelsen, två av Riksförsäkringsverket och en av Folkhälsoinstitutet. Endast en av utvärderingarna, gjord av en parlamentariskt sammansatt kommitté, har en lösare koppling till departementet.

Utgiftsområdet *Arbetsmarknad och arbetsliv* är ett exempel på ett område med en större spridning av utvärderare, trots att där finns en särskild utvärderingsmyndighet, nämligen Institutet för Arbetsmarknadspolitisk Utvärdering (IFAU). Institutet har dock bara producerat två av åtta departementsbeställda utvärderingar. I övrigt är spridningen stor; Statskontoret, RRV, Ungdomsstyrelsen, Växjö Universitet och IM-gruppen är de utvärderare inom området som inte tillhör Näringsdepartementet.

En svår avvägning

Exemplen illustrerar en svår avvägning. En hög utvärderingskapacitet inom departementet, särskilda utvärderingsmyndigheter knutna till departementet och kanske också starka sektorsmyndigheter ger en hög sakkunskap i utvärderingarna. Samtidigt finns risken att utvärderingarna inte kommer att betraktas som oberoende.

Motsatsen gäller när utvärderingarna istället läggs ut på fristående organisationer. Då minskar risken att utvärderingarna kommer att betraktas som partsinlagor men å andra sidan finns risken att sakkunskapen ifrågasätts.

Förmodligen är en optimal strategi varken den där alla utvärderingar görs "nära" departementet eller den där alla utvärderingar läggs ut på från departementet fristående organisationer. En mer flexibel lösning är förmodligen bättre där man, beroende på vilken fråga utvärderingarna gäller, från fall till fall, avgör om "sakkunskapsargumentet" eller "objektivitetsargumentet" ska väga tyngst.

Statskontorets förslag

Departementens bör, i sina utvärderingsstrategier, behandla hur man hanterar avvägningen mellan oberoende och sakkunskap när man beslutar om vilken organisation som ska genomföra utvärderingarna.

4. Utvärderingarna i beslutsfattandet

I detta kapitel analyserar vi på vilket sätt regeringen använder utvärderingarna i budgetpropositionen. Kommenteras de bara allmänt eller används de som underlag för något beslut? Innan vi ger oss in i analysen är det viktigt att påpeka två saker. För det första är det möjligt att utvärderingarna kan leda till beslut i ett senare skede som redovisas någon annanstans än i budgetpropositionen. För det andra har utvärderingar ett vidare syfte än att leda till direkta beslut, nämligen som en länk i en långsiktig kunskapsuppbyggnad inom ett visst ämne.

Fyra kategorier av kommentarer

I vår analys av hur utvärderingarna används har vi gjort en uppdelning i fyra kategorier, efter hur regeringen kommenterar utvärderingen i budgetpropositionen. I tabell 4.1 har utvärderingarna delats upp i de fyra kategorierna.

Med ”Kommenterar” avses att regeringen nämner utvärderingen i budgetpropositionen, utan att göra några egna värderingar. Detta kan betyda att utvärderingen används för att bygga upp kunskap, att den läggs på minnet för att kunna plockas fram vid framtida behov. Det kan också vara så att regeringen kommenterar utvärderingen som en del i redovisningen av vad myndigheterna presterat. Ett av budgetpropositionens syften är att redovisa vilka resultat förvaltningen åstadkommit.

Med ”Bekräftar” avses att regeringen inte fattar något nytt beslut med anledning av utvärderingen utan använder dess resultat för att bekräfta att den förda politiken inom det område som utvärderats fungerar bra.

Med kategorierna ”Återkommer” och ”Förändrar” avses att regeringen redovisar någon form av beslut med anledning av utvärderingen. I kategorin ”Återkommer” ingår beslut om ytterligare utredning, remissförfarande och liknande samt när regeringen lovar

att återkomma i ärendet. Med ”Förändrar” avses att man beslutat om någon förändring av politiken.

Tabell 4.1 Utvärderingarna fördelade efter hur de kommenteras i budgetpropositionen, antal och procent.

	Antal	Procent
Kommenterar	48	42%
Bekräftar	30	26%
Återkommer	22	19%
Förändrar	14	12%
Totalt	114	100%

Drygt 40 procent av alla utvärderingar som omnämns i budgetpropositionen, kommenteras endast allmänt av regeringen. Omkring en fjärdedel av utvärderingarna används för att bekräfta den förda politiken. En tredjedel av utvärderingarna nämns i samband med att regeringen redovisar något beslut.

Vill man vara positiv kan kanske även de fall där regeringen ”bekräftar” den förda politiken med hjälp av utvärderingens resultat ses som en form av beslut. Om utvärderingens resultat istället visat att politiken fungerar dåligt hade kanske regeringen beslutat om någon förändring. Med ett sådant synsätt används drygt hälften av de utvärderingar som omnämns i budgetpropositionen i beslutsfattandet.

Få utvärderingar leder till förändring

För drygt 30 procent av de utvärderingar vi kartlagt redovisar regeringen något slags beslut i samband med att utvärderingen nämns i budgetpropositionen.

I över hälften av dessa fall är beslutet att regeringen avser att återkomma till frågan. Oftast sker detta utan närmare precisering. Ibland nämns skäl som att ärendet är under remiss eller att man vill avvakta den slutliga utvärderingen som kommer senare. I några fall aviserar regeringen att den avser att tillsätta en utredning eller något liknande.

Av de totalt 114 utvärderingar vi kartlagt är det således endast i 14 fall (12 procent) som regeringen beslutar att göra någon förändring i politiken. I majoriteten av dessa fall handlar det om små förändringar. Några exempel är:

- RFV:s utvärdering av de nya reglerna för underhållsstöd resulterade i att regeringen beslutade om förändrat grundavdrag och om förändrade procentsatser som bestämmer återbetalningskyldigheten.
- Regeringen beslutar med anledning av Kungliga bibliotekets utvärdering att inordna SPPB (Statens Psykologisk Pedagogiska Bibliotek) under Stockholms universitetsbibliotek.
- Utvärderingen av Kulturnät Sverige visade enligt regeringen att verksamheten är en samhällsnyttighet som väl motiverar ett statligt åtagande. Regeringen beslutar att göra kulturnätet till en permanent del av Kulturrådets åtaganden.

Vi har inte noterat några större skillnader mellan utgiftsområdena när det gäller i hur hög grad utvärderingarna används för att fatta beslut.

Varför leder så få utvärderingar till beslut?

Vår analys visar att det är ganska få av utvärderingarna i budgetpropositionen som omnämns i samband med ett konkret beslut om en förändring av politiken. Det kan finnas flera förklaringar till detta. En förklaring kan vara att budgetpropositionen delvis har karaktären av resultatredovisning, d.v.s. där redovisas vad myndigheterna har åstadkommit under året. Det betyder att många utvärderingar kanske snarare omnämns som exempel på myndighetsresultat än för att de har relevans för regeringens beslutsfattande. Detta kan förklara den höga andel utvärderingar där regeringen nöjer sig med allmänna kommentarer.

En annan förklaring kan vara att utvärderingar oftare utgör en del i ett långsiktigt kunskapsuppbyggande än ett direkt beslutsunderlag. Mot detta talar att vi endast funnit enstaka exempel på att regeringen i budgetpropositionen nämner utvärderingar som är mer än något år gamla.

Ytterligare en förklaring kan vara att utvärderingarna leder till beslut som fattas i ett annat sammanhang än i budgetpropositionen. Det skulle kunna vara så att regeringen i många av de fall där de beslutat att återkomma till frågan fattar ett beslut i någon annan proposition.

Men det kan också vara så att regeringens användning av utvärderingarna inte är större än vad våra siffror indikerar. Det kan i såfall bero på att utvärderingarna håller för låg kvalitet. Men det kan också vara så att regeringens beställningar inte är tillräckligt genomtänkta så att utvärderingen svarar på andra frågor än de som är relevanta för beslutsfattandet. Vår översiktliga kartläggning är inte tillräcklig för att kunna säga något om varför så få utvärderingar används som beslutsunderlag, eller ens om detta verkligen är fallet.

En noggrant genomtänkt utvärderingsstrategi kan öka regeringens möjligheter att använda utvärderingar som beslutsunderlag. För detta behövs mer analys kring var i kedjan, beställning – utvärdering – mottagande - beslut, som bristerna finns.

Statskontorets förslag

Statskontoret föreslår att regeringen beslutar om en särskild utredning om varför inte fler utvärderingar används i beslutsfattandet. Vårt förslag är att denna utredning använder två olika angreppssätt som tillsammans bör utgöra ett bra underlag för förslag som ökar regeringens användande av utvärderingar.

- Det ena angreppssättet är att utgå från ett antal av regeringen fattade beslut och undersöka i vilken mån utvärderingar ingått i beslutsunderlaget, samt vilken typ av utvärderingar som ”borde” ha ingått i underlaget.

- Det andra angreppssättet är att utgå från ett antal utvärderingar gjorda för några år sedan och undersöka hur dessa har använts, och varför de inte i högre grad använts i beslutsfattandet.

5 Utvärderingarnas kvalitet

I detta kapitel gör vi, dels en allmän bedömning av hur utvärderingarna presenteras (5.1), dels en närmare granskning av två faktorer som har betydelse för utvärderingarnas användbarhet som beslutsunderlag för regeringen (5.2).

Olika målgrupper och oklarhet kring begreppet utvärdering

Innan vi ger oss in i analysen är det två saker vi vill betona som påverkar de bedömningar som görs i detta avsnitt:

1. Målgrupperna skiftar betydligt för de 114 utvärderingar som ingår i vår kartläggning. Det handlar om allt ifrån utförliga kommittébetänkanden i frågor där man kan tänka sig en relativt bred läsekrets, till ganska myndighetsinterna dokument på några få sidor, där kraven på en tydlig och utförlig metoddokumentation är lägre.

2. Det råder oklarhet kring vad begreppet utvärdering egentligen står för. Bland de rapporter som kallas "utvärdering" i budgetpropositionen och som därför ingår i vår analys finns det åtskilliga exempel som knappast skulle benämnas "utvärdering" om de vanligaste definitionerna av begreppet tillämpades. Det finns t.o.m. några exempel på rapporter som författarna inte på något ställe kallar för utvärdering, troligen p.g.a. att de anser att definitionen inte uppfylls, men som regeringen presenterar som en "utvärdering" i budgetpropositionen.

Vår statistik visar att antalet rapporter som benämns utvärdering ökat med 35 procent mellan budgetpropositionen för år 1999 och budgetpropositionen för år 2000. Det är inte alls säkert att hela denna ökning beror på att det görs fler utvärderingar eller att fler av de som görs omnämns i budgetpropositionen. Det kan också vara så att begreppet utvärdering fått en allt vidare definition i takt med att utvärderingarnas betydelse som beslutsunderlag alltmer betonats.

5.1 Hur utvärderingarna presenteras

Den bedömning vi gör i detta avsnitt avser inte utvärderingarnas metodologiska kvalitet utan hur de presenteras i skriftlig form.

Är det enkelt att förstå hur utvärderarna gått tillväga, vilka metoder de använt och därmed hur pass trovärdiga utvärderingens resultat är? Eller är detta svårt eller kanske t.o.m. omöjligt?

Skiftande kvalitet i presentationen

Vår bedömning är att kvaliteten i hur utvärderingarna presenteras är mycket skiftande. Det finns goda exempel där det som läsare är lätt att följa hur det hela gått till. Men det finns samtidigt alltför många exempel på mycket svårigenomträngliga rapporter där det är mycket svårt att bedöma kvaliteten i de metoder som använts och därmed tillförlitligheten i utvärderingens resultat. Det finns t.o.m. exempel på rapporter, som torde ha en relativt bred läsekrets, där det överhuvudtaget inte nämns vilken metod man använt och hur man gått tillväga. Man hoppas tydligen att läsarna trots detta ska acceptera de slutsatser som dras.

De brister vi noterat i hur utvärderingarna presenteras är av många olika slag. I vissa fall har man som läsare svårt att se kopplingen mellan de slutsatser som dras och de resultat som presenteras. I andra fall är det metoden bakom resultaten som är bristfälligt dokumenterad. Dåliga beskrivningar av vilken insats, åtgärd eller liknande som egentligen utvärderas förekommer också.

Brister i presentation minskar användbarheten

Tillsammans innebär detta att många utvärderingar, även om de egentligen är metodologiskt acceptabla, mister en del av den kraft de skulle kunna ha:

- svårigheter att bedöma slutsatsers grund eller tillförlitligheten i resultaten gör utvärderingarna mindre användbara i beslutsfattande och kunskapsuppbyggande

- svårigheter att förstå utvärderingsmetoderna gör att spridningen av goda metodexempel försvåras
- en dålig beskrivning av vilka insatser, åtgärder etc. som utvärderas gör det svårt för den icke initierade att dra slutsatser som kan föras över på andra områden

Statskontorets förslag

Genom att låta vissa grundläggande regler för hur en utvärdering ska presenteras bli en del av varje departementens utvärderingsstrategi kan den skiftande kvaliteten härvidlag minskas. En ökad enhetlighet underlättar också för den som tar del av många olika utvärderingar. Det måste finnas en tydlighet i såväl vilka resultat utvärderingens slutsatser grundas på som vilken metod som använts för att få fram resultaten.

5.2 Två betydelsefulla metodfaktorer

Det finns många faktorer som påverkar kvaliteten i en utvärdering. Vi har gjort en översiktlig granskning av två faktorer som har stor betydelse för regeringens möjligheter att använda utvärderingar som beslutsunderlag.

Den ena faktorn är *måttstocken*, d.v.s. vad man har jämfört den verksamhet eller åtgärd som utvärderas med. Jämförelser är grunden för att man ska kunna avgöra om något är bra eller dåligt. Vi har bedömt om utvärderingarnas måttstockar är tydliga eller otydliga. Den andra faktor vi granskat är om de resultatmått som utvärderingarna bygger på avser de statliga insatsernas *effekter på samhällsnivå* eller något annat.

Vi vill betona att våra bedömningar inte bara speglar kvaliteten i utvärderarnas arbete utan, kanske i minst lika hög grad, även hur utvärderingsvänlig omvärld utvärderingen gjorts i. Det är säkert många fall där vi underkänt utvärderingar, trots att utvärderarna gjort ett mycket bra jobb utifrån de förutsättningar de haft, men där t.ex. snäva tidsramar, brist på relevant statistik och liknande gjort deras uppgift mer eller mindre omöjlig.

Två tredjedelar av utvärderingarna har en otydlig måttstock

Med måttstock avses det ”nollalternativ” som den utvärderade verksamheten eller åtgärden jämförs med. Vår kartläggning visar att nästan två tredjedelar av utvärderingarna inte har någon tydlig måttstock. En tydlig måttstock underlättar bedömningen av om den utvärderade verksamheten är bra eller dålig. Saknas däremot en tydlig måttstock blir denna bedömning betydligt svårare.

De flesta utvärderingar som innehåller tydliga måttstockar har gjort jämförelser med, antingen tiden före den utvärderade åtgärden infördes, eller jämfört med någon ”kontrollgrupp” som inte omfattats av åtgärden.¹⁰ Till den sistnämnda kategorin räknas också utvärderingar där man jämför hur samma åtgärd påverkat olika grupper. En del utvärderingar kombinerar användningen av före-efter jämförelser och kontrollgrupper till experiment eller s.k. kvasiexperiment. Detta ger en ännu bättre måttstock.

Det finns också enstaka exempel på tydliga måttstockar som består av något ”mål” uttryckt i en enda siffra, som t.ex. att arbetslösheten ska vara högst fyra procent år 2000. I sådana fall är det oftast underförstått att det handlar om en sänkning eller höjning från någon tidigare känd nivå och resultatmåttan får därför ändå en karaktär av före-efter jämförelse.

I den majoritet av utvärderingar som har otydliga måttstockar lyser oftast före-efter jämförelser eller jämförelser med kontrollgrupper med sin frånvaro.

Ofta förekommande exempel på otydliga måttstockar är utvärderingar som grundas på attitydundersökningar, gjorda vid ett enda tillfälle och utan kontrollgrupper, där respondenterna visserligen fått uttala sig om den utvärderade verksamheten eller åtgärden fungerar ”bra” eller ”dåligt” men där det är oklart vad de som svarat egentligen jämfört med.¹¹

¹⁰ Om utvärderingen inte gäller en specifik åtgärd utan snarare en organisations löpande verksamhet passar inte begreppen ”före-efter” och ”kontrollgrupp”. Motsvarande måttstockar finns då i form av ”tidsserier” och jämförelser med ”liknande verksamheter”.

¹¹ Detta ska inte uppfattas som kritik mot attitydundersökningar som mätinstrument, utan mot att man mäter vid ett enda tillfälle och utan kontrollgrupper och dessutom utan att ställa frågorna på ett sådant sätt att den som svarar får en tydlig måttstock att utgå ifrån.

Det är också vanligt med utvärderingar som bygger på att ”utomstående bedömare” gör en, mer eller mindre ingående, granskning av en verksamhet ofta genom att läsa den dokumentation som finns om verksamheten samt samtala med personal och en del avnämare. Detta leder till kvalitativa värdebedömningar som t.ex. att ”verksamheten fungerar bra” men det är oklart vad de har jämfört med när de gjort bedömningen.

I båda dessa fall är det troligt att det finns måttstockar (i respondenternas eller bedömarnas huvuden) men det är mycket svårt för den som läser utvärderingen att avgöra vad bedömningen grundar sig på.

Det finns även utvärderingar som använt sig av före-efter jämförelser eller kontrollgrupper, men där vi ändå bedömt måttstocken som otydlig. Det kan handla om att man valt en kontrollgrupp som inte verkar riktigt relevant eller helt enkelt är otillräckligt beskriven i studien för att man ska kunna lita på jämförelsen.

Mäter utvärderingarna effekter?

Vi har också bedömt utvärderingarna utifrån några av de önskemål Finansutskottet haft på den resultatinformation som lämnas till riksdagen i budgetpropositionen. De vill bl.a. att resultatinformationen i högre grad ska var inriktad på de statliga insatsernas *effekter på samhällsnivå* snarare än på myndigheternas prestationer och interna göromål. De vill också att resultaten i högre grad redovisas i kvantitativa termer med hjälp av indikatorer eller nyckeltal.¹²

Vi har gjort en grov kategorisering av utvärderingarna efter i vilken mån de till en väsentlig del innehåller resultatmätt som avser effekter på samhällsnivå eller ej.

¹² se t.ex. bet 1998/99: FiU20, bet 1997/98:FiU20, bet 1995/96:FiU7 och bet 1993/94FiU10.

Var femte mäter effekter med sträng definition

Med en sträng definition av *effekter på samhällsnivå* är det mellan 20 och 25 procent av utvärderingarna som bygger på sådana mått. I denna definition innefattas att måtten ska avse faktiskt inträffade effekter, att de ska vara uppmätta på ett objektivt sätt, samt att de ska avse den mest relevanta länken i effektkedjan sett utifrån syftet med den utvärderade verksamheten.

Några exempel på utvärderingar som uppfyller denna definition är en studie av hur elektronisk övervakning påverkar återfallsfrekvensen för dem som dömts för rattfylleri och en studie av hur mängden lärartimmar i grundskolan påverkar elevernas snittbetyg. De flesta exemplen finns dock inom arbetsmarknadsområdet där vi funnit många studier av hur olika typer av åtgärder leder till att folk får arbete.

Med en vidare definition är nästan hälften effektmätningar

Vidgar man definitionen och räknar med alla slags mått som gäller effekterna på samhällsnivå stiger andelen utvärderingar som till en väsentlig del innehåller sådana mått till mellan 40 och 50 procent. Då räknar vi med även effektmått som vi bedömt som subjektiva, som mäter något annat än faktiskt inträffade effekter eller som inte nått tillräckligt långt i effektkedjan.

Exempel på subjektiva mått är en studie av ett antidrogprogram där resultaten bygger på hur mycket narkotika individerna i målgruppen uppger att de konsumerat och flera studier av glesbygdsprojekt där de "projektägare" som administrerar projekten bedömer hur många jobb som skapats eller räddats. Ett exempel på en utvärdering som "mäter" något annat än faktiskt inträffade effekter är en jämförelse mellan olika regionalpolitiska åtgärder där slutsatserna grundas på hypotetiska antaganden om företags agerande enligt mikroekonomisk teori. Ett exempel på en studie som inte går tillräckligt långt i effektkedjan är när de extra resurser regeringen anslagit för att höja kvaliteten i vård, skola och omsorg bedöms utifrån kommunernas enkätsvar om huruvida pengarna verkligen använts till dessa ändamål.

En dryg tredjedel gäller kvaliteten

Mellan 35 och 40 procent av utvärderingarna avser inte alls, eller bara till en liten del, effekter på samhällsnivå. Dessa kan klassificeras som kvalitetsstudier där fokus ligger på kvaliteten i myndigheternas prestationer snarare än på vilka effekter dessa prestationer åstadkommer.

Några av utvärderingarna i denna kategori använder olika slags nyckeltal som huvudsakliga mått. Ett exempel är Konjunkturinstitutet som jämför träffsäkerheten i sina prognoser med andra prognosinstitut. De flesta utvärderingarna i denna kategori bygger dock antingen på resultat från attitydundersökningar till kunder eller på utomstående bedömares granskning av verksamheten. I det senare fallet är redovisningen av resultat i kvantitativa termer ofta obetydlig.

Det finns stora skillnader i vilket syfte utvärderingarna i denna kategori har. Vissa, som t.ex. Konjunkturinstitutets jämförelser av träffsäkerheten i prognoserna, har inget annat syfte än att bedöma just kvaliteten. Många andra utvärderingar i denna kategori har dock som ett mer eller mindre uttalat syfte att bedöma effekter. Av olika skäl kommer man inte så långt. Det finns flera exempel på att man åberopar bristande statistik om effekter som skäl till att man måste nöja sig med mer ”verksamhetsnära” mått.

Var femte handlar om politikens genomförande

Återstår att redovisa en kategori. Det gäller de ca 20 procent som vi klassificerat som *genomförandevärderingar*. Dessa bedömer i första hand om verksamheten genomförs enligt uppdragsgivarens riktlinjer och om processen fungerar som det var tänkt. Många av genomförandevärderingarna har samband med igångsättandet av tre stora satsningar nämligen kunskapslyftet, EU:s strukturfondsbidrag och försöket med friare användning av arbetsmarknadspolitiska medel. I samtliga dessa fall har man i ett senare skede gjort (eller avser att göra) effektutvärderingar.

Genom att kombinera genomförande- och effektutvärderingar ökar möjligheterna att förklara vad resultaten i effektutvärderingen beror på.

Mindre än var femte utvärdering uppfyller båda kriterierna

Om vi väger samman bedömningarna i de föregående avsnitten finner vi att mindre än 20 procent av de utvärderingar vi granskat klarar den stränga definitionen på att mäta effekter samtidigt som de innehåller en tydlig måttstock.

Vad är förklaringen till att inte fler utvärderingar klarar båda kriterierna? Till en del handlar säkert förklaringarna om korta tidsramar, knappa resurser eller bristande metodkunskap hos dem som utfört eller beställt utvärderingen. En annan förklaring är att det inom vissa politikområden är mycket svårt att finna bra effektmått. Det gäller t.ex. ofta forskningsverksamhet där effekterna på samhällsnivå kan vara svåra att kvantifiera och dessutom ofta inträffar många år efter de insatser som ska utvärderas. Det är troligen en anledning till att de utvärderingar av forskningsverksamhet som vi funnit snarare koncentreras till forskningens kvalitet än dess effekter.

Vi är dock övertygade om att det går att göra fler högkvalitativa utvärderingar av effekter på samhällsnivå än det görs idag. Dessutom kan kvaliteten i de utvärderingar som inte avser sådana effekter höjas betydligt. Vi vill här lyfta fram två nyckelfaktorer som på ett väsentligt sätt skulle öka möjligheterna att åstadkomma utvärderingar som ger regeringen ett bättre beslutsunderlag.

Tänk efter före

Den ena nyckelfaktorn handlar om att tänka i utvärderingstermer redan i samband med att nya åtgärder eller verksamheter planeras. För att säkra tydliga måttstockar och relevanta resultatmått krävs ofta att utvärderingarna planeras redan i samband med att den åtgärd som ska utvärderas introduceras. För att kunna göra en bra före-efter studie behöver man säkra data om tillståndet före åtgärden kommit igång. För att kunna göra en bra jämförelse med en kontrollgrupp krävs ofta att man, i detta tidiga skede, utser försöksgrupper och kontrollgrupper. Sätter man däremot igång åtgärden utan att tänka på detta blir det betydligt svårare att göra en bra utvärdering när detta blir aktuellt.

Vi har funnit enstaka exempel där man verkligen tänkt efter före och dels säkrat god statistik, dels introducerat den verksamhet som ska utvärderas försöksvis i vissa regioner, medan man låtit andra regioner utgöra kontrollgrupper. Om liknande försöksverksamheter genomfördes oftare i samband med att nya åtgärder eller verksamheter introduceras skulle de som ska fatta beslut utifrån utvärderingarna få ett säkrare beslutsunderlag.

Bra uppföljning underlättar bra utvärdering

Den allra viktigaste faktorn för att öka möjligheterna att producera utvärderingar av hög kvalitet är att den löpande uppföljningen inom det område som ska utvärderas är anpassad till att göra utvärderingar. Med detta avses bl.a. att det, på ett så smidigt sätt som möjligt, skall gå att koppla ihop kostnader, insatser, prestationer och effekter för den verksamhet eller åtgärd som ska utvärderas. Ett, på detta sätt, flexibelt uppföljningssystem ökar t.ex. möjligheten att göra bra före-efter studier även om man "glömt" att planera för utvärderingen innan åtgärden infördes. I de fall man planerar för utvärderingen i god tid så blir datainsamlingen betydligt enklare och billigare än om den löpande uppföljningen på området inte är anpassad för utvärderingar.

Statskontorets förslag

En grundläggande förutsättning för att uppnå en hög kvalitet i utvärderingarna är en bra löpande uppföljning inom de politiskområden departementet ansvarar för. Departementen bör i sina uppföljnings- och utvärderingsstrategier lägga stor vikt vid att den löpande uppföljningen anpassas så att möjligheterna att göra utvärderingar förbättras.

I departementens utvärderingsstrategier bör riktlinjer för hur man planerar för utvärdering redan i samband med att nya verksamheter, åtgärder eller liknande introduceras, vara ett mycket viktigt inslag.

För att verkligen betona vikten av att utvärderingar kommer in tidigt i beslutsprocessen bör regeringen dessutom överväga att införa ett generellt krav på att framtida utvärderingar finns med i beredningen redan i samband med att nya åtgärder och verksamheter beslutas.

Bilaga, tabeller

Tabell B.1 Totala antalet utvärderingar (avslutade, pågående och kommande) i budgetpropositionerna för år 1999 och år 2000

Utgiftsområde	1999	2000	Totalt
Rikets styrelse (1)	5	8	13
Samhällsekonomi och finansförvaltning (2)	8	16	24
Skatteförvaltning och uppbörd (3)	3	3	6
Rättsväsendet (4)	13	22	35
Utrikesförvaltning och internationell samverkan (5)	0	2	2
Totalförsvaret (6)	5	6	11
Internationellt bistånd (7)	14	8	22
Invandrare och flyktingar (8)	1	6	7
Hälsovård, sjukvård och social omsorg (9)	30	38	68
Ekonomisk trygghet vid sjukdom och handikapp (10)	1	4	6
Ekonomisk trygghet vid ålderdom (11)	0	2	2
Ekonomisk trygghet för familjer och barn (12)	4	2	6
Ekonomisk trygghet vid arbetslöshet (13)	2	1	3
Arbetsmarknad och arbetsliv (14)	23	25	48
Studiestöd (15)	0	0	0
Utbildning och universitetsforskning (16)	44	57	101
Kultur, medier, trossamfund och fritid (17)	17	31	48
Samhällsplanering, bostadsförsörjning och byggande (18)	12	6	18
Regionalutjämnning och utveckling (19)	16	21	37
Allmän miljö- och naturvård (20)	6	4	10
Energi (21)	7	16	23
Kommunikationer (22)	5	6	11
Jord- och skogsbruk, fiske med anslutande näringar (23)	14	14	28
Näringsliv (24)	11	19	30
Allmänna bidrag till kommunerna (25)	1	5	6
Statsskuldräntor m.m. (26)	1	0	1
Avgiften till europeiska gemenskaperna (27)	0	0	0
Finansplan	5	12	17
Totalt	248	334	583

Tabell B.2 Samtliga, Avslutade, Pågående och Kommande utvärderingar i budgetpropositionerna för år 1999 och år 2000 fördelade inom de olika utgiftsområdena och finansplanen¹³

BP år 1999				
Uo	A	P	K	S:a
1	1	1	3	5
2	5	3	0	8
3	2	0	1	3
4	2	6	5	13
5	0	0	0	0
6	1	3	1	5
7	10	2	2	14
8	0	1	0	1
9	20	6	4	30
10	1	0	0	1
11	0	0	0	0
12	2	1	1	4
13	1	0	1	2
14	7	13	3	23
15	0	0	0	0
16	29	12	3	44
17	7	9	1	17
18	7	5	0	12
19	6	7	3	16
20	4	2	0	6
21	4	1	2	7
22	3	0	2	5
23	9	5	0	14
24	8	1	2	11
25	0	1	0	1
26	0	1	0	1
27	0	0	0	0
Fipl	2	3	0	5
S:a	131	83	34	248

BP år 2000				
Uo	A	P	K	S:a
1	2	3	3	8
2	12	2	2	16
3	1	0	2	3
4	4	15	3	22
5	2	0	0	2
6	1	1	4	6
7	3	1	4	8
8	0	4	2	6
9	16	16	6	38
10	0	3	1	4
11	2	0	0	2
12	1	1	0	2
13	0	1	0	1
14	14	7	4	25
15	0	0	0	0
16	31	13	13	57
17	17	7	7	31
18	1	5	0	6
19	12	7	2	21
20	1	2	1	4
21	3	1	12	16
22	3	2	1	6
23	5	7	2	14
24	8	10	1	19
25	2	2	1	5
26	0	0	0	0
27	0	0	0	0
Fipl	5	4	3	12
S:a	146	114	74	334

¹³ Att siffran över antalet avslutade utvärderingar i BP för år 2000 är betydligt högre än de 114 utvärderingar vi analyserat beror på att vi i tabellen räknat vissa utvärderingar flera gånger då de omnämns på flera olika ställen inom samma utgiftsområde.

Tabell B.3 Antal avslutade utvärderingar, utgifter, antal sidor i statsliggaren samt antal artiklar på DN-debatt, per utgiftsområde år 2000

Utgiftsområde (Nr)	Utv.	Miljard. Kr.	Sid. i statsligg.	DN-debatt
Utbildning och universitetsforskning (16)	30	32,573	270	18
Hälsovård, sjukvård och social omsorg (9)	13	25,363	110	16
Kultur, medier, trossamfund och fritid (17)	13	7,57	265	1
Arbetsmarknad och arbetsliv (14)	12	46,13	88	22
Regionalutjämning och utveckling (19)	11	3,31	38	1
Samhällsekonomi och finansförvaltning (2)	9	1,578	100	3
Näringsliv (24)	5	2,983	169	9
Jord- och skogsbruk, fiske med anslutande näringar (23)	4	9,726	114	1
Rättsväsendet (4)	4	22,667	85	14
Internationellt bistånd (7)	3	13,237	68	4
Totalförsvaret (6)	3	46,649	238	4
Allmänna bidrag till kommunerna (25)	2	97,667	11	6
Skatteförvaltning och uppbörd (3)	1	5,922	26	2
Ekonomisk trygghet vid ålderdom (11)	1	33,619	5	6
Ekonomisk trygghet för familjer och barn (12)	1	44,755	6	3
Samhällsplanering, bostadsförsörjning och byggande (18)	1	15,592	94	4
Allmän miljö- och naturvård (20)	1	1,721	60	5
Energi (21)	1	1,446	52	0
Kommunikationer (22)	1	25,532	106	2
Rikets styrelse (1)	1	4,461	57	6
Finansplan	1	0	0	0
Ekonomisk trygghet vid sjukdom och handikapp (10)	1	92,47	23	2
Utrikesförvaltning och internationell samverkan (5)	0	2,983	52	17
Invandrare och flyktingar (8)	0	4,901	41	4
Ekonomisk trygghet vid arbetslöshet (13)	0	30,573	7	6
Studiestöd (15)	0	20,981	9	0
Statsskuldsräntor m.m. (26)	0	81,81	3	0
Avgiften till europeiska gemenskapen (27)	0	23,083	6	0
Totalt	119	699,302	2103	156

Tabell B.4 Antal avslutade utvärderingar, utgifter, antal sidor i statsliggaren samt antal artiklar på DN-debatt, i procent av samtliga , per utgiftsområde år 2000

Utgiftsområde (Nr)	Utv.	Miljard. Kr.	Sid. I statsligg.	DN-debatt
Utbildning och universitetsforskning (16)	25%	5%	13%	12%
Hälsovård, sjukvård och social omsorg (9)	11%	4%	5%	10%
Kultur, medier, trossamfund och fritid (17)	11%	1%	13%	1%
Arbetsmarknad och arbetsliv (14)	10%	7%	4%	14%
Regionalutjämning och utveckling (19)	9%	0%	2%	1%
Samhällsekonomi och finansförvaltning (2)	8%	0%	5%	2%
Näringsliv (24)	4%	0%	8%	6%
Jord- och skogsbruk, fiske med anslutande näringar (23)	3%	1%	5%	1%
Rättsväsendet (4)	3%	3%	4%	9%
Internationellt bistånd (7)	3%	2%	3%	3%
Totalförsvaret (6)	3%	7%	11%	3%
Allmänna bidrag till kommunerna (25)	2%	14%	1%	4%
Skatteförvaltning och uppbörd (3)	1%	1%	1%	1%
Ekonomisk trygghet vid ålderdom (11)	1%	5%	0%	4%
Ekonomisk trygghet för familjer och barn (12)	1%	6%	0%	2%
Samhällsplanering, bostadsförsörjning och byggande (18)	1%	2%	4%	3%
Allmän miljö- och naturvård (20)	1%	0%	3%	3%
Energi (21)	1%	0%	2%	0%
Kommunikationer (22)	1%	4%	5%	1%
Rikets styrelse (1)	1%	1%	3%	4%
Finansplan	1%	0%	0%	0%
Ekonomisk trygghet vid sjukdom och handikapp (10)	1%	13%	1%	1%
Utrikesförvaltning och internationell samverkan (5)	0%	0%	2%	11%
Invandrare och flyktingar (8)	0%	1%	2%	3%
Ekonomisk trygghet vid arbetslöshet (13)	0%	4%	0%	4%
Studiestöd (15)	0%	3%	0%	0%
Statsskuldräntor m.m. (26)	0%	12%	0%	0%
Avgiften till europeiska gemenskapen (27)	0%	3%	0%	0%
Totalt	100%	100%	100%	100%