

C



FINANSDEPARTEMENTET

Kan myndigheter utvärdera sig själva?

Rapport till

ESO

Expertgruppen för studier
i offentlig ekonomi

Ds 1996:36



FINANSDEPARTEMENTET

KAN MYNDIGHETER UTVÄRDERA SIG SJÄLVA?

96. 00 14

EX C

Rapport till expertgruppen för
studier i offentlig ekonomi

Av Catharina Barkman
och Stefan Fölster

Ds 1996:36

PARTEMENTET

SOU och Ds kan köpas från Fritzes kundtjänst. För remissutsändningar av SOU och D svarar Fritzes, Offentliga Publikationer, på uppdrag av Regeringskansliets förvaltningsk

Beställningsadress: Fritzes kundtjänst
106 47 Stockholm
Fax: 08-20 50 21
Telefon: 08-690 91 90

Svara på remiss. Hur och Varför. Statsrådsberedningen, 1993.
- En liten broschyr som underlättar arbetet för den som skall svara på remiss.

Broschyren kan beställas hos:
Regeringskansliets förvaltningskontor
Distributionscentralen
103 33 Stockholm
Fax: 08-405 10 10
Telefon: 08-405 10 25

Förord

Enligt budgetordningen skall myndigheternas verksamhet utvärderas löpande; detta såsom ett självklart inslag i den budgetprocess som skall fastställa anslagen till verksamheten i fråga. Myndigheterna har själva en viktig uppgift i denna utvärdering; deras "självutvärdering" skall redovisa vad de presterat och vilka mätbara resultat verksamheten åstadkommit.

Någon systematisk granskning av dessa utvärderingar har dock hittills inte ägt rum. Vi har således inte vetat huruvida myndigheternas självutvärderingar verkligen förmått tillföra budgetprocessen den kunskap som förutsatts.

I denna rapport presenterar Catharina Barkman, RRV, och Stefan Fölster, Industriens Utredningsinstitut, en kartläggning och kvalitetsgranskning av myndigheternas självutvärderingar. Ett antal fristående experter – företrädesvis forskare eller utredare med egen erfarenhet av fältet – har genomfört granskningen av ett 20-tal myndigheters självutvärderingar. Resultatet är nedslående:

- Det finns många myndigheter som inte har genomfört några självutvärderingar alls.
- Bland de genomförda självutvärderingarna är kvaliteten ofta låg; det gäller såväl metodmässigt som resultatmässigt.
- Slående är att självutvärderingarna ofta helt saknar ett tillförlitligt resonemang om ändamålsenlighet och kostnadseffektivitet i den granskade verksamheten.

Visst finns några lysande undantag, och man kan också diskutera representativiteten i urvalet av självundersökningar i denna rapport. Bark-

mans och Fölsters slutsats är likväl att "de flesta av självvärderingarna är av så dålig kvalitet att de faktiskt inte bör användas som beslutsunderlag". Därmed saknas ett centralt inslag i den nya budgetprocessen – enligt Barkman/Fölster själva "motorn" i processen. Följaktligen är det svårt att hävda att resultatstyrningen fungerar som avsett.

Denna rapport är således såväl viktig som kontroversiell. Den reser en rad frågor rörande budgetprocessen, utvärderingsmekanismen och resultatstyrningen. Den reser också frågor om hur utvärderingsprocessen skall kunna förbättras och kvaliteten på självvärderingarna höjas. Författarna avslutar sin rapport med ett antal förslag om hur detta skall kunna åstadkommas.

Det är vår förhoppning att rapporten därmed skall bidra till en diskussion om hur budgetprocessen skall kunna bli mer stringent. Som vanligt i ESO-sammanhang är analysen och resultaten författarna egna.

Mörksund i juni 1996

Klas Eklund
Ordförande för ESO

Innehållsförteckning

1	Sammanfattning	7
2	Inledning	11
3	Självutvärderingarna och den nya budgetprocessen	15
4	Kartläggning av statliga myndigheters själv- utvärderingar	23
4.1	Vad är en självutvärdering?	23
4.2	Myndighetsurvalet	25
4.3	Självutvärderingarna – några generella drag	29
4.3.1	AMS	31
4.3.2	Banverket	32
4.3.3	Datainspektionen	32
4.3.4	Glesbyggsverket	33
4.3.5	Invandrarverket	33
4.3.6	Jordbruksverket	33
4.3.7	Konsumentverket	33
4.3.8	Kriminalvårdsstyrelsen	34
4.3.9	Livsmedelsverket	34
4.3.10	Naturvårdsverket	34
4.3.11	Närings- och teknikutvecklingsverket	35
4.3.12	Patent- och registreringsverket	35
4.3.13	Riksförsäkringsverket	36
4.3.14	Rikspolisstyrelsen	36
4.3.15	Riksskatteverket	37

4.3.16	Riksåklagaren	31
4.3.17	SIDA	31
4.3.18	Skogsstyrelsen	31
4.3.19	Skolverket	31
4.3.20	Socialstyrelsen	31
4.3.21	Tullverket	31
4.3.22	Vägverket	40
4.4	Några slutsatser	40
5	Metod för kvalitetsbedömning	41
5.1	Hur har självutvärderingarna valts ut?	42
5.2	Hur har de sakkunniga valts ut?	44
5.3	Vad skall de sakkunniga jämföra med?	46
6	Kvaliteten i statliga myndigheters självutvärderingar	49
6.1	Helhetsbilden	49
6.2	Dokumentation, data och metod	57
6.3	Slutsatser, kostnadseffektivitet och objektivitet	63
6.4	Några goda exempel på självutvärderingar	68
7	Slutsatser och rekommendationer	71
	Referenser	77
	Abstract	79
	<i>Bilagor:</i>	
Bilaga 1	Termen självutvärdering och projektets uppläggning	83
Bilaga 2	Instruktioner till sakkunnigbedömning	87
Bilaga 3	Förteckning över utvärderingar som bedömts av sakkunniga	93
Bilaga 4	Förteckning över de sakkunniga	95
Bilaga 5	Sakkunnigbedömningar	97
	Förteckning över ESO:s tidigare publicerade rapporter	317

1 Sammanfattning

Beslutet att införa den nya statliga budgetprocessen sedan 1988 har ibland kallats för "den största förvaltningsreformen sedan Axel Oxenstierna". Den skulle ge regering och riksdag möjlighet att styra de statliga myndigheterna på basis av kunskap om faktiskt uppnådda resultat. Tanken var att myndigheterna själva skulle utvärdera och redovisa vad de presterat och vilka mätbara effekter man åstadkommit.

Under ett antal år har myndigheterna nu presenterat sina resultat i årsredovisningar, fördjupade anslagsframställningar samt i separata utvärderingsrapporter. En fundamental fråga i detta sammanhang är om myndigheternas utvärderingar av den egna verksamheten *kan* användas. I vilken omfattning genomför myndigheter utvärderingar som verkligen belyser relevanta frågor om kärnverksamheten? Är utvärderingarna pålitliga och objektiva? Är kvaliteten i metodval och framställning godtagbar?

I denna rapport redovisas en kartläggning av 22 myndigheters självutvärderingar, dvs. utvärderingar av den egna verksamheten. Omfattningen och i viss mån karaktären av myndigheters självutvärderingsarbete under åren 1990–1995 bygger på RRV:s definition av begreppet utvärdering. Denna definition utgår från att utvärderingar bl.a. "klargör orsaksförhållanden när det gäller myndigheternas prestationer", vilket exkluderar en del enklare beskrivningar av verksamhetens utveckling som av myndigheterna själva ibland betraktas som självutvärderingar. Hälften av de utvalda myndigheterna tillhör de största, hälften är ett slumpmässigt urval av mindre och medelstora myndigheter.

Som ett led i arbetet har ett antal fristående sakkunniga personer anlåtats för att bedöma kvaliteten i ett tvärsnitt av självutvärderingarna. Totalt har 23 utvärderingar granskats. De utvalda utvärderingarna är i huvudsak sådana som myndigheterna själva framhåller bland sina bästa eller metodmässigt mest sofistikerade. För varje utvärdering har vi låtit två fristående och mot varandra anonyma sakkunniga göra en bedömning. Sammanlagt har 40 sakkunniga anlåtats. De sakkunniga är forskare från olika discipliner och utredare med mångårig erfarenhet från utvärderingar. Myndigheterna har själva deltagit i urvalet av sakkunniga, och har kunnat välja bort sakkunniga.

Kartläggningen visar att av 22 myndigheter har 10 över en femårsperiod gjort färre än 1 självutvärdering per 1 000 anställda. Sex av dessa, eller en dryg fjärdedel, har inte gjort några självutvärderingar alls. Bland de övriga 14 dominerar attitydundersökningar och enkla bedömningar av olika delverksamheter eller organisatoriska förändringar. Relativt sällan har myndigheterna tagit sig an uppgiften att mäta effekter eller effektiviteten som uppnåtts. Det bör här särskilt understrykas att analysen i rapporten vare sig handlar om hur väl en myndighet fungerar eller om att ge en bild av myndigheternas totala utvärderingsverksamhet. Rapporten behandlar enbart självutvärderingsarbetet.

När det gäller kvaliteten på utvärderingarna har de sakkunniga i allmänhet lämnat mycket detaljerade motiveringar till sina bedömningar. En majoritet av de sakkunniga är mycket kritiska till kvaliteten i fråga om dokumentation, metodval, frågeställning m.m. I flertalet fall anser de att slutsatser dras som det inte finns stöd för i det empiriska materialet och att det hade gått att göra en väsentligt bättre utvärdering med samma arbetsinsats. Ofta har de sakkunniga anmärkt på att utvärderingar med tystnad förbigått de viktigaste frågorna angående ända målsenlighet och kostnadseffektivitet i den utvärderade verksamheten.

Några fåtal goda exempel eller pärlor finns som har fått höga betyg. Dessa utvärderingar kännetecknas av att fokus ligger på en central frågeställning, att vetenskapligt acceptabla metoder kommit till användning och att man endast dragit sådana slutsatser som materialet tillåter.

Slutsatsen är att de flesta självutvärderingarna är av så dålig kvalitet att de faktiskt inte bör användas som beslutsunderlag. I den mening som kan sägas att budgetprocessens centrala ambition att möjliggöra resultatstyrning ännu inte kommit att uppnås, även om den i flera avseer

den varit framgångsrik. Utan en fungerande resultatmätning av godtagbar kvalitet saknas den motor i den nya budgetprocessen som krävs för att resultatstyrning skall kunna bli en realitet.

För att förverkliga budgetprocessens ambitioner krävs sannolikt en betydande uppstramning av självutvärderingsprocessen. Ett avgörande inslag i detta bör vara en kvalitetssäkring. Till en del måste ansvaret för kvalitetssäkringen självfallet ligga hos de berörda myndigheterna. Men till en del bör även en oberoende instans som ligger utanför myndigheterna utöva en kvalitetssäkrande roll. Det förefaller t.ex. rimligt att RRV i högre grad än tidigare kommenterar och granskar kvaliteten i myndigheters självutvärderingar.

Som metod för kvalitetssäkring av utvärderingar förefaller en vägvida överlägsen och även mest använd i olika utvärderingssammanhang. Det är samma metod som vi också har använt för denna rapport. Oberoende och anonyma (om de så önskar) sakkunniga anlitas för att bedöma utvärderingarna. Av särskild vikt är att dessa sakkunniga väljs ut av andra än av den som genomfört utvärderingen, den som utvärderas, eller närstående till dessa. Vi föreslår därför att någon instans ges uppgiften att agera som "Kvalitetssäkringsgrupp" med uppgift främst att förmedla sakkunniga. På begäran skulle Kvalitetssäkringsgruppen välja sakkunniga som passar till en viss självutvärdering, och återföra de sakkunnigas bedömning till de berörda, under bevarandet av de sakkunnigas anonymitet om de så önskar. Kvalitetssäkringsgruppen skulle kunna anlitas av myndigheter som utför självutvärderingar. Detta kan samtidigt fungera som en kvalitetsstämpel som ger läsaren en signal om pålitligheten. Men Kvalitetssäkringsgruppen skulle också kunna anlitas av t.ex. regeringskansliet och RRV i granskningen av självutvärderingars kvalitet. Det är tänkbart, och även önskvärt, att myndigheter skaffar sig en kvalitetsstämpel på sina självutvärderingar som ett sätt att förebygga den kritik som kan komma om någon annan i stället anlitar Kvalitetssäkringsgruppen. Självfallet är det också bra om myndigheter använder sig av kvalitetssäkringen i ett tidigt skede under en självutvärderings framställning och lyckas arbeta in de sakkunnigas synpunkter.

En sådan kvalitetssäkringsprocess fungerar sannolikt också som katalysator som ökar utvärderingskunnandet inom myndigheterna. Våra resultat indikerar att en förbättring av självutvärderingar handlar om kunnande, inte nödvändigtvis om högre utvärderingskostnader. En så-

dan kompetens är en förutsättning, inte bara för att utföra egna utvärderingar, utan även för att beställa och kontrollera utvärderingar gjorda utanför myndigheten av t.ex. utomstående konsulter.

Eftersom ökade kvalitetssäkringskrav kan medföra att myndigheterna gör färre självutvärderingar är det viktigt att kompletterande utvärderingar kan beställas av t.ex. revisionsmyndigheterna.

Först när kvaliteten i myndigheters självutvärderingar har förbättrats är det meningsfullt att återigen se på frågan om användningen av utvärderingarna. Det är mycket möjligt att en förbättrad kvalitet och tydligare information om enskilda utvärderingars kvalitet ökar användbarheten och därmed även användningen.

2 Inledning

Bakgrund

År 1988 fattade regering och riksdag beslut om en ny s.k. budgetprocess. Denna syftade till att låta regering och riksdag styra myndigheter utifrån kunskap om myndigheternas prestationer och effektivitet.

En viktig hörnsten i den nya budgetprocessen är att de statliga myndigheternas resultat skall mätas. I förordningen (1993:134) om myndigheters årsredovisning och anslagsframställan ställs t.ex. krav på att myndigheters resultatredovisning skall innehålla vad myndigheten "har presterat och mätbara effekter av prestationerna". Ansvar för resultatmätningen lades på de berörda myndigheterna själva.

Många myndigheter har som ett verksamhetsområde, ansvar för utvärderingar inom olika områden, dvs. utvärderingar som inte bedömer just den egna verksamheten. Utvärdering av den egna verksamheten, eller – som vi valt att kalla det i denna rapport – självutvärdering, utgör i normalfallet endast en bråkdel av myndigheters utvärderingsverksamhet. En bråkdel som dock är betydelsefull eftersom den utgör det fundament som regeringens styrning av de statliga myndigheterna vilar på.

Att myndigheter skall utvärdera sig själva har aldrig varit helt okontroversiellt. Trots detta har diskussionen kring hur självutvärderingar fungerar upptagit ett mycket begränsat utrymme, både i de utredningar som ledde fram till den nya budgetprocessen, och i de studier av budgetprocessen som har gjorts därefter. Mycket av arbetet har i stället fokuserat på omläggningen av de formella rutinerna. Uppföljningar som gjorts i efterhand har i korthet konstaterat att resultatänkandet hos myndigheter visserligen har ökat, men att regering och riksdag in-

te utnyttjar självvärderingarna. Även i efterhand har diskussionen till stor del gällt de formella rutinerna samt regeringskansliets förmåga

Syfte

Denna studie utgår från en mer grundläggande fråga. Kan myndigheterna utvärdera sin egen verksamhet? Hur många självvärderingar görs egentligen? Hur mycket kostar utvärderingsarbetet? Undersök de rätt saker? Är självvärderingarna pålitliga? Kan potentiella användare bedöma tillförlitligheten?

Det har inte funnits någon samlad bild över i vilken utsträckning myndigheter utvärderar sin verksamhet eller om utvärderingarna håller tillräckligt hög kvalitet för att resultaten skall vara användbara. En del indikationer har dock funnits på att det kan finnas problem med myndigheters självvärderingar. RRV:s granskning av NUTEK:s stöd till företag och forskning (RRV 1995:39) innehöll t.ex. kritik mot självvärderingar som verket gjort.¹

En mer impressionistisk observation är att forskarvärlden till stor del förbigår myndigheters självvärderingar med tystnad. Man bryr sig inte om att kritisera dem, men man refererar inte heller till dem.

Om det är så att myndigheters självvärderingar ofta är behäftad med betydande kvalitetsbrister, kan det innebära att icke-användningen av självvärderingar är en helt riktig reaktion. Att i ett sådant läge försöka tvinga fram en ökad användning av självvärderingarna skulle enbart öka risken att beslut fattas på felaktiga grunder. Därmed skulle också den nya budgetprocessens legitimitet undergrävas. I den meningen kan sägas att en granskning av självvärderingarnas kvalitet är en förutsättning för att framgångsrikt kunna ändra på användningen av myndigheters resultatmätningar, och därmed även för att den nya budgetprocessen skall kunna fungera som tänkt.

¹ En av slutsatserna är att effektivitetsfrågorna är försummade och att bristfällig utvärderingsmetodik ibland använts.

Några definitioner

I denna rapport har vi valt att utgå från den definition av begreppet utvärdering som används av t.ex. RRV och Statskontoret. Enligt den definitionen skall en utvärdering bidra till kunskap om en åtgärds ändamålsenlighet och kostnadseffektivitet. Flera myndigheter har som en verksamhetsuppgift att göra utvärderingar inom olika områden, dvs. utvärderingar som inte bedömer just den egna verksamheten. För att påminna om att det handlar om myndigheternas utvärderingar av den *egna* verksamheten använder vi benämningen "*självutvärderingar*" (se även bilaga 1). Det spelar emellertid ingen roll om myndigheterna själva har gjort utvärderingarna eller om konsulter, anlitate av myndigheterna, har utfört utvärderingarna.

Inriktning och avgränsning

Rapporten bygger på en kartläggning av 22 myndigheters självutvärderingar under perioden 1990–1995. I detta urval ingår många av de största myndigheterna, mätt i antalet anställda. Dessutom ingår ett slumpmässigt urval av små och medelstora myndigheter. För varje myndighet har vi samlat in självutvärderingar eller listor på självutvärderingar. De flesta myndigheterna har hjälpt till med att ta fram detta underlag. Försvaret, universitet och högskolor – de s.k. stabsmyndigheterna samt statliga bolag och stiftelser ingår däremot inte.

Att granska kvaliteten i utvärderingar är inte lätt. En del skulle kanske hävda att utvärderingar har likheter med konstverk, dvs. det kan finnas många legitima uppfattningar om kvaliteten. Därtill kommer frågan om bedömarens objektivitet. Redan av den anledningen skulle inte vi, med någon trovärdighet, kunna utgå bara från våra egna bedömningar.

En grundsyn som de flesta utvärderare, liksom forskare, delar är dock att det finns vissa grundläggande kriterier som en utvärdering bör uppfylla för att räknas som trovärdig. På flera områden har dessa kriterier formulerats explicit (se t.ex. Sanders 1994), mer eller mindre omfattande. Kriterierna gäller bl.a. att dokumentation, metodval, data och slutsatser uppfyller vissa minimikrav. Kriterierna blir meningsfulla därför att de flesta initierade personer i många fall gör en likartad

bedömning av huruvida kriterierna är uppfyllda. Däremot kan viss personer i enskilda fall göra avvikande bedömningar, och för vissa utvärderingar kan det uppstå en större spridning av bedömningar.

Därför har vi för vår kvalitetsgranskning valt ut ett stort antal självutvärderingar och flera sakkunniga som bedömt självutvärderingarna. Urvalet av självutvärderingar bygger i huvudsak på myndigheters förslag, och myndigheterna har även haft visst inflytande över valet av sakkunniga. En närmare diskussion av kvalitetsfrågan och vårt tillvägagångssätt förs i kapitel 5 (se även bilaga 1).

Vi för ingen diskussion om metodval i de olika utvärderingarna. Men de sakkunniga diskuterar i allmänhet metodfrågorna mycket grundligt och förtjänstfullt. Samtliga bedömningar finns i bilaga 5. Vi dare finns i bilaga 4 en förteckning över de sakkunniga som granska rapporterna och som inte har begärt att få vara anonyma.

Vi har inte gjort någon undersökning av i vilken mån självutvärderingar används internt eller av regeringskansli och riksdag.² Det är givetvis en intressant fråga i sig. Användningen av självutvärderingar är dock avhängig kvaliteten och relevansen vilket vi studerar i denna rapport.

² Exempel på studier om detta är Statskontoret, 1992:32, Att använda resultatmått med gott resultat.

3 Självtvärderingarna och den nya budgetprocessen

Under det senaste decenniet har några statliga verk och myndigheter bantats, lagts ner, bolagiserats eller privatiserats. Andra myndigheter har varit föremål för omstruktureringar. De flesta arbetar emellertid vidare i en oförändrad myndighetsform medan deras styrning har ändrats framför allt genom den nya statliga budgetprocessen.

Den nya budgetprocessen infördes successivt från 1988.³ Den innebär i korthet att riksdag och regering skall besluta om inriktning av verksamheten genom att formulera mål. Samtidigt får myndigheterna friare händer att välja medel för att uppnå dessa mål. Resultaten skall följas upp och utvärderas. Tanken är att bristande resultatuppfyllelse skall leda till någon form av reaktion och utkrävande av ansvar.

Genom den förändrade budgetprocessen har myndigheterna fått ett betydligt större ansvar än tidigare för uppföljning och utvärdering av den egna verksamheten och de regelsystem man administrerar. I förordningen (1993:134) om myndigheters årsredovisning och anslagsframställan formaliseras kraven på att myndigheterna skall redovisa bland annat "prestationer och mätbara effekter av prestationerna".

Betydelsen av resultatmätning har också lyfts fram i senare uppföljningar av den nya budgetprocessen. I en rapport om åtgärder för att stärka resultatstyrningen (ELMA-gruppen Ds 1995:6) fastslås t.ex. att myndigheternas information bör vara den viktigaste basen för rege-

på olika nivåer i beslutskedjan... Utan sådana utvärderingar går det inte att göra medvetna prioriteringar."

Principen att myndigheterna skall utvärdera sig själva har varit kontroversiell. De argument som talar för självutvärdering bygger på tanken att dessa inte bara fyller en *utåt* legitimerande funktion, utan även en viktig intern funktion. Myndigheter måste kunna utvärdera löpande som ett rutinemässigt led i styrningen av den egna verksamheten. Det är därför närmast en självklarhet att utvärdering av funktionen inte uteslutande kan vara en uppgift för externa organisationer.

En väl fungerande intern utvärderingsfunktion ger också rent praktiskt ett försprång t.ex. i termer av detaljkunskap vid genomförandet av utvärderingar. I bästa fall kan således myndigheten själv utföra utvärderingar billigare och bättre än någon extern utvärderare.

Kritiken av självutvärdering har ifrågasatt om myndigheter alltid har kompetens för att producera goda utvärderingar. Även beställning av utvärderingar kräver en hel del kompetens för att kunna välja en kunnig utförare och för att kunna formulera ett genomtänkt uppdrag. En annan viktig fråga är om myndigheter alltid har den oberoende objektivitet som krävs för att göra trovärdiga självutvärderingar.

Den formella styrningsprocessen

Mycket av det förberedande arbetet inför den nya budgetprocessen och uppföljningsarbetet har handlat om de formella procedurerna i budgetprocessen.

Åtminstone vart 6:e år gör regering och riksdag en fördjupad prövning av myndigheternas verksamhet. Myndigheterna förser regeringen med en fördjupad anslagsframställning som bl.a. skall ge svar på om de mål som uppställts för myndigheternas verksamhet har uppnåtts och vilka effekter de olika verksamheterna gett.

I förordningen (1993:134) om myndigheters årsredovisning och anslagsframställan finns allmänna föreskrifter om vilken information en fördjupad anslagsframställning skall innehålla. Ofta finns det emellertid skäl för regeringen att precisera kraven på den information myndigheten skall lämna. En sådan precisering görs med hjälp av myndighetsspecifika direktiv. I dessa begär regeringen in uppgifter som ä

viktiga för att kunna fatta beslut om myndighetens verksamhet under nästa treårsperiod.

I förhållande till gällande förordning och föreskrifter innehåller direktiven ofta skärpta krav på resultatanalys inom områden som är under omprövning. På andra, mindre angelägna områden kan i stället direktiven lätta på kraven på information. När myndigheterna lämnat sina fördjupade anslagsframställningar granskas dessa i regeringskansliet.

Myndigheterna skall också varje år redovisa sina resultat för regeringen. I dessa årliga resultatredovisningar ges en fortlöpande återkoppling till verksamhetens resultat, kostnader, avvikelser från planer och ändrade förutsättningar. Årsredovisningarna innehåller förutom resultatredovisning även balans- och resultaträkning, anslagsredovisning och finansieringsanalys.

Årsredovisningen har utvecklats till ett av resultatstyrningens viktigaste dokument samtidigt som anslagsframställningens betydelse minskat. Ett annat styrningsdokument i budgetprocessen som har fått ökad vikt är regleringsbrevet som sedan 1991 innehåller både en verksamhetsdel och en finansieringsdel. I verksamhetsdelen anges mål och krav på resultat i verksamheten.

ELMA-gruppen diskuterar de olika styrningsdokumenten och menar att både dokumenten och själva styrningen i högre grad än tidigare bör verksamhetsanpassas. Det framhålls också att de myndighetsspecifika direktiven och de fördjupade anslagsframställningarna inte passar in i sin nuvarande form i resultatstyrningen och de centrala kraven på dessa dokument bör därför tas bort.

Erfarenheter av den nya budgetprocessen

Flera uppföljningar av den nya budgetprocessen konstaterar ett ökat resultattänkande både hos myndigheterna och inom regeringskansliet.⁴ Inom flera myndigheter har införandet av resultatstyrning följts av interna förändringar. ELMA-gruppen skriver bl.a. att myndigheterna har

⁴ För en sammanfattning av studier se Barkman (1993), kapitel 9 i "Sveriges systemskifte i fara?", IUI.

varit aktivt kunskapssökande och engagerat sig i reformprocesser. Myndigheternas anslagsframställningar har i det långa perspektivet blivit mycket bättre.

En studie där RRV har jämfört anslagsframställningar över en 25 årsperiod visar att det har skett stora förbättringar. Anslagsframställningarna har utvecklats från att vara "synnerligen detaljerade tiggare brev" till att bli "analytiska produkter" med högt informationsvärde. Myndigheternas resultatanalyser har utvecklats.

Samtidigt har en rad studier av den nya budgetprocessen emellertid lyft fram en del problem som den nya budgetprocessen inte lyckades lösa i enlighet med förhoppningarna. Riksdagens revisorer har gjort en studie, som bl.a. visade att myndigheternas resultatredovisningar inte gav det underlag för omprövningar av verksamheten som förutsattes. Underlag hämtades i stället från fristående utredningar. Viktiga beslut om omprövning av statlig verksamhet fattades helt eller delvis vid sidan av budgetprocessen.

Liknande synpunkter finns också i en uppföljning av den s.k. ELMA-gruppen (Ds 1995:6). I ELMA-gruppens rapport konstateras bl.a. att de nya "styrformerna inte lett till ett bättre underlag för departementens omprövningar av verksamheter genom den information som lämnas av myndigheterna". Gruppen menar också att större tonvikt bör läggas på redovisning av effekter och uppnådda resultat.

Bristande användning av resultatmått

Flera tidigare undersökningar har konstaterat att regeringskansliet inte har klarat av att hantera den informationsmängd som kom från myndigheterna, och därmed inte heller att agera och reagera på basis av myndigheternas redovisningar (t.ex. Arbetstagarkonsult, 1992; RRV, 1991; Statskontoret, 1990:24).

Dessa undersökningar har några år på nacken och vissa åtgärder för att förbättra regeringskansliet med avseende på resultatstyrning har sedan dess initierats, bland annat kompetensutveckling. Även ELMA-gruppen (Ds 1995:6) betonar emellertid vikten av att reagera på missförhållanden och att utkräva ansvar. RRV har en viktig roll att granska effektivitet och resultat i statlig verksamhet. "Så måste t.ex. revisionens anmärkningar leda till konkreta åtgärder från regeringen. Så-

dana åtgärder kan vara att ompröva myndighetsledningens uppdrag eller att återta delar av den omfattande delegation som givits myndigheterna." ELMA-gruppen fastslår också att det fortfarande finns brister främst vad gäller att formulera övergripande mål och verksamhetsmål.

Inom ramen för denna rapport har samtal förts med utvärderingsansvariga eller motsvarande på ett stort antal myndigheter. Det har då framkommit att det fortfarande finns en betydande skepsis mot att departementen är effektiva i sin roll som beställare och mottagare av resultat.

Flera myndighetsrepresentanter har också framfört att de upplever att de får olika signaler från skilda instanser. Departementen ger uttryck för en viss inriktning i regleringsbrevet men anger en annan muntligen, mer informellt. RRV å andra sidan måste och bör i sin granskning av årsredovisningar (där även effekter skall redovisas) ha budgetförordningen (1993:134) och regleringsbrevet som utgångspunkt. Myndighetsrepresentanter menar att detta kan skapa viss förvirring.

Ett annat problem som myndigheterna gett uttryck för är begreppsapparaten. Vissa myndigheter upplever att departementen använder andra definitioner, eller för allmänt hållna definitioner, på begrepp som t.ex. "effekter" än myndigheterna själva.

Varför används inte självutvärderingarna?

Mycket av diskussionen kring den nya budgetprocessen har handlat om vad som kan ändras inom regeringskansliet för att öka användningen. Det s.k. "non-utilization" problemet har också diskuterats mycket i utvärderingslitteraturen (t.ex. Tarschys 1982). Följande orsaker till att utvärderingar inte används betonas:

1. Politiker har svårt att bedöma trovärdigheten av olika utvärderingar.
2. Politikens handlingar styrs inte enbart av effektivitetstänkande utan även av hänsyn till särintressen, den allmänna opinionen, förhandlingsspelet mellan partierna och inte minst det pressade tidsschemat.

3. Även en väl genomförd resultatmätning ger ibland inga entydiga signaler om vilka åtgärder som bör vidtas.

I den omfattande utvärderingsforskningen formuleras slutsatser om hur utvärderingar bör se ut för att öka chansen att de används (t.ex. Seidel 1983). Förslagen handlar ofta om att öka utvärderingarnas läslighet och att förankra de hos de berörda parter.

Inriktningen på denna rapport är emellertid att undersöka i vilken mån självutvärderingar är användbara. En fjärde tänkbar förklaring till icke-användning av resultatmätningar, som således står i centrum för vår rapport, är att många självutvärderingar är behäftade med olika problem som gör dem oanvändbara, eller att de mäter saker som inte är relevanta för de politiska besluten.

I den pedagogiska kampanj som bedrevs under införandet av budgetprocessen användes ett mycket begränsat antal resultatmått som exempel. Diskussionen om vilka resultatmått som verkligen var meningsfulla sköts på framtiden. Kvalitativa skillnader mellan olika resultatmått tonades också ner.

De statliga myndigheterna har haft betydande frihet att utforma resultatmått själva. Detta har gett upphov till misstanken att en del verksamheter medvetet valt sådana resultatmått som har begränsat deras användbarhet för resultatstyrning. I utvärderingsforskningen framhävs att så kallade "eye wash"-utvärderingar är mycket vanliga. Även ovana och låg kompetens avseende resultatmätning kan ha bidragit till val av mindre användbara resultatmått.

Resultatmåttens användbarhet måste bedömas med avseende på i vilken utsträckning de kan förbättra beslutsfattandet. Användbarheten ökar ju bättre måttet lyckas ge tydliga signaler om lämpliga åtgärder.

Resultatmått kan i det avseendet uppdelas i tre kategorier vars användbarhet kan rangordnas.

1. *Verksamhetsutveckling*: Information till huvudmännen om kostnadsutveckling, verksamhetsutveckling, utveckling av enkla produktivitetstal (typ kostnad per ärende) och beskrivning av omvärldsförändringar.
2. *Effekter av olika åtgärder*: Information om att en åtgärd eller verksamhet har lett till vissa effekter, t.ex. förbättrad produktivitet, ändrat beteende hos allmänheten, m.m.

3. *Kostnadseffektivitet*: Information om verksamhetens effekter inkl. kvalitet i förhållande till kostnader.
4. *Relativ kostnadseffektivitet*: Information om kostnadseffektivitet enligt punkt 3 i förhållande till vad som är möjligt eller vad andra uppnår.

Generellt kan sägas att det alltid är möjligt att genomföra resultatmätningar i kategorin "verksamhetsutveckling". Kostnaderna är kända eller kan beräknas med förbättrade redovisningsmetoder. Mått på verksamhetens volym som t.ex. antal ärenden, antal elever och dylikt kan också alltid beräknas.

Även produktivetsmått som mäter volym/kostnader tillhör kategorin verksamhetsutveckling. Som ett exempel har produktivitet i skolväsendet mätts som kostnader per elev. Ett sådant mått kan emellertid inte skilja mellan förändringar i effektivitet, förändringar i kvalitet (t.ex. fler skolpsykologer) eller förändringar i omvärlden som påverkar kostnaderna. När ett produktivetsmått är det enda möjliga resultatmättet kan det vara värdefullt för en viss typ av styrning där pengar fördelas i förhållande till verksamhetens omfattning, t.ex. per elev. Det är emellertid tveksamt om detta kan kallas för resultatstyrning eftersom skillnader i kvalitet inte tas med i beräkning.

Den andra kategorin av resultatmått ger information om effekter, dvs. en mätning av hur vissa kvalitativa eller kvantitativa aspekter ändras till följd av åtgärder eller förändringar i verksamheten. Om relevanta effekter har valts, kan detta användas för resultatstyrning, t.ex. för beslut om åtgärders lämplighet. Effektmätningen ligger också till grund för beräkning av kostnadseffektivitet. Här krävs en mer detaljerad mätning av kvalitetsförändringar.⁵ Om kvalitetsmätningen lyckas ger ett sådant mått bättre information än mått på enbart verksamhetsutveckling. Kostnadseffektivitet kan i vissa situationer ge entydiga signaler för resultatstyrning. Det kan t.ex. visa sig att en verksamhet ger så obetydliga effekter att den bör läggas ned. Ibland kan det också vara möjligt att hänföra ändringar i kostnadseffektiviteten till dess orsaker som sedan kan åtgärdas. I många fall ger kostnadseffekti-

⁵ I fall där outputmättet tar hänsyn till kvaliteten är kostnadseffektivitet detsamma som den totala faktorproduktiviteten.

vitet emellertid inte heller riktigt entydiga signaler som kan användas i resultatstyrningen därför att det förblir okänt hur hög kostnadseffektivitet som skulle kunna uppnås med t.ex. en bättre ledning eller en mer förnuftig arbetsmetod.

Därför är ett ännu bättre resultatmått ett som ställer prestationsrelation till vad som är möjligt att prestera, eller vad andra presterar eller vad som presteras med olika arbetsmetoder inom samma organisation. Detta fångas upp av den relativa kostnadseffektiviteten. Sådana resultatmått förutsätter inte bara att kvaliteten kan mätas utan även att det finns jämförbara verksamheter. De ger också alltid ett entydigt signal till resultatstyrningen, nämligen att flytta över verksamhet till den som har den bästa kostnadseffektiviteten. Därmed är måttet på relativ kostnadseffektivitet klart mest användbart för resultatstyrning.

Denna rangordning av resultatmåttens användbarhet är inte särskilt stridd. I en mer systematisk undersökning har t.ex. Statskontoret (1990:24) visat att offentliga verksamheter sällan använder sig av resultatmått som mäter effektiviteten. I Statskontorets rapport visas också att resultatmått i den offentliga verksamheten sällan används i åtgärder. Resultatmått används främst som allmän information och bland annat som underlag för allokering av medel. I de företag som undersöktes i Statskontorets studie betonades däremot att resultatmätning av allmänhet omfattade även effektiviteten och att de ledde till åtgärder.

En möjlig orsak till resultatmåttens bristande användning kan därför vara att resultatmått i offentlig verksamhet oftast tillhör kategorin verksamhetsutveckling.

Sammanfattningsvis har av olika utredningar konstaterats att myndigheters självframtagna resultatmätningar inte används mycket. Diskussionen om detta problem har fokuserat på dels frågor om hur utvärderingsresultaten presenteras, dels olika blockeringar hos mottagarna. I denna rapport reses den mer grundläggande frågan om själva utvärderingarna i nuvarande form går att använda för resultatstyrning. Om så inte är fallet, är innebörden att den kanske viktigaste länken i den nya budgetprocessen ännu inte fungerar.

⁶ Detta bekräftas av undersökningen "Bäst i världen" (IVA 1990) som visar också att de mest framgångsrika företagen utvärderar mer och omsätter resultaten i åtgärder än konkurrenterna.

4 Kartläggning av statliga myndigheters självutvärderingar

I detta kapitel redovisas en kartläggning av självutvärderingar i ett urval av myndigheter. Hur många utvärderingar som har genomförts under perioden 1991–1995 vad har utvärderats, vilka metoder har använts osv. Vi inleder med en diskussion om vad en självutvärdering är.

4.1 Vad är en självutvärdering?

Flera myndigheter har enligt sin instruktion uppgiften att göra utvärderingar inom olika områden, dvs. utvärderingar som inte bedömer just den egna verksamheten. För att understryka att det handlar om myndigheternas utvärderingar av den *egna* verksamheten använder vi benämningen "*självutvärderingar*". Utvärderingarna kan vara utförda av myndigheterna själva, eller av konsulter anlitade av myndigheterna. Vi bortser således från utvärderingar som har tillkommit på initiativ av externa aktörer.

Inom utvärderingslitteraturen florerar en rad olika definitioner av "utvärdering". Vedung (1994) t.ex. använder en vid definition: "Utvärdering är en noggrann efterhandsbedömning av utfall, slutprestationer och genomförande i offentlig politik, vilken avses spela en roll i praktiska handlingssituationer".

RRV:s definition av utvärdering inom ramen av den statliga budgetprocessen är att "bedöma och förklara prestationer och händelseförlopp" i en myndighetsverksamhet. Utvärderingen "skärskådar och klar-

gör orsaksförhållanden när det gäller myndighetens prestation och/eller prestationernas effekter i samhället". (RRV 1994: Resultatverksamhet, ekonomi, en handledning för myndigheterna).

En kommitté som bildats i USA (The Joint Committee on Standards for Educational Evaluation) har tagit fram standards eller kriterier (The Program Evaluation Standards) för hur en bra utvärdering ska se ut, både vad avser metodik, användning etc. Dessa standards blivit accepterade i USA som nödvändiga för välgjorda och relevanta utvärderingar. Inom ramen för dess arbete finns en övergripande definition på utvärdering (evaluation) som är "The systematic investigation of the worth or merit of an object".

Visst kan det skapa förvirring med olika definitioner på utvärdering och det är givetvis eftersträvänsvärt med en enhetlig begreppsapparat. Några utvärderingsansvariga inom olika myndigheter som vi diskuterat med inom ramen för denna studie har också anfört att det finns risk för att myndigheterna använder sig av en definition samtidigt som regeringen (fackdepartementen) avser någonting annat än vad spektive myndighets definition rymmer.

Vi har i denna rapport valt att utgå ifrån RRV:s definition. Detta innebär att vi inte tar med enkla mått på verksamhetsutveckling, till exempel redovisning av kostnadsutveckling, antal ärenden och enkla produktivitetsmått. Ofta redovisar myndigheter sådana mått i årsredovisningar där de också fyller en viktig funktion att beskriva hur verksamheten har utvecklats. De uppfyller dock inte RRV:s definition av en utvärdering att "bedöma och förklara...", än mindre övriga definitioner. Enligt diskussionen i det föregående är enkla mått på verksamhetsutveckling inte heller tillräckliga eller i sig särskilt användbara för resultatstyrning.

Som självutvärderingar betraktar vi således utvärderingar enligt RRV:s definition som *avser att bedöma eller mäta effekter av, och utvärdera effektiviteten i, åtgärder och verksamheter som en myndighet ansvarar för*. I definitionen på självutvärderingar lägger vi inga krav på metodik etc. Det kan avse en enkel bedömning, en attitydundersökning eller en mer sofistikerad empirisk undersökning.

4.2 Myndighetsurvalet

Urvalet omfattar totalt 22 statliga myndigheter. Hälften av myndigheterna representerar de största mätt i antalet anställda. Försvaret, högskolor och universitet, stabsmyndigheterna, statliga bolag och stiftelser ingår däremot inte. Övriga 11 myndigheter bygger på ett slumpmässigt urval av små och medelstora myndigheter.

I tabellerna 4.1 och 4.2 redovisas som bakgrund antalet årsarbetskrafter och ungefärligt anslag från regeringen. Vi har dessutom (tabell 4.2) även angett driftskostnad och transfereringar.

Myndigheterna skall i de flesta fall utvärdera den egna verksamheten, de bidrag man hanterar och de regelsystem man administrerar.

Tabell 4.1 Myndighetsurvalet. Årsarbetskrafter och anslag (mkr).

	Antal anställda 1995	Anslag (mkr) 1995/96
AMS	12 000	4 285
Banverket	6 532	0
Datainspektionen	40	22
Glesbygdsverket	17	20
Invandrarverket	2 850	589
Jordbruksverket	900	165
Konsumentverket	180	70
Kriminalvårdsstyrelsen	8 000	4 000
Livsmedelsverket	500	90
Naturvårdsverket	470	0
Närings- och Teknikutvecklingsverket	360	180
Patent- och registreringsverket	760	*
Riksförsäkringsverket	640	824
Rikspolisstyrelsen	27 000	11 000
Riksskatteverket	15 000	6 066
Riksåklagaren	1 370	582
SIDA	506	315
Skogsstyrelsen	1 500	279
Skolverket	250	243
Socialstyrelsen	390	347
Tullverket	3 500	1 200
Vägverket	7 000	13 000

* = s.k. 1 000-kronorsanslag

Källa: Myndighetskalender 1995, RRV. Siffrorna är ungefärliga. Exakta siffror återfinns i myndigheternas årsredovisningar.

Tabell 4.2 Myndigheternas driftskostnader och transfereringar (mkr)

	Drifts- kostnader	Transfe- reringar
AMS	4 285	75 645
Banverket	11 500	0
Datainspektionen	23	0
Glesbygdsverket	20	0
Invandrarverket	611	8 439
Jordbruksverket	650	9 000
Konsumentverket	111	0
Kriminalvårdsstyrelsen	4 000	0
Livsmedelsverket	260	0
Naturvårdsverket	203	105
Närings- och Teknikutvecklingsverket	180	3 000
Patent- och registreringsverket	380	0
Riksförsäkringsverket	924	167 000
Riksskatteverket	6 349	-613 000
Rikspolisstyrelsen	11 020	0
Riksåklagaren	651	0
SIDA	315	16 500
Skogsstyrelsen	1 189	0
Skolverket	246	2 169
Socialstyrelsen	365	3 728
Tullverket	1 200	-90 000
Vägverket	13 600	1 560

Källa: Myndighetskalender 1995, RRV. Siffrorna är ungefärliga. Exakta siffror återfinns i myndigheternas årsredovisningar.

För kartläggningen av myndigheternas självutvärderingar kontakades i de flesta fall den utvärderingsansvarige eller motsvarande på respektive myndighet och ombads att hjälpa till med att sammanställa en lista över självutvärderingarna. Vid ungefär en tredjedel av myndigheterna fanns redan färdiga listor. Flertalet myndigheter har varit mycket hjälpsamma och gjort sammanställningar av självutvärderingarna åt oss.

Detta förfarande är av naturliga skäl beroende av myndigheternas agerande. En risk är att den person vi har haft kontakt med på respektive myndighet inte känt till alla självutvärderingar. Detta kan vara en tänkbar felkälla. Dock kan man ifrågasätta i vilken mån självutvärderingar som inte gjorts kända inom myndigheten, för utvärderingsansvariga eller motsvarande, uppnår syftet med självutvärderingar. För att gardera mot sådana felkällor har vi dessutom brevlades frågat generaldirektören (motsv.) om våra uppgifter uppfattas som

korrekta. Några myndigheter har då kommit med kompletterande material som vi beaktat.

En annan risk kan vara att en myndighet valt att inte redovisa vissa självutvärderingar. I den mån sådant skulle ha förekommit, innebär det att vår sammanställning har med något färre, men förmodligen kvalitativt bättre, självutvärderingar än vad som faktiskt genomförts.

Med dessa reservationer redovisas i det följande *hur många* självutvärderingar respektive myndighet har genomfört, *vad dessa mäter* (produktivitet, effekter, kostnadseffektivitet, eller annat) samt vilken typ av *metod* som använts.

Tabell 4.3 Antal självutvärderingar per myndighet/1 000 anställd, 1990–1995.

	Antal själv- utvärderingar	Antal själv- utvärderingar per 1 000 anställd
AMS	19	1.58
Banverket	0	0
Datainspektionen	0	0
Glesbygdsverket	0	0
Invandrarverket	17	6
Jordbruksverket	7	8
Konsumentverket	23	127
Kriminalvårdsstyrelsen	3	0.4
Livsmedelsverket	16	32
Naturvårdsverket	2	4.2
Närings- och Teknikutvecklingsverket	25-49 ¹	69-136
Patent- och registreringsverket	0	0
Riksförsäkringsverket	6 ²	9
Rikspolisstyrelsen	5	0.19
Riksskatteverket	13	0.86
Riksåklagaren	0	0
SIDA	130 ³	257
Skogsstyrelsen	2	1.33
Skolverket	10	40
Socialstyrelsen	12	30
Tullverket	0	0
Vägverket	3	0.42

¹ För NUTEK faller 24 rapporter i en gråzon för vad som definieras som självutvärdering. Detta förklaras närmare längre fram.

² Alla sex avser dock Försäkringskassan snarare än RFV själv.

³ Uppskattat antal.

Beräkningen av antal utförda självutvärderingar måste tas med försiktighet. Vissa självutvärderingar är väldigt smala medan andra täcker en stor del av verksamheten. Att enbart se på antalet självutvärde-

ringar ger således ingen bra bild över i vilken utsträckning en myndighet verkligen utvärderar sig själv.

En viktig slutsats är emellertid att sex myndigheter överhuvudtaget inte redovisar några självvärderingar. Tre av dessa tillhör dessutom kategorien stora myndigheter med mer än 1 000 anställda. Om vårt urval är representativt skulle det innebära att cirka en dryg fjärdedel av myndigheterna inte har gjort någon självvärdering alls. Tio myndigheter, eller närmare hälften, redovisar under femårsperioden färre än en självvärdering per 1 000 anställda.

På ett grovt sätt kan tabell 4.3 också användas för att få en bild av kostnaden för självvärderingar. Vi har uppgifter om kostnaden för några självvärderingar. Utifrån det kan man utgå ifrån att kostnader för de flesta självvärderingarna ligger mellan 30 000–150 000 kronor. Om man utgår från ett medelvärde på 70 000 kronor per självvärdering, så utgör de totala kostnaderna för självvärderingar 0,02 procent av myndigheters driftskostnader, eller en tusendels procent av myndigheters totala driftskostnad och transfereringar.

Det måste betonas att detta enbart avser kostnaden för självvärdering. Det är sannolikt att kostnaden för relaterade uppgifter som verksamhetsbeskrivning, uppföljning och myndighetens spridning av information om sig själv vida överstiger kostnaden för självvärdering

4.3 Självvärderingarna – några generella drag

I ett nästa steg delas självvärderingarna in enligt vad de syftar till att mäta. Detta framgår av tabell 4.4. En svårighet här är att en grupp av självvärderingar som i huvudsak består av en utvärderares bedömning inte tydligt mäter någonting specifikt. I stället gör de i allmänhet explicita eller implicita bedömningar av såväl produktivitet som effekter och effektivitet. Som framgår av tabellen är merparten av självvärderingarna av denna karaktär.

Endast några få självvärderingar mäter produktivitet. Vi påminner återigen om att vi inte räknar in enkla produktivetsmått, som ofta återfinns i årsredovisningarna, om det inte görs ett försök att "bedöma och förklara" varför produktiviteten har utvecklats på ett visst sätt.

Något vanligare är mätningar av effekter av åtgärder. Däremot förekommer mätningar av den relativa effektiviteten inte alls.

Tabell 4.4 Vad mäter självvärderingarna?

	Bedömn/ uppfölj- ning	Produk- tivitet	Effektr/ effekti- vit	Relativ effekti- vit	Annat
AMS	15		3		1
Banverket					
Datainspektionen					
Glesbygdsverket					
Invandrarverket		14	3		
Jordbruksverket	5		2		
Konsumentverket	18		5		
Kriminalvårdsstyrelsen		6		1	
Livsmedelsverket	3		13		
Naturvårdsverket	2				
NUTEK	13-33	2	11		4
Patent- och registreringsverket					
Riksförsäkringsverket	2	1	3		
Rikspolisstyrelsen	4		1		
Riksskatteverket	5	5	3		
Riksåklagaren					
SIDA	121	3	6		
Skogsstyrelsen		1		1	
Skolverket	5		5		
Socialstyrelsen		3		7	2
Tullverket					
Vägverket	3				

Av tabell 4.5 framgår vilken metod som använts i självvärderingarna. En metodkategori, professionella bedömningar, överensstämmer av naturliga skäl nästan helt med motsvarande kolumn i tabell 4.4.

I övrigt tycks intervjuundersökningar vara förhållandevis vanliga bland självvärderingar, medan empiriska bearbetningar grundade på andra datakällor är mindre vanliga.

Tabell 4.5 Vilken metod används i självutvärderingarna?

	Profession- nella be- dömare	Intervjuer	Statistisk/ empiriskt	Annat
AMS	12	3	3	
Banverket				
Datainspektionen				
Glesbygdsverket				
Invandrarverket	8	6	8	
Jordbruksverket	5	1	1	
Konsumentverket	6	13	4	
Kriminalvårdsstyrelsen			1	6
Livsmedelsverket	4	7	5	
Naturvårdsverket		2		
NUTEK	13-33	13	5	3
Patent- och registreringsverket				
Riksförsäkringsverket	2	1	3	
Rikspolisstyrelsen	4		1	
Riksskatteverket	1	4	8	
Riksåklagaren				
SIDA	121	3	6	
Skogsstyrelsen		1	1	
Skolverket	3	7		
Socialstyrelsen		3	9	
Tullverket				
Vägverket	1	2		

Tabellerna 4.4 och 4.5 beskriver ganska schablonmässigt självutvärderingarnas omfattning. Därför följer en kort beskrivning av särskilt utmärkande drag i varje myndighets självutvärdering samt vår något subjektiva bedömning av hur väl självutvärderingarna täcker respektive myndighets verksamhet. Beskrivningarna gör inget anspråk på fullständighet.

4.3.1 AMS

Totalt har AMS redovisat ett 50-tal utvärderingar under tidsperioden. Av dessa kan 19 betecknas som självutvärderingar i vid bemärkelse. Tolv av dessa är beskrivningar/bedömningar av olika försöksverksamheter, samt uppföljningar där t.ex. redovisas hur stor andel av personer som genomgått en åtgärd är fortsatt arbetslösa efter en tidsperiod

utan att jämförelse görs med någon kontrollgrupp. Tre av självvärderingarna är s.k. servicemätningar där kunder på arbetsförmedlingarna ombeds att bedöma servicekvaliteten. En är en genomgång av empirisk litteratur på arbetsmarknadsområdet. Resterande tre är självvärderingar med ambitionen att mäta effekter av arbetsmarknadspolitiska åtgärder. Två av dessa har valts (se kapitel 6).

Ganska få självvärderingar belyser verksamhetens effekter. Till bilden hör dock att en hel del utvärderingar av arbetsmarknadspolitiska åtgärder har genomförts utanför AMS av fristående forskare och av EFA (Expertgruppen för arbetsmarknadspolitiska utvärderingsstudier). AMS har också byggt upp databaser som görs tillgängliga för utomstående forskare. Under undersökningsperioden syns en ökning av ambitionsnivån att göra självvärderingar. Detta reflekteras bl.a. i en förteckning över planerade och pågående utvärderingar som AMS sammanställt.

4.3.2 Banverket

Banverket har inte redovisat några självvärderingar alls. Några samhällsekonomiska kostnads-intäkts kalkyler över olika projekt har dock gjorts. Dessa kalkyler bygger emellertid i huvudsak på antagna eller faktiska förändringar i t.ex. antalet resande utan att utvärdera i vilken mån förändringen beror på en genomförd investering eller på externa faktorer. I vår (och två oberoende experters) bedömning bör kalkylerna därför inte räknas som fullgoda självvärderingar enligt RRV:s definition.

4.3.3 Datainspektionen

Datainspektionen har inte redovisat några självvärderingar.

3.4 Glesbygdsverket

esbygdsverket har inte redovisat några självvärderingar. Däremot har en omvärldsanalys enligt Delfitekniken gjorts.

3.5 Invandrarverket

invandrarverket har redovisat 17 självvärderingar. Av dessa är tre enskilda granskningar av asylärenden. Tre är beskrivningar av produktutveckling och handläggningstider. De övriga är bedömningar vanligtvis grundade på intervjuer – av olika delar av verksamheten; ofta försöksverksamheter. Av dessa ingår två i denna studie. Bedömningen av verksamheten genom utvärderingarna förefaller god.

3.6 Jordbruksverket

Jordbruksverket har redovisat sju självvärderingar av enskilda åtgärder som verket administrerar. De flesta utgörs enbart av beskrivningar och bedömningar, i synnerhet de som avser mer centrala delar av livsmedels- och jordbrukspolitiken. Några använder sig dock av empiriska metoder, och mäter effekten av olika kampanjer, angående bekämpningsmedel. En av dessa ingår i denna studie. Täckning av verksamhetsområdet genom självvärderingar förefaller något begränsad.

3.7 Konsumentverket

Konsumentverket har redovisat 23 självvärderingar eller uppföljningar av olika satsningar, informationskampanjer, Råd och Rön m.fl. Bedömningen av verksamheten förefaller god. Många av utvärderingarna förhållandevis enkla attitydundersökningar, men Konsumentverkets utvärderingar kan också sägas vara att påverka attityder. Därutöver finns också några entydiga effektstudier där effekter på konsumentbeteende eller annat mäts. Två av dessa ingår i vår studie.

4.3.8 Kriminalvårdsstyrelsen

Kriminalvårdsstyrelsen har redovisat sju uppföljningar. En av dess är en attitydundersökning. De resterande är oftast beskrivningar av antal återfall efter behandling. I vissa fall redovisas också antal återfall från olika behandlingar. I allmänhet görs dock inte några försök att förklara varför återfallsfrekvensen varierar, t.ex. genom att jämföra antal återfall med likartade grupper, eller att på annat sätt behandla selektionseffekter. Enbart tre av studierna har därför kunnat klassas som självutvärderingar. Mest ambitiös verkar en studie som Kriminalvårdsstyrelsen beställt av Väg- och Trafikinstitutet där ett mer systematiskt försök görs att undersöka effekten av en behandling. Denna studie har valts till sakkunnigbehandling.

4.3.9 Livsmedelsverket

Få självutvärderingar har redovisats för åren 1990–1993. Därefter sker en markant ökning, ungefär åtta stycken per år. Efter 1994 är täckningen av verkets verksamhet god. De flesta utvärderingarna kan betecknas som effektstudier. Självutvärderingarna utgör oftast intervjustudier med frågor om attityder, information och beteende. Två av dessa ingår i denna studie.

4.3.10 Naturvårdsverket

Naturvårdsverket har redovisat två självutvärderingar, båda intervjustudier bland användare av Naturvårdsverkets information. En genomgång av 9 miljömål har emellertid gjorts samt en beskrivning av vad olika aktörer gör på området. Verket redovisar också en sammanställning av regeringsuppdrag som verket fått under åren samt ett antal andra uppföljningar.

4.3.11 Närings- och teknikutvecklingsverket

En del av NUTEK:s uppföljningar och utvärderingar har tidigare granskats i en revision av RRV (RRV 1995:39). Där behandlades två verksamhetsområden, NUTEK Teknik och NUTEK företag. Därtill har vi här kartlagt NUTEK Regional och NUTEK Energi. Sammanlagt redovisar NUTEK ett 50-tal självutvärderingar för perioden. Drygt hälften av dessa är bedömningar gjorda av experter av projekt som får en del av sin finansiering av NUTEK. I vår mening är dessa endast i begränsad omfattning självutvärderingar. Experterna som gör bedömningen har i allmänhet inte information om hur stor andel av projektet som finansieras av NUTEK, och de uttalar sig endast ibland och då perifert om hur de bedömer NUTEK:s roll och betydelse för projektet.⁷

Bland de övriga förekommer en hel del uppföljningar som oftast med hjälp av intervjuer bland företag och forskare som fått stöd försöker skapa en bild av stödets effekter. De räknas således som effektstudier. I några studier ställs också effekterna i relation till kostnaderna och utgör således effektivitetsstudier. Täckningen av verksamheten med självutvärderingar är förhållandevis god.

I RRV:s revision påpekades metodproblem i en del utvärderingar. För att inte överlappa arbete som redan gjorts, har här valts en utvärdering av det regionalpolitiska stödet som NUTEK genomförde 1993 för sakkunnigbedömning. Utvärderingen består av en huvudrapport och ett antal delrapporter.

4.3.12 Patent- och registreringsverket

Patent- och registreringsverket har inte redovisat några självutvärderingar. Kvartalsvis görs dock en resultatuppföljning som också ger underlag till årsredovisningen. Där beräknas produktivitet och tidsåtgång för behandling av ärenden.

⁷ Jfr SIDA.

4.3.13 Riksförsäkringsverket

Riksförsäkringsverket redovisar 25 uppföljningar och 12 utvärderingar av socialförsäkringssystemet. Inga av dessa är emellertid entydiga självutvärderingar av verkets centrala aktiviteter. En typ av utvärdering som verket gör är att beräkna det s.k. ohälsotalet och att statistiskt försöka bestämma orsaksfaktorer bakom förändringar i detta mått. Antydning görs att förändringar som inte kan hänföras till de undersökta exogena faktorer (som t.ex. demografiska förändringar) kan tolkas som effekter av ändringar i socialförsäkringssystemet. I vår bedömning är dock detta en allt för indirekt härledning för att kunna räknas som en självutvärdering av verkets centrala aktiviteter.

RFV är tillsynsmyndighet över försäkringskassorna som dock har egen juridisk status. Jämfört med andra myndigheter med regional organisation har också kassorna en större självständighet. Försäkringskassorna tillhör inte myndigheten RFV. Med denna utgångspunkt är täckningen av Riksförsäkringsverkets aktiviteter begränsad. Även om kassorna räknas in i RFV:s aktiviteter är täckningen ändå inte överväldigande. Man kan däremot beskriva det som att RFV och kassorna tillhör socialförsäkringsadministrationen. Kassornas arbete är i nära anknytning till RFV:s aktiviteter. Med tanke på detta kan sex av ovan nämnda uppföljningar och utvärderingar räknas som självutvärderingar, varav flertalet berör olika försöksverksamheter. Två av dessa ingår i vår studie.

4.3.14 Rikspolisstyrelsen

Rikspolisstyrelsen har redovisat fem självutvärderingar, varav flertalet är bedömningar och intervjuer som rör försöksverksamheter. En av dessa har valts ut till sakkunniggranskning. Täckningen av den centrala verksamheten förefaller vara dålig.

4.3.15 Riksskatteverket

Riksskatteverket har redovisat 13 självvärdering under perioden. Av dessa är fyra attitydundersökningar där människor ombeds bedöma sina kontakter med Riksskatteverket. En självvärdering är en expertbedömning av en del av verksamheten. Fem självvärderingar är produktivitetsstudier där behandling av olika typer av skatteärenden eller upptaxeringar till följd av granskning ställs i relation till kostnader. De tre övriga självvärderingarna behandlar med olika metoder frågor som förekomsten av skattefusk och effekter av skattereformen. En av självvärderingarna, som i vår bedömning tillhör de mest sofistikerade har valts till sakkunniggranskning.

Det är inte helt entydigt vad som skall betraktas som effekter av skattemyndigheters verksamhet. Upptäckt av fler fall av skatteundandragande behandlas i viss mån som en effekt, och i den mening är flera av självvärderingarna också effektstudier, med en ganska god täckning av verksamheten. Men den vidare effekten på omfattning av försök till skatteundandragande behandlas i ganska liten omfattning. Detsamma gäller andra frågor som rör hur skattestrukturen och administrationen påverkar människors beteende. Sådana frågor undersöks dock även av utomstående forskare, bl.a. på uppdrag av Finansdepartementet.

4.3.16 Riksåklagaren

Riksåklagaren har inte genomfört några självvärderingar.

4.3.17 SIDA

SIDA har genomfört ett 30-tal självvärderingar per år. Dessa utvärderingar avser i allmänhet ett biståndsprojekt, som oftast bedöms av någon eller flera experter.⁸

Av 20 SIDA-utvärderingar som vi har gått igenom fanns en som innehöll vissa mått på kostnadseffektivitet. I enskilda fall förekommer också produktivetsmått. Vanligen utgörs de dock huvudsakligen av verbala bedömningar. Sedan 1994 görs emellertid försök att i ökad utsträckning använda sig av kostnads-intäkts analyser.

I utvärderingarna beaktas oftast inte (och kan kanske inte beaktas) indirekta och makroekonomiska aspekter. En del studier tyder på att dessa effekter förtar positiva effekter av bistånd. En fråga som måste resas är emellertid om SIDA:s projektutvärderingarna skulle kunna ge sådan information om den är sann.

4.3.18 Skogsstyrelsen

Skogsstyrelsen har redovisat två självvärderingar. Den ena är i huvudsak en attitydundersökning bland skogsägare. Den andra är en effektstudie av särskilda skogsvårdsinsatser. Den senare ingår i vår studie.

4.3.19 Skolverket

Skolverket har redovisat tio självvärderingar. Tre av dessa är bedömningar av hur organisationsförändringar har påverkat den interna effektiviteten. Två är attitydundersökningar angående hur skolhuvudmän och andra såg på eller reagerat på insatser från Skolverket. Slut-

⁸ En jämförelse kan här göras med de utvärderingar som NUTEK låtit göra av projekt som finansieras. Även här är frågan om utvärderingarna kan ses som självvärderingar mer än i högst begränsad omfattning. I vår bedömning är emellertid SIDA:s operativa ansvar för utformning och genomförande av de flesta projekten större än t.ex. NUTEK Tekniks roll i projekt som får stöd av NUTEK. I den mån den är också utvärderingar av SIDA:s projekt mer att betrakta som självvärderingar än motsvarande hos NUTEK Teknik.

ligen har ett antal läsvärdesundersökningar av olika av Skolverkets skrifter och informationssatsningar gjorts.

Den mest avancerade undersökningen förefaller vara en angående genomslaget av den nationella utvärderingen av grundskolan. Den, samt en utvärdering av den nya fältorganisationen, har valts för sakkunnigbedömning.

Skolverket redovisar också en planerad fördjupning av utvärderingsverksamheten. Detta gäller bland annat effekter av tillsynsverksamhet.

4.3.20 Socialstyrelsen

Socialstyrelsen har redovisat 12 självutvärderingar. Av dessa är två bedömningar gjorda av egen personal, och en är en bedömning gjord av en extern expert. Resterande nio är intervjuundersökningar. Av dessa behandlar sju i vilken omfattning Socialstyrelsens skrifter har lästs och har haft genomslag. En avser allmänhetens attityder gentemot Socialstyrelsen, och en är en pilotstudie om hur Socialstyrelsen uppfattas vid Lex Maria-anmälningar. En utvärdering som tar reda på om människor har läst en skrift kan i vår bedömning inte betecknas som effektutvärdering. Tre av utvärderingarna undersöker dock även effekter i andra termer. Två av dessa ingår i vår studie och en betecknas som "pilotstudie". Vi har valt den ändå, eftersom den metodmässigt förefaller mer gedigen än de andra utvärderingarna. Det faktum att den beskrivs som pilotstudie beaktas dock i kvalitetsbedömningen.

4.3.21 Tullverket

Tullverket har ett omfattande uppföljningssystem. Man har redovisat ett antal resultatmått som t.ex. utvecklingen av antal beslag, antal visitationer och mängden beslagtagna droger. Dessa relateras också till kostnader och utgör därför ett produktivetsmått. De relateras dock inte till arbetsmetoder och inte heller systematiskt till förändringar i omvärldsförhållanden. Det förekommer inte heller utvärderingar av hur tullens arbete påverkar omfattningen av den illegala handeln. Effekt- eller effektivitetsstudier görs således inte.

4.3.22 Vägverket

Vägverket redovisar en rad olika uppföljningar av nationella vägtrafikmål, där t.ex. utvecklingen av antal döda och skadade eller bärighet på vägarna beskrivs. I dessa uppföljningar görs emellertid i allmänhet inga försök att undersöka varför förändringar sker. Därför har endast tre kunnat klassas som självvärderingar.⁹

4.4 Några slutsatser

Genomgången av myndigheternas självvärdering tyder på en mycket ojämn fördelning i flera avseenden. Vissa myndigheter gör många självvärderingar, andra inga alls. Bland självvärderingarna använder sig de flesta av mycket enkla metoder, men några mer sofistikerade studier förekommer.

Till en del skulle denna ojämna fördelning kunna bero på olika förutsättningar på olika verksamhetsområden. Huvudintrycket är emellertid att ambitionsnivån oftast är avgörande.

⁹ Dessa tre redovisades först vid efterkontrollen. Ingen av dessa har därför valts ut för sakkunnigbedömning.

5 Metod för kvalitetsbedömning

En utvärderings kvalitet har många dimensioner. Den är också i viss mening subjektiv. Det som är viktigt för en läsare kan verka oväsentligt i en annans ögon. Olika utvärderare kan också göra olika bedömningar om lämpligheten i olika metoder eller vad rimliga tolkningar av resultat kan vara. Av dessa skäl kan det förefalla omöjligt att exakt mäta eller beskriva kvaliteten på en utvärdering.

Samtidigt finns ofta en betydande samstämmighet bland utvärderare och forskare om grundläggande principer för vad som utgör en bra utvärdering. Sådana principer har också formulerats på olika sätt. Ett är de s.k. "Program Evaluation Standards" (The Program Evaluation Standards. 2nd Edition. Sage Publications), en skrift som beskriver kriterier (standards) som man menar att en god utvärdering skall uppfylla. Ett sådant kriterium är t.ex. att dokumentationen skall vara så utförlig att läsaren i princip skulle kunna kontrollera de empiriska bearbetningarna. Därtill skall data och metod vara lämpliga och slutsatserna relevanta och användbara. Andra kriterier handlar om vikten av objektiva och trovärdiga utvärderare.

Kriterierna blir meningsfulla därför att de flesta initierade personer i de flesta fallen gör en likartad bedömning av huruvida de är uppfyllda. Däremot kan vissa personer i enskilda fall göra avvikande bedömningar, och för vissa utvärderingar kan det uppstå en större spridning av bedömningar.

Konsekvensen av denna grundsyn är att vi dels har valt att granska ett stort antal självutvärderingar, dels att två sakkunniga bedömt varje självutvärdering. Urvalet av självutvärderingar bygger i huvudsak på myndigheternas förslag om de bästa, och myndigheterna har dessutom haft visst inflytande över val av sakkunniga.

Det normala förfarandet i många utvärderings- och forskningssammanhang är att kvaliteten bedöms av oberoende experter eller utvärderare, med avseende på de uppställda kriterierna. Ett sådant förfarande har emellertid hittills aldrig tillämpats på de svenska statliga myndigheternas utvärderingar.

Inför vår tillämpning av en sådan sakkunniggranskning eller -bedömning aktualiserades flera frågor. Hur skall utvärderingar väljas ut? Hur skall sakkunniga väljas ut? Vilka kriterier skall de sakkunniga tillämpa, dvs. vad skall de jämföra med? Det är uppenbart att dessa frågor måste hanteras på ett sätt som inte snedvrider granskningen. Vi har därför lagt ner en hel del möda för att utveckla och tillämpa ett så rättvist förfarande som möjligt. Detta beskrivs i avsnitten 5.1–5.3.

5.1 Hur har självutvärderingarna valts ut?

En möjlighet hade varit att helt slumpmässigt göra ett urval av myndigheters självutvärderingar. Vår kartläggning visade emellertid uppenbart att ett sådant förfarande skulle kunna ha lett till felaktiga slutsatser. Ett problem är att fördelningen mellan myndigheter skulle ha blivit sned. De få myndigheter med många utvärderingar hade blivit överrepresenterade, medan flera med få utvärderingar inte hade kommit med alls. Ett än större problem är att ett slumpmässigt urval hade lett till att många enkla attitydundersökningar, som antalsmässigt dominerar, hade tagits ut till sakkunniggranskning, medan endast ett fåtal av de mer sofistikerade utvärderingarna hade kommit med.

Vi har här strävat efter att få med de mest kvalificerade självutvärderingarna. De mest sofistikerade utvärderingarna valdes ut. Konkret innebär detta att i första hand valdes utvärderingar som mätte effektivitet eller effekter på ett statistiskt/empiriskt sätt (enligt tabellerna 6.3 och 6.4).¹⁰ Om det fanns flera sådana för en myndighet så valdes två slumpmässigt. Om det inte fanns några sådana valdes två slumpmässigt av resterande utvärderingskategorier (enligt tabellerna

¹⁰ I de fall respektive myndighet framhöll att de själva betraktade vissa utvärderingar som bäst accepterade vi den bedömningen.

4.4 och 4.5). Ambitionen var således att välja två från varje myndighet.¹¹

Tabell 5.1 visar vilka utvärderingar som har valts ut för respektive myndighet (i bilaga 2 ges fullständiga referenser till respektive utvärdering). I två fall består utvärderingen av flera rapporter. Totalt har 23 utvärderingar valts ut för sakkunnigbedömning.

Tabell 5.1 Vad handlar de uttagna självutvärderingar om?

Självutvärderingens frågeställning	
Starta-eget bidrag	AMS
Rekryteringsstödet resultat	AMS
Bosnier hjälper Bosnier	Invandrarverket
Kontakt nät för invandrare	Invandrarverket
Minskade hälso-miljörisiker av bekämpningsmedel	Jordbruksverket
Varudistribution	Konsumentverket
Värderad service	Konsumentverket
Kurs för fängelsedömda rattfyllerister	Kriminalvårdsstyrelsen
Infokampanj om fisk till gravida och ammande	Livsmedelsverket
Utvärdering av "Startpaket Nyckelhål"	Livsmedelsverket
Regionalpolitiska stödets effekter (Huvudrapport och 7 delstudier)	NUTEK
Försöksverksamhet inom rehabområden	Riksförsäkringsverket
Rehabarbetet på försäkringskassan	Riksförsäkringsverket
RAR-systemet (lokalt ADB-stöd)	Rikspolisstyrelsen
Kvalitetsutveckling inom skatteförvaltningen	Riksskatteverket
Importstöd till några u-länder:	
– Importstöd till Guinea-Bissau m.fl.	SIDA
– Importstöd till Tanzania	SIDA
– Importstöd till Cap Verde	SIDA
Särskilda skogsvårdsinsatser i norra Sverige	Skogsstyrelsen
Utvärdering av nya fältorganisationen	Skolverket
Nationell utvärdering av grundskolan (3 st. i en bok)	Skolverket
Lex Maria	Socialstyrelsen
Kvalitetssäkring inom vården	Socialstyrelsen

¹¹ I några fall valdes bara en, t.ex. Rikspolisstyrelsen.

5.2 Hur har de sakkunniga valts ut?

Bland experter inom olika områden finns i vissa fall sådana som är kända för att vara mer positiva eller mer negativa till en myndighet eller en myndighets verksamhet. Detta kan naturligtvis vara välgrundade bedömningar baserade på tidigare erfarenheter med respektive myndighet. Samtidigt kan det också finnas ett inslag av förutfattade meningar, eller personliga konflikter. Av dessa skäl har vi försökt välja sakkunniga som inte är kända för en positiv eller negativ inställning till respektive myndighet. I urval av sakkunniga har varje myndighet fått tillfälle att utöva ett inflytande.

Andra krav som ställs på de sakkunniga är att de dels har en gedigen kunskap om utvärderingsmetoder, dels har kunskap om sakområdet.

I en första omgång har vi sammanställt listor av mellan 5 och 10 tänkbara sakkunniga för varje myndighet. Listorna byggde till stor del på förslag från den referensgrupp som ESO knutit till vårt projekt.

I en andra omgång sändes listorna till respektive myndighet. Varje myndighet fick då tillfälle att stryka namn som de inte ville ha. Det fanns ingen explicit restriktion för hur många namn en myndighet fick stryka. Däremot gjordes klart att om myndigheten inte accepterade något av namnförslagen eller om den person som en myndighet helst vill ha inte ställer upp som sakkunnig så skulle nya sakkunniga väljas utan att myndigheten tillfrågades en gång till.

Tre myndigheter anförde klagomål mot samtliga sakkunniga eller mot urvalsförfarande för sakkunniga. I efterhand kan vi dock konstatera att det inte finns något samband mellan antal strykningar eller klagomål och utfallet. Självutvärderingar av myndigheter som var mer negativa till sakkunniga bedömdes inte i genomsnitt mer negativt av de sakkunniga.

Två myndigheter kommenterade också den stora andelen nationalekonomer bland sakkunniga. Dessa utgjorde 58 % av de som genomfört kvalitetsgranskning. Självutvärderingarna fokuserar emellertid ofta på begrepp som produktivitet eller effektivitet som i hög grad är ekonomiska begrepp. Det är därför försvarbart om det blir en viss övervikt av ekonomiskt orienterade personer bland sakkunniga. En

efterhandskontroll visar att nationalekonomerna bland de sakkunniga inte varit mer negativa än andra sakkunniga.

Sammanlagt valdes 40 sakkunniga. (En förteckning på sakkunniga finns i bilaga 3). I ett fall gjordes en bedömning av två forskare gemensamt. Tre av de sakkunniga har granskat två självutvärderingar var. Normalt valdes emellertid två sakkunniga per utvärdering. En sakkunnig har inte haft kännedom om vem mer som haft till uppgift att granska samma utvärdering. Detta gjorde att vi kunde göra en avstämning för att se om bedömningarna skilde sig åt i någon större utsträckning. Två av varandra oberoende utvärderare gör också bedömningarna i genomsnitt "säkrare". I efterhand har det visat sig att i de flesta fall var de båda sakkunnigas bedömningar av samma utvärdering i hög grad samstämmiga. Detta kommenteras mer längre fram.

Sakkunniga tilläts vara anonyma om de så önskade. Det diskuterades ingående i projektets referensgrupp¹² om anonymiteten skulle kunna leda sakkunniga till mindre gedigna granskningar. Mot det anfördes dock att en del sakkunniga inte skulle tordas skriva sin mening om deras namn offentliggjordes. Detta slutade i kompromissen att låta de sakkunniga avgöra frågan själv. Endast ett fåtal valde att förbli anonyma (fyra stycken). I efterhand kan sägas följande om farhågorna angående anonymiteten. Enligt vår bedömning kan endast några få bedömningar beskrivas som mindre genomarbetade. Någon enstaka av dessa är bland de anonyma. Det finns också sakkunnigbedömningar av de anonyma som tillhör de mest välskrivna och genomarbetade rapporterna. Nästan samtliga som har avgett positiva omdömen om en självutvärdering har gått med på att låta sitt namn offentliggöras. Å andra sidan finns sakkunniga som avgett kritiska bedömningar och har valt att inte vara anonyma. Med tanke på att de sakkunniga i allra flesta fall är mycket kunniga och trovärdiga är vårt subjektiva intryck att fördelarna med att tillåta anonymitet väger tyngre än risken för mindre gedigna utlåtanden som detta kan föra med sig.

De sakkunniga rekryterades från tre grupper med olika bakgrund. Den största gruppen bestod av forskare eller personer med forskaran-

¹² I referensgruppen ingick Göran Arvidsson, SNS, Peter Bohm, Stockholms universitet, Lars Hultkrantz, Högskolan Dalarna och Uppsala universitet, Marja Lemne, ESO, Ingrid Munck, SCB och Göran Schubert, ESO.

knytning inom ett område som låg nära det som utvärderingen fokuserade på. Den andra, och mycket mindre, gruppen bestod av utvärderingskunniga inom vissa utredande myndigheter t.ex. Statskontoret och SCB. Den tredje och minsta gruppen, bestod av konsulter med inriktning på utvärdering.

I vår bedömning har det generellt funnits en tydlig skillnad mellan konsultgruppen och de andra. I konsultgruppen tycktes frågor som layout och läsvänlighet ibland tillmätas större vikt än metodfrågorna. Det verkade inte heller finnas en erfarenhet av att skriva bedömningar om utvärderingar.

Vår slutsats av dessa observationer inför framtida granskningar är dels att anonymitet är viktigt, dels att urvalet av sakkunniga är så viktigt att det *inte* kan göras rutin- eller slumpmässigt.

5.3 Vad skall de sakkunniga jämföra med?

De sakkunnigas svar hade kunnat standardiseras med t.ex. en poängsättning. Vår referensgrupp gjorde emellertid den bedömningen att detta skulle styra in de sakkunniga alltför mycket. Därför har de sakkunniga haft stora möjligheter att formulera sin bedömning själv, styrd endast av några brett formulerade delfrågor.

När det gäller instruktioner till sakkunniga har vi utgått från "Program Evaluation Standards", men sedan ansett det nödvändigt att formulera om, slå ihop och lämna bort en del. Instruktionerna har gått genom två "pilotstudieomgångar" med sammanlagt sex testpersoner. Instruktionerna redovisas i *bilaga 2*.

De sakkunniga ombads att göra sin bedömning utifrån en jämförelse med "vad de bedömde som god praxis på området". Detta är ganska vagt hållet, och skulle kunna leda till en osäkerhet om vad de sakkunniga jämför med. Det har emellertid visat sig att de sakkunniga nästan alltid tydligt anger vad de jämför med. Samtidigt är de sakkunniga erfarna forskare och/eller utvärderare/utredare vilket gör att dom har en god grundkännedom om hur en godtagbar utvärdering bör se ut. Ofta anger de också hur de anser att en utvärdering förhåller sig till olika referenspunkter. En sakkunnig har t.ex. påpekat att rådande "praxis" (av utvärderingar inom området) inte är av god kvalitet och bedöm-

ningarna är därför baserade på vad den sakkunnige anser bör vara godtagbar kvalitet på en utvärdering. En annan sakkunnig skriver att utvärderingen bedöms vara bättre än de flesta av myndighetens utvärderingar, men ändå otillfredsställande jämfört med god praxis. En sakkunnig utgår från RRV:s definition av utvärdering för budgetprocessen nämligen att "bedöma och förklara prestationer och händelseförlopp i en myndighetsverksamhet" samt att "klargöra och skärskåda myndighetens prestationer och/eller prestationernas effekter i samhället". En annan sakkunnig utgår ifrån "en vägledande princip för en utvärdering av ett projekt är att undersöka om nytta av projektet överstiger uppoffringarna (kostnaderna) för att genomföra det". Den sakkunniga påpekar samtidigt svårigheten med att göra en perfekt utvärdering men har ändå angivit viktiga utgångspunkter.

I övrigt ombads de sakkunniga att kommentera följande punkter¹³:

¹³ Flera sakkunniga har påpekat att dessa punkter är heltäckande och genomtänkta.

A Dokumentation

Är verksamheten som utvärderas, tidigare utvärderingar, data som används, metoder, resultaten och annat av relevans beskrivet på ett sådant sätt att läsaren kan skapa sig en rimlig bild av utvärderingens tillförlitlighet?

B Data

Är de data som ligger till grund för utvärderingen lämpliga och tillräckligt tillförlitliga?

C Metod

Är utvärderingsmetoden lämplig? Är den konsekvent och riktigt tillämpad?

D Slutsatserna

Är de slutsatser som dras väl underbyggda i analysen? Finns andra rimliga tolkningar av analysen som inte beaktas?

E Kostnadseffektivitet

Anser du att det hade varit möjligt att göra en väsentligt mycket bättre utvärdering med en bättre uppläggning, men – såvitt du kan bedöma – ungefär samma resursinsats när det gäller t.ex. datainsamling, bearbetning, skrivande osv ?

F Objektivitet

Anser du att tillvägagångssättet i utvärderingen, val av frågeställning, tolkning av resultat m.m. ger uttryck för en godtagbar objektivitet hos utvärderaren?

6 Kvaliteten i statliga myndigheters självutvärderingar

6.1 Helhetsbilden

Eftersom de sakkunniga har bedömt självutvärderingarna med egna ord och utan någon standardiserad mall, så kan helhetsbilden inte återges i exakta siffror. Vi har därför valt att göra en enkel kodning av resultaten och sammanställt dessa i tabellform för att få en överskådlig bild. Vi har dessutom kommenterat och beskrivit viktiga dimensioner kvalitativt. Detta innebär givetvis en viss risk att bilden färgas av våra glasögon. För att minska den risken har vi dessutom låtit en oberoende expert koda de sakkunnigas bedömningar. Denna kontrollkodning gav nästan identiska resultat som vår kodning.¹⁴

Samtidigt kan dock vår bild kontrolleras genom att läsa sakkunnig-utlåtanden (se bilaga 5). I många fall är de mycket välskrivna och utgör intressant läsning även för den som inte läst de underliggande självutvärderingarna. Vi har också angett utgångspunkter för den enkla kodningen och exemplifierat vårt tillvägagångssätt.

Vi har genomgående valt att koda bedömningar efter följande mall när det gäller de övergripande bedömningarna:

¹⁴ Kontrollkodningen genomfördes av ekonomie doktor Roger Svensson. För den övergripande bedömningen skildes inte kodningen från vår egen med mer än ett tecken i något fall.

- (--) = mycket kritisk (mycket stora brister, uppfyller inte minimikraven, beskrivande rapport, någon utvärdering görs ej)
- (-) = kritisk (mindre bra, utvärderingen har en del brister)
- (0) = neutral (varken eller; rimlig nivå, finns både bra och dåliga punkter i utvärderingen)
- (+) = positiv (bra)
- (++) = mycket positiv (mycket bra, utmärkt, ett mycket viktigt bidrag)

I tabell 6.1 ger vi en bild av det övergripande resultatet. Vi har fått in totalt 46 olika utlåtanden (två per utvärdering). De övergripande bedömningarna fördelar sig på följande sätt.¹⁵

Tabell 6.1 Totalbild av de sakkunnigas helhetsbedömningar

(--)	(-)	(0)	(+)	(++)	Totalt
21	16	1	6	2	46

Av de utvärderingar vi har tagit ut har endast sex stycken av totalt 23 utvärderingar erhållit minst en positiv bedömning (+) eller (++)). Två utvärderingar (RFV, SIDA) har av de sakkunniga bedömts med (++)). Den totalt sett låga kvaliteten i utvärderingarna är något förvånande. De sakkunniga gav ett negativt slutomdöme till mer än två tredjedelar av självutvärderingarna.

¹⁵ Det är värt att notera att kontrollkodningen inte skiljer sig mer än ett steg i något fall från vår kodning.

I tabell 6.2 visas de övergripande bedömningarna av respektive utvärdering. Kodningen bygger på följande resonemang.

Det finns fall där de sakkunniga menar att de mest elementära kraven på en utvärdering saknas, avsaknad av försök att mäta resultat av en verksamhet, avsaknad av diskussion om i vilken utsträckning målen för verksamheten uppfylls, osv. Vi har då klassat den övervägande bedömningen som kritisk (-) eller (--).

Låt oss ge några exempel. En sakkunnig skriver i den övergripande bedömningen att:

"Den granskade "utvärderingen" är ett gott exempel på en ...relativt enkel uppföljningsinsats kring ett projekt...men svarar inte på de grundläggande frågor som ställs".

Här har vi gett bedömningen (--), dels för att den sakkunniga inte anser rapporten vara en utvärdering, dels för att den inte svarar på de frågor som ställs.

En annan bedömning som vi har tolkat som kritisk (-) men som kanske t.o.m. skulle kunna klassas som (--) har följande sammanfattning:

"Utvärderingens styrka är att det finns en intresseväckande beskrivning....Utvärderingens svaghet är att man kring de centrala frågeställningarna saknar en riktig datainsamling och en genomtänkt metod för att göra jämförelser. Därtill kan läggas att utvärderaren har en okritisk positiv hållning till hela projektet, vilket resulterar i att slutsatserna är lite distanslösa."

En sakkunnig skriver i den sammanfattande bedömningen att:

"...Då det dessutom finns mycket få mätbara kriterier för output i den organisation som utvärderas blir det utifrån det underlag jag erhållit omöjligt att bilda sig en välgrundad uppfattning om hur väl verksamheten fungerat".

Den bedömningen har vi tolkat som mycket kritisk (--) eftersom utvärderingen inte verkar bidra till kunskap om åtgärdens ändamålsenlighet (se kapitel 2 för en diskussion om begreppet utvärdering).

De sakkunniga som gett utvärderingen "väl godkänt" eller uttryckt att "huvudintrycket är mycket positivt" i sammanfattningen har vi klassat som (++).

I tolkningen av tabellerna skall man komma ihåg att den *inte* avser myndigheters *genomsnittliga* självutvärdering. Man bör därför inte övertolka utfallet för enskilda myndigheter.¹⁶

¹⁶ De sakkunnigas bedömningar återfinns i bilaga 5. Två granskningar per utvärdering har gjorts. I bilagan benämns dessa utlåtanden 1 och utlåtande 2. I tabellerna motsvarar kodningen benämningen i bilagan. Den första bedömningen i tabellerna kommer från utlåtande 1.

Tabell 6.2 Sakkunnigas övergripande bedömningar av självutvärderingarna¹

Starta eget-bidrag. (AMS)	(--)	(--)
Rekryteringsstödet resultat. 1995:1 (AMS)	(--)	(+)
Bosnier hjälper Bosnier (SIV)	(--)	(--)
Tusentals kontakter - nätverk för framtiden. (SIV)	(--)	(-)
Minskade hälso- och miljörisker vid användning av bekämpningsmedel (Jordbruksverket)	(--)	(+)
Varudistribution. (Konsumentverket)	(-)	(-)
Värderad service (Konsumentverket)	(-)	(-)
"Rattfällan." Utvärdering av kurs för fängelsedömda rattfyllerister (KVS)	(0)	(-)
"Goda råd om fisk" – utvärdering av information (Livsmedelsverket)	(--)	(-)
Utvärdering av "Startpaket Nyckelhål" (Livsmedelsverket)	(--)	(--)
"Värdet av regionalpolitiken" (NUTEK)	(--)	(-)
Försöksverksamhet inom rehabiliteringsområdet i Malmöhus, Norrbottens och Blekinge län (RFV)	(++)	(++)
Rehabiliteringsarbetet på försäkringskassan (RFV)	(--)	(-)
Utvärdering av RAR-systemet (RPS)	(--)	(-)
Kvalitetsutveckling inom skatteförvaltningen (RSV)	(-)	(--)
Import support to Guinea-Bissau, Mozambique, Nicaragua and Vietnam. (SIDA)	(--)	(+)
Fungibility and Sustainability. Import Support to Tanzania (SIDA)	(+)	(+)
Swedish Support to Cap Verde. (SIDA)	(-)	(-)
Resultat av Särskilda skogsvårdsinsatser i norra Sverige (Skogsstyrelsen)	(--)	(--)
Utvärdering av den nya fältorganisationen (Skolverket)	(--)	(+)
Den nationella utvärderingen av grundskolan (Skolverket)	(--)	(-)
Pilotstudie av Lex Maria anmälningar och beslut. (Socialstyrelsen)	(-)	(--)
Rapport till Socialstyrelsen. 8 sept 1995 (Socialstyrelsen)	(-)	(--)

¹ I bilaga 2 återfinns en fullständig förteckning över utvärderingarna.

Det är intressant att se hur ofta två sakkunniga ger samma bedömning om samma utvärdering. Endast i 4 av 23 fall gjorde två sakkunniga bedömningar som skiljde sig mer än ett steg från varandra.

I flera fall är det givetvis förenat med svårigheter att göra en bra utvärdering som ger kunskap om verksamhetens ändamålsenlighet och kostnadseffektivitet. Vissa verksamheter är mer komplicerade att utvärdera än andra. Att göra en jämförelse mellan utfallen för en experimentgrupp, som utsätts för en åtgärd och en kontrollgrupp, som inte

deltar, kan t.ex. innebära praktiska problem. I flera granskningar lyfter också sakkunniga fram problemen med att göra goda utvärderingar. Men samtidigt anför många konkreta tillvägagångssätt för att höja kvaliteten på respektive utvärdering betydligt.

I en utvärdering av självhjälpsverksamhet – Bosnier hjälper Bosnier (Invandrarverket) – undersöks hur bosniska flyktingar i Sverige i egen regi sökt tillvarata den egna gruppens kompetens för att underlätta deras integration och möjliggöra eventuell återutvandring. Syftet med utvärderingen var bl.a. att ange vad Invandrarverkets stöd betydde för verksamheten. En sakkunnig skriver att själva utvärderingsmetoder främst går ut på att beskriva vad som utträttats i olika projekt. Trots det angivna syftet görs inte några försök att utröna vilken betydelse SIV:s stöd egentligen haft. I det sammanfattande omdömet anför följande:

"Studien kan inte kallas för en utvärdering i gängse vetenskaplig mening. En sådan utvärdering är nämligen en jämförelse mellan utfallen för en experimentgrupp, som utsätts för åtgärden, och en kontrollgrupp, som inte deltar. De beskrivningar av utfallen i experimentgruppen som redovisas är därmed inte ointressanta, men man måste naturligtvis vara mycket försiktig med långtgående slutsatser. Ofta är det förenat med stora praktiska eller administrativa svårigheter att använda sig av en kontrollgrupp. I stödverksamhetsprojektet borde man dock ha kunnat använda sig av en sådan grupp, eftersom denna verksamhet inte kom till stånd på mer än två orter – Borås och Lund – och det finns andra orter med många bosnier som man ganska lätt kunnat jämföra med. Intervjuundersökningar skulle kunna ha genomförts med deltagare respektive icke-deltagare. En intressant fråga är t.ex. om icke-deltagare anlitate den reguljära sjukvården i större utsträckning än deltagare. I så fall skulle man ha kunnat peka på möjligheten till kostnadsbesparingar av självhjälpsprojektet. Enligt rapporten skall en intervjuundersökning genomföras, dock endast med bosnier på orter där verksamheten bedrivits."

I detta exempel på granskning framkommer en medvetenhet om metodsvårigheter, samtidigt som det på ett tämligen konkret sätt ges förslag på hur kvaliteten kan förbättras i utvärderingen.

I de olika granskningarna förekommer ofta hänvisningar till olika jämförelsenormer. I flera fall uttrycks en viss uppskattning av myndigheters försök, eller förbättringar i standarden trots att utvärderingar fortfarande beskrivs som otillfredställande. I en del fall sägs också att

en utvärdering skulle kunna vara acceptabel som en intern, "quick and dirty", uppföljning, men inte som en riktig utvärdering.

I en del självutvärderingar som vi tagit ut sägs i syftet att man vill "ge en bild" av någonting. Denna skrivning är viktig att diskutera. I syftet framgår att det är en kartläggning man vill göra vilket givetvis är intressant i olika sammanhang. Trots det försiktiga syftet kan det sedan komma slutsatser där kausala samband och effektiviteten av den undersökta åtgärden fastslås. Slutsatserna måste då givetvis prövas mot den metod som används i rapporten. Ett exempel är en AMS-rapport ("Starta eget-bidrag: En totalundersökning över de personer som avslutade bidragsperioden i oktober -93") som ingår i vår studie. Rapporten beskriver situationen för de personer och därigenom de företag som beviljats starta-eget bidrag (SEB). Dess fastslagna syfte är att kartlägga dessa personers arbetsmarknadssituation och deras uppfattningar om bidragshandlingen efter avslutad bidragsperiod. Sedan slås det emellertid fast i denna "kartläggande studie" att "Vart fjärde företag har mer än fem anställda. Eftersom antalet anställda i företagen överstiger antalet personer som ej längre bedriver verksamhet har SEB därmed resulterat i ett sysselsättningstillskott som överstiger antalet beviljade SEB". En av de sakkunniga menar att slutsatsen saknar grund och förutsätter en rad inte alltid rimlig antaganden. Detta diskuteras mer utförligt nedan där bedömningarna av utvärderingarnas slutsatser presenteras. Intressant är emellertid att studien är ett exempel där man i syftet "friskriver sig" från att göra en "riktig" utvärdering men ändå med bestämdhet slår fast olika resultat. I detta fall bör givetvis resultatet i studien bedömas utifrån om det i studien finns empiri och metod som håller för sådana slutsatser.

Just frågan om vad som kan räknas som en riktig utvärdering återkommer ofta i sakkunnigutlåtanden. Många konstaterar exempelvis att en utvärdering reser viktiga och relevanta frågor i inledningen, men sedan använder empiri som inte ger svar på dessa frågor utan på perifer, mindre relevanta frågor. En del anför att en utvärdering leds fel av att den inte utgår från någon klar frågeställning. Ett exempel är Skolverkets nationella utvärdering av grundskolan 1987-1994. En sakkunnig menar att flera olika syften dyker upp. Syftet som anges i utvärderingens inledning är inte detsamma som vad som används längre fram i samma utvärdering. En sakkunnig anför följande:

"Syftet har därför inte i första hand varit att bedöma Skolverket på sätt att genomföra den nationella utvärderingen (som de själva skriver är syftet). Det har därför varit oerhört frustrerande att försöka läsa denna text som en utvärdering av genomförandet".

Det finns sakkunniga som anser att utvärderingen över huvud taget inte nämner de i sammanhanget relevanta frågeställningarna. Ett exempel är en utvärdering av ett lokalt ADB-stöd (RAR – Rationell Anmälan Rutin) för brottsanmälan (Rikspolisstyrelsen). Av utvärderingen framgår inget annat än att valet stod mellan att sprida RAR till övriga polismyndigheter eller att behålla de manuella rutinerna. Båda sakkunniga påpekar (oberoende av varandra) att det är tveksamt om detta är den rätta frågeställningen. Den ena sakkunnige skriver:

"En fråga ... är om Rikspolisstyrelsen valt rätt huvudfrågeställning för utvärderingen. Stod Rikspolisstyrelsens val verkligen mellan att sprida just RAR till övriga polismyndigheter eller att behålla de manuella rutinerna för brottsanmälan?"

Den andra påpekar:

"Att man skall ha ett ADB-system för administrativa handläggningsrutiner i samband med brottsanmälan är en självklarhet i dagens IT-samhälle men att det skall utformas enligt RAR-systemet är en helt obesvarad fråga".

Ett annat exempel där frågeställningens betydelse diskuteras är en bedömning av Livsmedelsverkets utvärdering "Goda råd om fisk". Utvärderingen går ut på att belysa om verkets information i en viss fråga – kostråd om fisk till gravida och ammande kvinnor – nått ut till målgruppen, varit begriplig och lett till ändrade kostvanor. En sakkunnig har gett utvärderingen godkänt eller bra betyg för de enskilda aspekterna såsom dokumentation, data, metod, slutsatser osv. Men den övergripande omdömet blir ändå negativt. Den sakkunniga för en diskussion om utvärderingens nytta och åtgärdens kostnadseffektivitet och konstaterar att:

"Studien är så avgränsad att den inte ger något meningsfullt svar på frågan om den verksamhet som den utvärderar varit framgångsrik".

När frågeställningar diskuteras undrar en del sakkunniga om den utvärdering som de granskats skall ses som ett komplement till andra utvärderingar som undersöker de centrala frågorna. Utifrån vår kartläggning kan vi dock konstatera att det i inget av dessa fall funnits andra utvärderingar som har undersökt de frågor som sakkunniga lyft fram som de centrala.

En mycket tydlig slutsats av att så många sakkunniga reser dessa frågor är att urvalet av självutvärderingsfrågor är ett problem.

En annan aspekt som diskuteras i sakkunnigbedömningarna är avsaknaden av försök till kostnads- och intäktsvärdering. Ett stöd eller en ny verksamhet kan utvärderas utan att utvärderaren nämner den samhällsekonomiska kostnaden respektive förväntad intäkt för åtgärden. Detta problem påpekas av flera sakkunniga. Ett exempel är en utvärdering av rekryteringsstödet resultat (AMS) där en av de sakkunniga skriver att frågan om vad rekryteringsstödet kostat och om det är kostnadseffektivt relativt andra åtgärder helt har försumrats. Ett annat exempel är RFV:s studie av rehabiliteringsarbetet på försäkringskassan. Resultaten är av stort intresse menar en sakkunnig.

"Däremot ... är det en påtaglig brist att begrepp som kostnader och intäkter/nyttor helt lyser med sin frånvaro. Rapportens värde som beslutsunderlag skulle tillta dramatiskt om kostnaderna för olika delar av den diskuterade verksamheten redovisats."

6.2 Dokumentation, data och metod

De sakkunnigas bedömningar av dokumentation, data och metod framgår av tabell 6.3. Kodningen i tabellen bygger på följande resonemang. Om datamaterialet/underlaget bedömts som tillförlitligt och relevant har vi betraktat detta som positivt. En sakkunnig menar att

"...de data som används, metoden och resultaten är beskrivna på ett sätt så att läsaren kan skapa sig en rimlig bild av utvärderingens resultat".

Här har vi tolkat bedömningen av dokumentationen som positiv (+), dock inte mycket positiv (++) eftersom den sakkunniga samtidigt gör vissa anmärkningar på dokumentationen. Om bedömningen däremot

gett uttryck för att "materialet är för litet för att vara statistiskt tillförlitligt" har vi tolkat detta som negativ kritik (-). Om metoden bedöms som lämplig har vi angett detta som (+) eller (++) beroende på övriga kommentarer om metoden. Om den sakkunniga däremot beskriver metoden som den "enklast möjliga statistiska metod ... där "inga samband undersökts på ett tillfredsställande sätt" har detta tolkats som negativ kritik.¹⁷

¹⁷ De sakkunnigas bedömningar återfinns i bilaga 5. Två granskningar per utvärdering har gjorts. I bilagan benämns dessa bedömning 1 och bedömning 2. I tabellerna motsvarar kodningen benämningen i bilagan. Den första bedömningen i tabellen kommer från bedömning 1.

Tabell 6.3 Dokumentation, data och metod i utvärderingarna.

	Dokumen- tation	Data	Metod
Starta eget-bidrag. (AMS)	(--)(+)	(--)(--)	(--)(--)
Rekryteringsstödet resultat. 1995:1 (AMS)	(+)(++)	(++)(++)	(-)(+)
Bosnier hjälper Bosnier (SIV)	(-)(0)	(?)(0)	(+)(--)
Tusentals kontakter - nätverk för framtiden. (SIV)	(++)(--)	(-)(-)	(-)(--)
Minskade hälso- och miljörisker vid användning av bekämpningsmedel (Jordbruksverket)	(--)(+)	(?)(+)	(--)(0)
Varudistribution. (Konsumentverket)	(0)(-)	(-)(-)	(?)(-)
Värderad service (Konsumentverket)	(+)(0)	(+)(-)	(0)(-)
"Rattfällan." Utvärdering av kurs för fängelsedömda rattfyllerister (KVS)	(++)(+)	(+)(+)	(+)(0)
"Goda råd om fisk" - utvärdering av information (Livsmedelsverket)	(++)(+)	(+)(+)	(0)(+)
Utvärdering av "Startpaket Nyckelhål" (Livsmedelsverket)	(++)(0)	(?)(-)	(-)(-)
"Värdet av regionalpolitiken" (NUTEK)	(0)(+)	(--)(0)	(--)(0)
Försöksverksamhet inom rehabiliteringsområdet i Malmöhus, Norrbottens och Blekinge län (RFV)	(+)(++)	(+)(++)	(+)(++)
Rehabiliteringsarbetet på försäkringskassan (RFV)	(0)(0)	(+)(0)	(?)(+)
Utvärdering av RAR-systemet (RPS)	(-)(0)	(-)(-)	(-)(0)
Kvalitetsutveckling inom skatteförvaltningen (RSV)	(-)(--)	(--)(0)	(-)(0)
Import support to Guinea-Bissau, Mozambique, Nicaragua and Vietnam. (SIDA)	(+)(+)	(?)(?)	(?)(-)
Fungibility and Sustainability. Import Support to Tanzania (SIDA)	(++)(+)	(?)(?)	(+)(-)
Swedish Support to Cap Verde. (SIDA)	(-)(--)	(-)(-)	(-)(+)
Resultat av Särskilda skogsvårdsinsatser i norra Sverige (Skogsstyrelsen)	(-)(+)	(0)(+)	(-)(-)
Utvärdering av den nya fältorganisationen (Skolverket)	(--)(+)	(-)(-)	(--)(?)
Den nationella utvärderingen av grundskolan (Skolverket)	(--)(-)	(?)(-)	(-)(-)
Pilotstude av Lex Maria anmälningar och beslut (Socialstyrelsen)	(+)(-)	(-)(-)	(?)(+)
Rapport till Socialstyrelsen. 8 sept 1995 (Socialstyrelsen)	(--)(-)	(-)(-)	(+)(?)

¹ I bilaga 2 återfinns en fullständig förteckning över utvärderingarna.

utlåtanden om RFV:s utvärdering av försöksverksamhet inom rehabiliteringsområdet i två län (1992:12) anför följande:

"Tillgången till relevanta och tillförlitliga data synes ha varit imponerande god, och det är uppenbart att RFV satsat betydande resurser på att bygga upp tillförlitliga databaser som kan utnyttjas för en utvärdering av den egna verksamheten."

I ett mindre positivt utlåtande om AMS utvärdering av starta-eget-bidrag" påpekar en sakkunnig följande om datamaterialet:

"Undersökningen kallas totalundersökning vilket ger en felaktig bild eftersom materialet är ett urval av den underliggande populationen på 16 500 personer och urvalskriteriet är en månad ... Detta urvalskriterium diskuteras inte. Huruvida dessa är representativa för den underliggande populationen diskuteras inte heller. Man borde redovisa enkla statistiska mått för denna grupp och populationen för att kunna ge en bild av hur detta urval är relaterat till populationen ..."

Bedömningarna om metoden varierar. I ett utlåtande om Socialstyrelsens "Pilotstudie av Lex Maria-anmälningar och beslut" skriver en sakkunnig att:

"Metoden med personliga intervjuer bedömer jag som lämplig. Där emot kan så väl urvalet av ärenden ..., som urvalet av respondenter ifrågasättas."

RFV:s utvärdering av försöksverksamheten inom rehabiliteringsområdet får mycket beröm för metodval. En sakkunnig framhåller följande:

"Som tidigare angivits kan utvärderingen uppdelas i en statistiskt deskriptiv del, en ekonometrisk analys samt en mycket utförlig konkret beskrivning av hur arbetet i olika avseenden organiserats och förändrats vid försökskassorna under projektets gång. Enligt min uppfattning är detta en mycket lämplig metod för den aktuella utvärderingen. Min kritik gäller framför allt att den ekonometriska analysen redovisats så styvmoderligt i utvärderingsrapporten."

Ungefär hälften av granskningsrapporterna kritiserar den valda metoden. I en bedömning av NUTEK:s utvärdering (Värdet av regionalpolitiken) anför följande att:

"Det är i första hand de företag som får eller har fått stöd som intervjuats, men även organisationer och statliga verk i stödområdena får säga sin mening. Utomstående företag/personer/organisationer utanför stödområdena får sällan eller aldrig komma till tals om de regionalpolitiska stöden. Det är inte otroligt att de företag som fått stöd eller kommuner i stödområdena, vars sysselsättning och skatteintäkter är beroende av dessa företag, kommer att säga att stöden har varit till större hjälp än vad de egentligen har varit ...".

Vidare påpekar en sakkunnig att:

"Statistiska simuleringar utförs för att estimerar effekterna av företagsstöd på företagets betalningsförmåga på kort och lång sikt (likviditet och soliditet). Man kommer fram till slutsatsen att ett positivt samband finns. Detta är dock inte förvånande, ty om man ger företag stöd i form av pengar så kommer naturligtvis deras finanser att förbättras."

I ett utlåtande om Skolverkets rapport om den nya fältorganisationen skriver den ena sakkunnige följande i sin bedömning av metoden:

"Utvärderingsmetoden är knappast lämplig, eftersom den är alltför journalistisk och impressionistisk. Utvärderarna har "plockat" aspekter ur bakgrundsmaterialet på ett så pass osystematiskt sätt att en utomstående näppeligen kan bedöma om en rättvisande bild".

I några fall framhålls att metoden är beskrivning. I ett utlåtande av Invandrarverkets utvärdering av "Tusentals kontakter – Nätverk för framtiden. Kontaktnätets första tid" framhålls följande om metoden:

"Metoden är en mycket noggrann beskrivning av genomförandet av Kontaktnätet. Man redogör för olika enkätundersökningar som gjorts samt lägger in ett stort antal intervjuer med guider och med flyktingar. Man beskriver också olika myndigheters roll och uppträdande och hur olika frivilligorganisationer har engagerats ... Detta är inte en utvärdering i egentlig mening för att bedöma effektiviteten av åtgärder."

Ett utlåtande om den metod som används i en utvärdering av självhjälpssamsamhet framkommer liknande kommentarer om metoden:

"Utvärderingsmetoden går främst ut på att beskriva vad som utrettats i de olika projekten."

6.3 Slutsatser, kostnadseffektivitet och objektivitet

Vidare framgår av bedömningarna huruvida slutsatserna är välgrundade, om självutvärderingen varit kostnadseffektiv och om utvärderarna förefaller objektiva (tabell 6.4). Om slutsatserna anses väl underbyggda tolkas detta som en positiv bedömning (+) eller (++) beroende på övriga kommentarer om slutsatserna. Om slutsatserna däremot saknar grund eller är bristfälliga har detta tolkats som negativ kritik. I de fall sakkunniga menar att man kunde gjort en bättre utvärdering givet samma kostnad eller att man satsat onödigt stora resurser på utvärderingen (t.ex. onödigt stort antal intervjuer eller datamaterial) har detta tolkats som (-) eller (--). Om utvärderaren eller utvärderarens tillvägagångssätt anses ha godtagbar objektivitet har vi tolkat bedömningen av objektiviteten som positiv. Om den sakkunniga däremot skriver att "Trots att det lyser igenom att den negativa informationen överväger i bakgrundsunderlaget får ej detta fullt genomslag i texten..." har vi tolkat detta som negativ kritik. I de fall sakkunniga inte har bedömt eller kommenterat någon aspekt har vi markerat med (?) i tabellen.¹⁸

¹⁸ De sakkunnigas bedömningar återfinns i bilaga 5. Två granskningar per utvärdering har gjorts. I bilagan benämns dessa bedömning 1 och bedömning 2. I tabellerna motsvarar kodningen benämningen i bilagan. Den första bedömningen i tabellen kommer från bedömningen 1.

Tabell 6.4 Självutvärderingarnas slutsatser, kostnadseffektivitet och objektivitet.

	Slutsatser väl- grundande	Kostnads- effekti- vit	Objektivi- tet
Starta eget-bidrag. (AMS)	(--)(--)	(--)(-)	(-)(0)
Rekryteringsstödet resultat. 1995:1 (AMS)	(--)(+)	(?)(0)	(0)(++)
Bosnier hjälper Bosnier (SIV)	(-)(--)	(?)(--)	(-)(0)
Tusentals kontakter – nätverk för fram- tiden. (SIV)	(+)(-)	(-)(?)	(0)(-)
Minskade hälso- och miljörisker vid an- vändning av bekämpningsmedel (Jord- bruksverket)	(?)(+)	(-)(+)	(?)(+)
Varudistribution. (Konsumentverket)	(-)(-)	(-)(?)	(-)(--)
Värderad service (Konsumentverket)	(-)(?)	(0)(-)	(-)(-)
"Rattfällan." Utvärdering av kurs för fängelsedömda rattfyllerister (KVS)	(0)(+)	(?)(--)	(++)(+)
"Goda råd om fisk" – utvärdering av information (Livsmedelsverket)	(0)(+)	(-)(-)	(+)(+)
Utvärdering av "Startpaket Nyckelhål" (Livsmedelsverket)	(?)(-)	(?)(-)	(-)(+)
"Värdet av regionalpolitiken" (NUTEK)	(--)(0)	(-)(?)	(-)(0)
Försöksverksamhet inom rehabilite- ringsområdet i Malmöhus, Norr- bottens och Blekinge län (RFV)	(+)(++)	(+)(+)	(+)(+)
Rehabiliteringsarbetet på försäkrings- kassan (RFV)	(++)(+)	(-)(--)	(0)(+)
Utvärdering av RAR-systemet (RPS)	(-)(-)	(-)(-)	(-)(0)
Kvalitetsutveckling inom skatteför- valtningen (RSV)	(-)(?)	(--)(-)	(0)(?)
Import support to Guinea-Bissau, Mozambique, Nicaragua and Vietnam. Fungibility and Sustainability.	(--)(-)	(?)(-)	(?)(0)
Import Support to Tanzania (SIDA) (SIDA)	(+)(-)	(?)(-)	(0)(0)
Swedish Support to Cap Verde. (SIDA)	(-)(-)	(?)(?)	(?)(?)
Resultat av Särskilda skogsvårdsinsatser i norra Sverige (Skogsstyrelsen)	(--)(-)	(--)(-)	(--)(-)
Utvärdering av den nya fältorganisationen (Skolverket)	(--)(-)	(-)(?)	(-)(0)
Den nationella utvärderingen av grund- skolan (Skolverket)	(-)(-)	(?)(0)	(?)(+)
Pilotstude av Lex Maria anmälningar och beslut (Socialstyrelsen)	(+)(-)	(?)(0)	(0)(-)
Rapport till Socialstyrelsen 8 sept 1995 (Socialstyrelsen)	(-)(--)	(-)(-)	(0)(0)

¹ I bilaga 2 återfinns en fullständig förteckning över utvärderingarna.

Precis som bedömningarna av de övriga enskilda aspekterna är även låtanden om slutsatser, kostnadseffektivitet och objektivitet i allmänhet mer positiva jämfört med de övergripande bedömningarna. Som framgår av tabell 6.4 anser många emellertid att de slutsatser som dras inte är välgrundade. Flera sakkunniga anser därtill att det har gått att göra en bättre utvärdering med samma resursinsats, i flera fall t.o.m. med exakt samma data.

En del ger försiktiga skrivningar eller diskuterar problemet med objektivitet i självutvärderingar i allmänhet. I de fall de sakkunniga klarar på objektiviteten kan det handla om att utvärderarna har en alltför kritisk inställning till den åtgärd (t.ex. bidragssystem) som utvärderas eller att kritiska synpunkter inte getts samma tyngd som positiva omtyggen. De sakkunnigas försiktighet angående objektiviteten står i viss motsats till deras bedömningar angående hur välgrundade slutsatserna är. I nästan samtliga fall där slutsatserna inte anses välgrundade har slutsatser trots detta dragits som är till myndighetens fördel.

Låt oss ge några exempel på vad de sakkunniga skriver om de slutsatser som dras i utvärderingarna. I en bedömning av Invandrarverkets utvärdering (av projektet Bosnier hjälper Bosnier) anför en sakkunnig att:

"Slutsatserna kring delprojekten är underbyggda i analysen men det saknas en direkt koppling till de tre grundfrågorna som anges som utvärderingens syften. De generella slutsatserna är dock sämre underbyggda och saknar även de en direkt koppling till grundfrågorna. Bl.a. saknar en analys av vad som skiljer den bosniska gruppen, som i utvärderingen beskrivs som ... relativt välutbildad ..., från andra invandrar- och flyktinggrupper. Denna analys borde vara nödvändig för att beskriva hur "majoritetssamhället mött gruppens mobilisering" eller för att "beskriva under vilka förutsättningar stöd till "etisk" mobilisering skall vara effektiv."

I en utvärdering av Lex Maria anmälningar och beslut (Socialstyrelsen) är syftet att ge underlag för konstruktion av en enkät avsedd för en mer omfattande undersökning och värdering av Lex Maria verksamheten. En sakkunnig anför följande om slutsatserna:

"De slutsatser som dras beträffande det sannolika utfallet av en större studie samt beträffande rekommendationen att vänta några år med en sådan studie känns inte tillräckligt väl underbyggda: "En fortsatt,

större och mer representativ undersökning i anslutning till pilotstudien skulle sannolikt ge ett liknande utslag. Utredaren rekommenderar SoS att avvakta något eller några år med en ny undersökning av effekterna av Lex Maria utredningar och annan kvalitetssäkringsarbete." Det kan inte uteslutas att ett annat urval av ärenden och andra respondenter hade givit en åtminstone delvis annan bild och andra slutsatser."

I en bedömning av KVS som avser en kurs för fängelsedömda rattfyllerister menar en sakkunnig:

"Givet den kvasi-experimentella ansatsen och givet den analysmetod som valts (i huvudsak signifikansprövningar av skillnader mellan deltagargrupp och jämförelsegrupp med hjälp av chi2-test) kan slutsatserna som dras i utvärderingarna betraktas som underbyggda."

Om objektivitet skriver en sakkunnig följande om Konsumentverket utvärdering av varudistribution:

"I två avseenden brister rapporten i objektivitet: dels ifråga om urvalet av verksamhetsområden, där rapporten – visserligen öppet redovisat – begränsar sig till de områden där Konsumentverket anses ha goda resultat. Dels ifråga om rapportens upprepande slutsatser att insatserna lett till goda resultat utan att just denna rapport tydligt visar att så är fallet."

Om Skolverkets utvärdering av den nya fältorganisationen skriver en sakkunnig om objektiviteten:

"Trots att det lyser igenom att den negativa informationen överväger i bakgrundsunderlaget får ej detta fullt genomslag i texten genom att man i flera fall refererar enskilda positiva bedömningar och inslag oavsett om dessa är representativa."

Om Jordbruksverkets utvärdering av "minskade hälso- och miljörisker vid användning av bekämpningsmedel" skriver en sakkunnig att:

"Jag har ingen anledning att ifrågasätta objektiviteten hos utvärderaren. P.g.a att det s.k. halveringsprogrammet är kvantitativt uttryckt finns små möjligheter att tolka resultaten på andra sätt."

I ett utlåtande om utvärderingen av kurs för fängelsedömda rattfyllerister anför en sakkunnig följande:

"Utvärderingen av Rattfällan visar prov på godtagbar objektivitet. I den första rapporten diskuteras t.ex. svagheter med en kvasi-experimentell ansats där matchning görs utifrån observerbara egenskaper. ... Även bortfallsanalys har genomförts. Utvärderaren sticker inte heller under stol med att de stora bortfallen kan ha påverkat de uppmätta skillnaderna mellan kursdeltagare och jämförelsegrupp, varför de erhållna resultaten bör tolkas med försiktighet."

När det gäller en utvärdering av särskilda skogsvårdsinsatser i norra Sverige anför en sakkunnig att:

"Utredarens objektivitet skulle kunna ifrågasättas eftersom huvudslutsatsen förefaller bristfälligt underbyggd. Man kan ju inte heller bortse från möjligheten att Skogsstyrelsen, för vilken utvärderingen genomförts, har ett intresse av att läsa om positiva effekter av subventioner till skogsbruk. Å andra sidan skall poängteras att utredaren lagt sig vinn om att uppmärksamma svagheter i utvärderingen, även om dessa senare inte vägts in i den heltäckande bedömningen. Det är inte heller helt lätt att komma runt de svårigheter som här påtalats."

Slutligen vill vi ge några exempel på vad de sakkunniga skriver om kostnadseffektiviteten, dvs. om självutvärderingen i sig verkar ha varit kostnadseffektiv. AMS-rapporten om starta-eget-bidrag kommenteras på följande sätt avseende kostnadseffektiviteten:

... På basis av mina synpunkter ..., vill jag påstå att det är möjligt att åstadkomma bättre resultat med de resurser som rapporten tycks ha använt."

I en bedömning av RFV:s studie av rehabiliteringsarbetet på försäkringskassan skriver en sakkunnig att:

"Enligt ... uppgifter ... har RFV lagt ned betydande resurser på att samla in det datamaterial som ligger till grund för studien. Det är svårt att se att de slutsatser som dras i rapporten skulle motivera ett så stort kostnadsåtagande. Det bör dock beaktas i sammanhanget att datamaterialet antagligen skulle kunna användas för andra syften än den aktuella studien."

Livsmedelsverkets utvärdering av startpaket Nyckelhål (en informationskampanj) kommenteras av en sakkunnig på följande sätt:

"Det som har krävt mest resurser i utvärderingsarbetet är troligen att genomföra drygt 200 telefonintervjuer. I stället för att intervjua så många som möjligt borde kraft ha lagts ned på att få ett representativt urval och att utforma frågorna mer precist...Med ett bättre upplägg av undersökningen kunde troligen mer tillförlitliga resultat ha erhållits samtidigt som antalet intervjuer kunde ha minskats betydligt...Man borde således ha låtit kvalitet gå för kvantitet."

6.4 Några goda exempel på självutvärderingar

Pärulan bland utvärderingarna är en självutvärdering av *Riksförsäkringsverket* (RFV anser 1992:12. "Försöksverksamhet inom rehabiliteringsområdet i Malmöhus, Norrbottens och Blekinge län – en utvärdering"). Båda sakkunniga är mycket positiva till denna studie.

Särskilda medel tilldelades tre län för att vidareutveckla försäkringskassans roll i rehabiliteringsarbetet, dels för köp av rehabiliteringsinsatser. Syftet med försöksverksamheten var att utveckla nya arbetssätt och samarbetsformer i rehabiliteringsarbetet som på sikt kan leda till att människor snabbare än hittills får möjlighet att återgå i arbete efter sjukdom eller skada. Syftet med utvärderingen var att bedöma utvecklingen vid de aktuella försäkringskassorna samt uppnådda resultatet. I utvärderingen har man jämfört de områden där försäkringskassorna tilldelats extra medel med andra områden som inte tilldelats medel. Man har med andra ord gjort ett gediget försök att jämföra en ny situation (nya arbetssätt) med alternativsituationen, dvs. vad skulle ha hänt om man *inte* hade genomfört dessa åtgärder. Eftersom det är omöjligt att veta vad som skulle ha hänt i försöksområden om man inte hade genomfört nya arbetssätt har andra likartade områden, där inga förändringar på försäkringsområdet inträffat, använts som jämförelse. Jämförelsen är sedan bl.a. baserad på regressionsanalys för ohälsotalet där eventuella områdesskillnader avseende kön, ålder, civilstånd, inkomstfördelning m.m. beaktats. En samhällsekonomisk kalkyl har gjorts av försöksverksamheten m.m. Dessutom har en kvalitativ ansats använts där intervjuundersökningar har genomförts med personalen på försökskassorna och deras medaktörer.

De sakkunniga menar att det är positivt att både kvalitativa och kvantitativa metoder har använts. Båda sakkunniga menar att slutsatserna av den genomförda analysen förefaller mycket väl underbyggda. En av de sakkunniga framhåller i sin sammanfattande bedömning att:

"Huvudintrycket av utvärderingen är mycket positivt. Detta beror framför allt på att utvärderingen innehåller en kärna av mera kvalificerad statistisk analys baserad på en mycket relevant databas".

Den andra ger utvärderingen "väl godkänd".

En annan utvärdering som fått positiva kommentarer är SIDA:s "Importstöd till Tanzania". De sakkunniga tog i det fallet ställning till tre utvärderingar som gällde importstöd till tre olika länder. Utvärderingen av importstöd till Tanzania fick positiva omdömen, medan de andra två fick mer negativ kritik. Syftet med utvärderingen är att kartlägga de makroekonomiska effekterna av svenskt importstöd till Tanzania. En fråga gäller hur stödet påverkat tillväxten och tillväxtförutsättningarna i Tanzania. En sakkunnig betonar vikten av ett långsiktigt perspektiv när man skall dra slutsatser om tillväxteffekter. Samtidigt bör man kontrollera för andra faktorer som kan tänkas påverka tillväxten t.ex. konjunkturer, kapitalackumulering, den offentliga sektorns storlek, regleringar och handelshinder. Ett sådant långsiktigt perspektiv finns i utvärderingen av importstöd till Tanzania. I utvärderingen diskuteras hur betydelsen av importstöd empiriskt kan underbyggas genom att t.ex. studera hur pass känslig produktionen är för förändringar i tillgång på utländsk valuta. Importstödet effekter på makroekonomiska variabler behandlas. Båda sakkunniga är positiva till utvärderingen. Den ena sakkunniga skriver att:

"... utvärderingen till stora delar är mycket bra och relevant."

Den andra menar att:

"... detta är en väl genomarbetad utvärdering som uppfyller de mål som stipulerats i direktiven".

Det är intressant att notera att de övriga SIDA-utvärderingarna av importstöd till andra länder (Guinea-Bissau, Mozambique, Nicaragua, Cap Verde m.fl.) kommer till motsatt resultat. I båda dessa utvärde-

ringarna görs inga ambitiösa försök att bedöma de makroekonomiska implikationerna såsom det görs i utvärderingen av importstöd till Tanzania.

En tredje utvärdering som erhållit bra kritik är Kriminalvårdsstyrelsens utvärdering (Rattfällan. Utvärdering av kurs för fängelsedömda rattfyllerister). Syftet med studien var att utvärdera en frivillig kurs för fängelsedömda rattfyllerister. Dels gjordes en analys av i polisregistret registrerade återfall i trafiknykterhetsbrott. De personer som gick kursen jämfördes vad gäller återfall i registrerade trafiknykterhetsbrott inom en 2-årsperiod efter fängelsestraffet med en lika stor kontrollgrupp av individuellt matchade individer som, i likhet med kursdeltagarna, var dömda till fängelse. och där trafiknykterhetsbrott ingick i domen. Matchningsvariabler var ålder, intagningsmånad, kön, tidigare kriminalitet samt huvudbrott i aktuell dom. Dels gjordes en enkätstudie (hos båda grupper) där alkoholvanor samt attityder till, och kunskaper om alkohol och alkohol/trafik analyserades. Den ena sakkunniga bedömer det som försvarbart att utvärderingen inte lagts upp som ett experiment.

"Genom att välja en kontrollgruppsansats samt att utgå från det rimliga antagande att försöks- och kontrollgruppen inte skiljer sig med hänseende på upptäcktsrisk är det faktiskt möjligt att åtminstone delvis studera eventuella faktiska beteendeförändringar efter genomgången kurs."

Variablerna bedömdes som helt tillräckliga för att standardisera jämförelsegrupperna på ett tillfredsställande sätt. Båda sakkunniga påpekar samtidigt med emphasis problemet med selektion. Om det föreligger skillnader mellan grupperna avseende icke observerbara egenskaper, t.ex. skillnader i motivation (kursen är ju frivillig) uppstår ett selektionsproblem. En sakkunnig bedömer dock, trots selektionsproblem, att slutsatserna som dras i utvärderingen kan betraktas som underbyggda.

7 Slutsatser och rekommendationer

Ungefär en fjärdedel av myndigheterna kan inte redovisa några självvärderingar under den sex-årsperiod vi undersökt. Av de självvärderingar som myndigheter genomfört under perioden 1990–95 har vi gjort ett urval av de självvärderingar som myndigheterna redovisat/föreslagit eller som vi bedömt vara de mest sofistikerade. Dessa självvärderingar har sedan granskats av flera sakkunniga. Flertalet av dessa självvärderingar har fått mycket kritik av de sakkunniga.

Det har gått sju år sedan den nya budgetprocessen infördes. En bärande idé i denna reform var just införandet av resultatstyrningen som bygger på att resultat faktiskt mäts och redovisas. Många myndigheter redovisar visserligen enkla mått på verksamhetsutvecklingen i sina årsredovisningar. Dessa ger dock i sig inget underlag för resultatstyrning eftersom de oftast inte ger någon indikation som kan användas i styrningen, t.ex. om varför ändringar i verksamhetens utveckling har skett, hur betydelsefulla verksamhetens effekter eller effektivitet, och i vilken mån och hur det skulle gå att öka effektiviteten. Det är mot bakgrund av de sakkunnigbedömningar som redovisas i denna rapport svårt att undvika slutsatsen att det som borde vara en motor i den nya budgetprocessen ännu inte fungerar.

Olika utredningar har visat att riksdag och regering i liten utsträckning använder sig av myndigheternas självvärderingar. Vår slutsats är att självvärderingarnas fåtal och skiftande kvalitet faktiskt gör att de inte kan användas annat än av experter inom respektive område, vilka kan genomskåda dåligt underbyggda slutsatser.

Under den sex-årsperiod som vi granskat (1990–1995) finns visserligen tecken på en viss förbättring i antal självvärderingar såväl som

i kvaliteten. Ändå förefaller förbättringstakten inte så hög att man kan utgå ifrån att problemet löser sig själv. Under tiden finns också en risk att självvärderingar får dåligt rykte och att intresset för resultatmätning därmed avtar.

I själva verket finns redan tecken på att systemet med självvärderingar håller på att undergrävas. En linje, som flera myndigheter driver, är att kraven på myndigheterna att mäta sina resultat bör sänkas. Ett argument som anförs för detta är att det i många fall skulle vara svårt eller omöjligt att mäta resultat och att resultatmätning då utgör en onödig utgift.

Dessa argument motsägs emellertid av de resultat som redovisas i denna rapport. De flesta sakkunniga som vi har anlitat anser att de skulle gå att göra väsentligt bättre självvärderingar med samma resursinsats. Detta, i samband med att vi ändå valt utvärderingar som tillhör myndigheternas bättre, tyder på att myndigheterna befinner sig långt ifrån den punkt bortom vilken det inte går eller lönar sig att utvärdera längre. I själva verket tyder det på att det mycket väl skulle gå att sänka kostnaden för självvärdering och ändå göra bättre resultatmätningar.

Vår egen bedömning mot bakgrund av kartläggningen och sakkunnigutlåtanden är att det framförallt är tre misstag som gör självvärderingar onödigt dyra och minskar deras styrka:

- A) Utvärderingar läggs upp eller beställs i efterhand. Många självvärderingar skulle bli väsentligt billigare om de planerades i förväg eller i samband med införandet av något program.¹⁹ Som ett exempel kan insamling av data då göras till en automatisk del av den normala administrativa hanteringen. Vidare kan jämförelsegrupper ofta definieras i förväg, men inte i efterhand.
- B) Den dyra delen av en utvärdering är ofta datainsamlingen. I självvärderingarna görs datainsamlingen ofta med för lite tanke p

¹⁹ Ett exempel är uppläggningsenheten av ett stöd till biografier, det tidigare fraktstödet. Införandet av stödet gjordes på ett föredömligt sätt när det gäller att lägga upp verksamheter så att de blir utvärderingsbara. Stödet var vid införandet begränsat till vissa delar av landet, vilket gjorde det lätt att hitta en jämförelsesituation. Fraktstödet var utformat som en försöksverksamhet just för att effekterna skulle kunna analyseras på ett relativt enkelt sätt.

metoden. I många fall skulle det gå att nå en bättre utvärderingen med en mindre, men mer genomtänkt datainsamling.

- C) Val av konsulter som genomför utvärderingar på uppdrag verkar i flera fall ske för lättvindigt, vilket leder till höga arvoden och klena resultat.

Mot bakgrund av våra resultat och observationer föreslår vi ett antal reformer som skulle kunna fungera som katalysator för resultatstyrningen

1. Stärk den interna kvalitetssäkringen

Är det lämpligt att myndigheter utvärderar sig själva? I vår bedömning är det ofrånkomligt. Myndigheter måste kunna genomföra självutvärderingar som en del av det interna arbetet med att förbättra verksamheten. Till viss del kommer dessa självutvärderingar också att användas för att legitimera verksamheten utåt. I flera avseenden bör dock myndigheterna själva ta steg för att förbättra utvärderingskvaliteten.

I första hand kan detta göras genom att rutinmässigt genomföra kvalitetssäkring av självutvärderingar. En sådan kvalitetssäkring bör byggas på sakkunnigbedömningar på ett sätt som beskrivs närmare nedan under punkt 3. Därutöver bör utvärderingskompetens och utvärderingsfunktionen stärkas.

En god kompetens är nödvändig, inte bara för att utföra egna utvärderingar, utan även för att beställa och kontrollera utvärderingar gjorda utanför myndigheten t.ex. av utomstående konsulter.

En stark utvärderingsfunktion är också viktigt, t.ex. i form av ett aktivt stöd och satsning från ledningen. Inom ramen för denna rapport har vi haft telefonsamtal med utvärderingsansvariga inom många myndigheter. Inte sällan påpekar de en brist på uppbackning i ledningen vilket gör att de får föra en kamp på flera håll; dels att förankra själva idéerna till olika utvärderingar och dels att sedan få dessa riktigt utförda.

Man kan inte blunda för att utvärderare inom en myndighet befinner sig i en prekär situation. En utvärdering som avslöjar brister i en del av myndighetens verksamhet har kan leda till en svår sits för utvär-

deraren. Det finns åtskilliga exempel på utvärderare som har känt sig tvungna att sluta efter en sådan händelse. Även fristående konsulter som utför utvärdering på beställning av myndigheten kan självfallet hamna i en liknande konfliktsituation.

2. Förstärk den externa kvalitetssäkringen

Den interna kvalitetssäkringen kan emellertid inte lösa alla problem. I vår bedömning förekommer alltför många blockeringar inom myndigheterna för att man helt skall kunna lita på den interna kvalitetssäkringen. Objektiva utvärderingar uppfattas i många fall uppfattas som ett hot mot delar av en myndighets verksamhet. Det är därför att begära mycket om en myndighet förväntas både genomföra och kvalitetssäkra sådana utvärderingar helt själv.

En extern kvalitetssäkring kan stå på flera ben. Det ligger nära till hands att RRV i högre grad än tidigare kommenterar och bedömer självutvärderingars kvalitet och omfattning hos de myndigheter som granskas.

Därtill kan det vara rimligt att regeringskansliet i högre grad deltar i kvalitetssäkringen av myndigheternas utvärderingar.

Under nästa punkt utvecklas ett praktiskt och enkelt tillvägagångssätt på hur de olika instanser kan genomföra kvalitetsgranskningen.

3. Metod för kvalitetssäkring

Som metod för kvalitetssäkring av utvärderingar förefaller en väg vida överlägsen och även mest använd i olika utvärderingssammanhang. Det är samma metod som vi också har använt i denna rapport. Oberoende och anonyma (om de så önskar) sakkunniga anlitas för att kommentera utvärderingar. Av särskild vikt är att dessa sakkunniga inte väljs av den som genomfört utvärderingen, den som utvärderas, eller närstående till dessa. Vi föreslår därför att någon eller några tar på sig eller ges uppgiften att agera som "Kvalitetssäkringsgrupp" med uppgift främst att förmedla sakkunniga. På begäran skulle Kvalitetssäkringsgruppen välja sakkunniga som passar till en viss självutvärdering,

och återföra de sakkunnigas bedömning till de berörda, under bevarandet av de sakkunnigas anonymitet om de så önskar. Kvalitetssäkringsgruppen skulle kunna anlitas av myndigheter som utför självutvärderingar. Detta kan samtidigt fungera som en kvalitetsstämpel som ger läsaren en signal om tillförlitligheten. De sakkunnigbedömningar som kvalitetssäkringsgruppen förmedlar skall självfallet inte bestå av ett enkelt godkännande, utan kan innehålla krav på förbättringar som så småningom leder till ett godkännande.

Men Kvalitetssäkringsgruppen skulle också kunna anlitas t.ex. av ansvariga departement eller av RRV till hjälp i granskningen av självutvärderingars kvalitet. Det är tänkbart, och även önskvärt, att myndigheter skaffar sig en kvalitetsstämpel på sina självutvärderingar som ett sätt att förebygga den kritik som kan komma om någon annan istället anlitar Kvalitetssäkringsgruppen. Självfallet är det också bra om myndigheter använder sig av kvalitetssäkringen i ett tidigt skede under en självutvärderings framställning och lyckas arbeta in de sakkunnigas synpunkter.

En sådan kvalitetssäkringsprocess fungerar sannolikt också som katalysator som ökar utvärderingskunnandet inom myndigheterna. Våra resultat indikerar att en förbättring av självutvärderingar handlar om kunnande, inte om högre utvärderingskostnader. En sådan kompetens är nödvändig, inte bara för att utföra egna utvärderingar, utan även för att beställa och kontrollera utvärderingar gjorda utanför myndigheten av t.ex. utomstående konsulter. En sådan kvalitetssäkringsprocess underlättar också arbetet betydligt för regeringskansliets tjänstemän som ofta är mottagare av självutvärderingarna. Finns det en kvalitetsstämpel på utvärderingarna kan mottagarna känna sig säkra på att resultaten är relevanta och korrekta.

4. Beställ kompletterande utvärderingar

Ökade kvalitetssäkringskrav kan medföra att myndigheterna gör färre självutvärderingar. Eftersom många myndigheter dessutom inte genomför några självutvärderingar alls, och andra inte har en bra täckning av sin verksamhet, förefaller det finnas behov av externa organ som kan beställa eller genomföra kompletterande utvärderingar på

eget initiativ. I princip finns sådana möjligheter redan idag. I praktiken fungerar de dock inte på ett tillfredsställande sätt.

De ansvariga fackdepartementen kan t.ex. beställa kompletterande utvärderingar. Det är dock ingen hemlighet att fackdepartementen alltför ofta "värnar" om sina myndigheter, anställer personal från sina myndigheter, osv. vilket gör att objektiviteten inte alltid finns. Andra instanser som RRV och Riksdagens Revisorer granskar visserligen myndigheter, men i en särskild form, som hittills inte innefattat att beställa eller genomföra kompletterande självutvärderingar. Slutligen finns även forskare som ibland utför utvärderingar. Ofta har emellertid forskare svårt att få tillgång till de data som myndigheten förfogar över, och även om de får det så styrs deras intresse ofta av frågeställningar i den akademiska forskningslitteraturen, vilket exkluderar många typiska självutvärderingsfrågor.

Förekomsten av kompletterande självutvärderingar kan stärkas på flera sätt. En möjlighet är att ge revisionsmyndigheterna ett tydligare mandat att beställa kompletterande självutvärderingar. Därtill skulle sektorsforskningsmyndigheter kunna få ett tydligare mandat att även satsa på projekt som utgör utvärdering av myndighetsarbete. Ett exempel på sådant är att Kommunikationsforskningsberedningen beställt en kostnads/intäktsanalys av Dennispaketet, som därmed blev den första oberoende utvärderingen av projektet.

Referenser

- Arbetstagarkonsult, 1992. En ny start för regeringskansliet.
Ds 1995:6 ELMA – Åtgärder för att stärka resultatstyrningen.
Finansdepartementet
- Ds 1995:14 Företagsstödet – Vad kostar det egentligen? Rapport till
ESO, Finansdepartementet
- Fölster Stefan m.fl., 1992. Sveriges systemskifte i fara? Erfarenheter
av privatisering, avreglering och decentralisering, Industriens
Utredningsinstitut
- IVA, 1991. Bäst i världen, Vad kan vi lära av världens ledande
företag?
- Riksdagens revisorer, 1991/92. Att ändra statsförvaltning: Styrnings-
problem i ljuset av fyra fallstudier.
- Riksdagens revisorer, 1991/92. Resultatorienterat budgetarbete
- RRV, 1991. Resultatredovisningar. Några erfarenheter från myndig-
hetsförsök.
- RRV, 1993. Resurser och resultat. En analys av innehållet i statliga
myndigheters anslagsframställningar under 25 år. Av Joachim
Schäfer
- RRV 1994:10 Resultat, Verksamhet, Ekonomi – en handledning för
myndigheterna.
- RRV 1995:39 NUTEK:s stöd till företag och forskning.
- Sanders, James R., 1994. The program evaluation standards. 2nd
Edition. The Joint Committee on Standards for Educational Evalua-
tion.

- Seidel, A.D., 1993, Producing Usable Research: A Selected Review
Policy Studies Review, 3.
- Statskontoret, 1990:24. "Att mäta eller inte mäta resultat".
- Statskontoret 1991:32. Myndigheternas förnyelsearbete.
- Statskontoret, 1992:23. Att använda resultatmått med gott resultat.
- Tarschys, D. (1982) För döva öron? Politikern och utvärderaren,
Palmlund, I. (red.), Utvärdering av offentlig verksamhet, Liber.

Abstract

Can Government Agencies Evaluate themselves?

A major reform of the budgetary process in 1988 introduced an entirely new approach to governance of state agencies in Sweden. The central aim of the reform was to allow the government and Parliament to govern by performance management. Government agencies are now expected to evaluate and report their own performance and results.

The new budgetary process has itself been evaluated a number of times since its inception. A frequent finding has been that, while many agencies have become more performance-oriented, their performance evaluations are rarely utilized as a basis for decision-making by management, government or Parliament.

Non-utilization may have many explanations. One fundamental question is whether public agencies' self-evaluations should be used. Do self-evaluations examine relevant questions? Are they objective and trustworthy? Do the evaluations employ valid methods?

This report maps the extent and quality of self-evaluation in 22 government agencies. Self-evaluations are defined as attempts to establish causal links between an agency's efforts and results, that are initiated or conducted by the agency itself. Quality of self-evaluations was judged by independent referees. The self-evaluations chosen for refereeing were generally those the government agencies themselves considered to be among their best. Each self-evaluation was examined by two referees. In total, 40 referees were engaged. All of them were experts and researchers in the various fields with many years of evaluation experience. Public agencies participated in choosing referees, and could veto referees that they considered biased.

The mapping of self-evaluations showed that 10 out of 22 government agencies had conducted fewer than one self-evaluation per 1000 employees over the period 1990–1995. Six agencies had conducted no self-evaluations at all. Among the self-evaluations conducted, simple

surveys of attitudes and judgements dominate. Empirical measurement of effects or effectiveness is relatively rare.

As far as quality is concerned, referees generally supplied detailed motivations of their judgements, that are reprinted as an appendix to this report. A majority of referees are very critical of the quality in the examined self-evaluations. The criticism concerns documentation, choice of method, choice of evaluation question and other aspects. Most referees argued that conclusions drawn in the self-evaluation were not supported by the empirical evidence. Many also felt that much better self-evaluations could have been produced with the same effort and resource limit. Referees often pointed out that self-evaluations had entirely neglected central questions concerning how adequate and efficient the evaluated project might be.

A few self-evaluations were praised by the referees. These generally focused on a central question, employed acceptable empirical methods, and reported conclusions that were supported by the evidence.

In sum, this report indicated that most government agencies' self-evaluations in fact should not be utilized. This implies that the central aim of the new budgetary process – to allow management based on performance and results – has yet to be achieved.

A much improved self-evaluation process is probably necessary if performance management is to be realized. Quality control of self-evaluations ought to be a crucial ingredient. In part responsibility for quality control must lie with the government agencies themselves; in part organisations independent of the self-evaluating agency should participate more in quality control. It would, for example, appear reasonable that the National Audit Office to a greater extent comment and examine quality in government agencies' self-evaluations.

One approach to quality control is most frequently used, and probably most reliable. That is the approach used in preparing this report. Independent and anonymous (if they desire) referees examine self-evaluations. In this process it is particularly important that referees are not chosen by the government agencies themselves. It is therefore proposed that a "Quality Control Group" be organized with the main task of choosing and engaging referees. Upon request the Quality Control Group would choose referees for a certain self-

evaluation and forward the referees' report while maintaining the referees' anonymity.

The Quality Control Group could be engaged by government agencies themselves. This would provide a quality stamp that would signal to readers that an evaluation is reliable. The Quality Control Group could also be engaged for example by the ministries or the National Audit Office.

Finally, this report recommends that evaluations of government agencies' performance more often are initiated by external interests, such as the ministries or the National Audit Office.

When the quality of self-evaluations has been improved it will again be meaningful to consider the issue of non-utilization. Quite possibly improved quality, and improved information about quality, will prove sufficient to make self-evaluations usable.

Bilaga 1

Terminen självutvärdering och projektets uppläggnig

1 Rapportens definition av självutvärdering

Inom utvärderingslitteraturen förekommer en rad olika definitioner av "utvärdering". Vedung (1994) t.ex. använder en vid definition: "Utvärdering är en noggrann efterhandsbedömning av utfall, slutprestationer och genomförande i offentlig politik, vilken avses spela en roll i praktiska handlingssituationer".

Riksrevisionsverkets definition av utvärdering inom ramen för den statliga budgetprocessen är att "bedöma och förklara prestationer och händelseförlopp" i en myndighets verksamhet. Utvärderingen "skärskådar och klargör orsaksförhållanden när det gäller myndighetens prestationer och/eller prestationernas effekter i samhället".

En kommitté som bildats i USA (The Joint Committee on Standards for Educational Evaluation) har tagit fram standards och kriterier (The Program Evaluation Standards) för hur en bra utvärdering bör se ut, både vad avser metodik, användning etc. Dessa standards har blivit accepterade i USA som nödvändiga för välgjorda och relevanta utvärderingar. Inom ramen för deras arbete finns en övergripande definition på utvärdering (evaluation) som är "The systematic investigation of the worth or merit of an object".

Visst kan det skapa förvirring med olika definitioner på utvärdering och det är givetvis eftersträvänsvärt med en enhetlig begreppsapparat. Några utvärderingsansvariga inom olika myndigheter som vi talat med inom ramen för denna studie har också framfört att det händer att

myndigheterna använder sig av en definition samtidigt som regeringen (departementen) menar något annat.

Vi har i huvudsak i denna studie utgått från RRV:s definition. Detta innebär att enkla mått på verksamhetsutveckling, som redovisning av kostnadsutveckling, antal ärenden eller enkla produktivitetmått inte ingår. Många gånger redovisar myndigheter sådana mått i årsredovisningen, där de också fyller en viktig funktion genom att beskriva hur verksamheten utvecklats. De uppfyller dock inte kraven i RRV:s definition av en utvärdering att "bedöma och förklara...", än mindre de andra definitionerna som gavs ovan.

Med självutvärdering avses sådana utvärderingar som syftar till att bedöma eller mäta effekter av, eller effektiviteten i, åtgärder och verksamheter som en myndighet ansvarar för.

I rapporten behandlas endast sådana utvärderingar som avser den *egna* verksamheten. Vi har valt att använda begreppet *självutvärdering* för dessa utvärderingar. Flera myndigheter har som uppgift att genomföra utvärderingar inom olika verksamhetsområden, dvs. utvärderingar som inte bedömer just den egna verksamheten.

I definitionen på självutvärdering lägger vi inga krav på metoder. Det kan avses en enkel bedömning, en attitydundersökning eller en mer sofistikerad empirisk undersökning. I uttrycket självutvärdering ligger att myndigheten själv utför eller beställer utvärderingen som dessutom avser den egna verksamheten. Vi bortser således från utvärderingar som har tillkommit på initiativ av externa aktörer.

2 Projektets genomförande

I det följande beskrivs närmare hur arbetet i princip varit upplagt (enskilda avvikelser kan ha förekommit), och vilka kontakter som förekommit mellan projektet och berörda myndigheter under arbetet med rapporten. De olika stegen i arbetet har också successivt förankrats i projektets referensgrupp som har bestått av Göran Arvidsson, SNS, Peter Bohm, Stockholms universitet, Lars Hultkrantz, Högskolan Dalarna och Uppsala universitet, Marja Lemne, ESO, Ingrid Munck, SCB och Göran Schubert, ESO.

1 Myndighetsurvalet

I kartläggningen ingår totalt 22 myndigheter, dels de 11 största (räknat i antalet anställda), dels ett slumpmässigt urval om 11 av de små och medelstora myndigheterna. Försvarssektorn, universitet och högskolor samt de s.k. stabsmyndigheterna ingår inte i studien.

2 Urvalet av utvärderingar

Samtliga av dessa myndigheter kontaktades per telefon (som regel den utvärderingsansvariga vid resp. myndighet) och ombads skicka in dels en lista över utvärderingar som genomförts under perioden 1990-95, dels ett eller flera ex. på sådana rapporter. I några fall hade vi till följd av tidigare utredningar redan tillgång till sammanställningar över utvärderingsrapporter.

I vissa fall ombads vi skicka en skriftlig framställan, vilket också gjordes.

Utifrån det underlag som myndigheterna tillställde projektet gjorde författarna till denna rapport ett urval av utvärderingar som fyllde kraven på självutvärdering (se ovan). Ursprungligen var avsikten att slumpa fram ett antal rapporter, men detta bedömdes som mindre lämpligt, då ett stort antal av de insända rapporterna inte fyllde kraven. Urvalet kan därför anses ha viss positiv bias. I många fall har de rapporter tagits med som respektive myndighet själv lyft fram som bra exempel på utvärderingar.

3 Urvalet av sakkunniga

Sverige är ett litet land. Många av de personer som har kunskap om ett visst område kan ha varit anlitade av en myndighet, alternativt i den allmänna debatten öppet kritiserat en verksamhet hos en viss myndighet. För att minimera detta problem fördes även i detta fall en dialog med berörda myndigheter.

Samtliga myndigheter (kontaktpersonen) fick ett brev från ESO (under tecknat av Marja Lemne, ESO:s sekretariat) där mellan 5 och 10

tänkbara sakkunniga var förtecknade. Myndigheterna fick ca 2 veckor på sig att höra av sig med ev. invändningar.

4 Inför rapportens publicering

Flera av myndigheterna hade inte redovisat några utvärderingar. Inför slutskrivningarna beslöts därför att skriva till samtliga myndigheter och informera om att rapporten beräknades bli klar före sommaren och i brevet angavs för respektive myndighet dels resultatet av kartläggningen (om myndigheten genomfört någon självutvärdering 1990-95). Vidare angavs i om en eller två av myndighetens utvärderingar skulle ingå i rapporten.

Flertalet myndigheter kontaktade i samband därmed författarna eller ESO:s sekretariat och det förde konstruktiva diskussioner kring detta. I vissa fall har vi efter dessa kontakter gjort förändringar i fråga om kartläggningsdelen.

Bilaga 2

Instruktioner till sakkunnigbedömning

Tack för att du ställer upp som sakkunnig.

Finansdepartementet expertgrupp för studier i offentlig ekonomi (ESO) har inlett ett särskilt projekt för att belysa myndigheters utvärderingsverksamhet.

Sedan den nya budgetprocessen infördes har det blivit allt vanligare att myndigheter själva utvärderar sin egen verksamhet, eller anlitar konsulter för detta ändamål. Det finns nu tillräckligt många sådana självutvärderingar för att kunna bedöma hur väl utvärderingsprocessen fungerar: Hur pålitliga är utvärderingarna? Undersöker de rätt saker? Ger de en god täckning av myndigheters verksamhet? Underkastas de en kvalitetssäkring? Används de av beslutsfattare? Kan användare av utvärderingarna bedöma kvaliteten? Syftet med projektet är att undersöka dessa frågor. I en första fas har vi kartlagt ett tjugotal myndigheters självutvärderingar. I andra fasen vill vi slussa ungefär 15 myndighetsutvärderingar genom sakkunnig-bedömning.

Till varje självutvärdering har valts två sakkunniga. Sakkunniga får vara anonyma om de så önskar. Sakkunnigbedömningarna skall sedan sammanställas för att ge en helhetsbild över självutvärderingarnas kvalitet.

För att underlätta sammanställningen, och för göra sakkunnigbedömningarna någorlunda jämförbara ber vi dig att svara på frågorna nedan.

- Ange ditt namn, telefonnummer och utvärderingens titel. För utbetalning av arvodet vill vi också ha ditt personnummer. Om du har F-skatt kan du även fakturera arvodet (lägg då till Moms). Om inget annat är överenskommit är arvodet SEK 2 000 (exkl. moms).
- Din och andras bedömningar kommer att offentliggöras, förmodligen i form av en bilaga till rapporten. Vi vill helst att bedömnarnas namn också skall vara offentliga. Du får dock själv välja om

du vill vara anonym eller om ditt namn skall offentliggöras. Om du vill vara anonym, ange detta.

3. Beskriv kortfattat utvärderingens syfte.
4. Beskriv kortfattat utvärderingens metod.
5. I det följande ber vi dig att bedöma utvärderingen i olika avseenden. Gör bedömningen utifrån en jämförelse med vad du bedömer som god praxis på området. Beskriv gärna kortfattat problem eller poänger i utvärderingen som bör uppmärksammas.

A) Dokumentation

Är verksamheten som utvärderas, tidigare utvärderingar, data som används, metoder, resultaten och annat av relevans beskrivet på ett sådant sätt att läsaren kan skapa sig en rimlig bild av utvärderingens tillförlitlighet?

B) Data

Är de data som ligger till grund för utvärderingen lämpliga och tillräckligt tillförlitliga?

C) Metod

Är utvärderingsmetoden lämplig? Är den konsekvent och riktigt tillämpad?

D) Slutsatserna

Är de slutsatser som dras väl underbyggda i analysen? Finns andra rimliga tolkningar av analysen som inte beaktas?

E) Kostnadseffektivitet

Anser du att det hade varit möjligt att göra en väsentligt mycket bättre utvärdering med en bättre uppläggning, men - såvitt du kan bedöma - ungefär samma resursinsats när det gäller t.ex. datainsamling, bearbetning, skrivande osv ?

F) Objektivitet

Anser du att tillvägagångssättet i utvärderingen, val av frågeställning, tolkning av resultat m.m. ger uttryck för en godtagbar objektivitet hos utvärderaren?

6. Sammanfattande bedömning och ytterligare kommentarer.

Vi vill också gärna ha ett kort helhetsomdöme av utvärderingen. Om du dessutom vill kommentera någonting som vi inte frågat efter, så gör gärna det.

Bilaga 3

Förteckning över utvärderingar som bedömts av sakkunniga

1 Arbetsmarknadsstyrelsen

- a) "Starta eget-bidrag." En totalundersökning över de personer som avslutade bidragsperioden i oktober-93. Av Ann-Christine Bystedt. AMS.
- b) Uppföljning och utvärdering 1995:1. "Rekryteringsstödet resultat." Av Ann-Christine Bystedt, Ulf Göransson och Anders Harkman.

2 Invandrarverket

- a) Tusentals kontakter -nätverk för framtiden. SIV. Av Erland Bergman.
- b) Bosnier hjälper Bosnier. En utvärdering av självhjälpsverksamhet. 1995-08-20. Av Ragnar Landén.

3 Jordbruksverket

- a) Rapport 1994:18. "Minskade hälso- och miljörisker vid användning av bekämpningsmedel". Resultat av handlingsprogrammet.

4 Konsumentverket

- a) Värderad service. Utvärdering av Konsumentverkets stöd till den kommunala konsumentverksamheten. Rapport 1994/95:16.
- b) Varudistribution. Uppföljning av resultat. Bilaga 1 till Konsumentverkets särskilda rapport 1992/93-1994/95.

5 Kriminalvårdsstyrelsen

- a) VTirapport 371, 1992. "Rattfällan. Utvärdering av ku fängelsedömda rattfyllerister." Inklusiv VTImeddelande n 1993. "Rattfällan. Utvärdering av kurs för fängelsec rattfyllerister: återfall inom en fyra-årsperiod." Av Jan Tö beställd av Kriminalvårdsstyrelsen.

6 Livsmedelsverket

- a) Rapport 3/94. "Goda råd om fisk" - utvärdering av Livsmede kets information till gravida och ammande. Av Birgitta Skeppstedt.
- b) Rapport 45/94. Utvärdering av "Startpaket Nyckelhål". Av Lundqvist.

7 NUTEK

- a) R 1993:43 "Värdet av regionalpolitiken". Inklusiv 5 bilag 1993:44, 46, 48, 49, 50.

8 Riksförsäkringsverket

- a) RFV redovisar 1994:5. "En studie av rehabiliteringsarbet försäkringskassan".
- b) RFV anser 1992:12. "Försöksverksamhet inom rehabilitering rådet i Malmöhus, Norrbottens och Blekinge län - en utvärde

9 Rikspolisstyrelsen

- a) Utvärdering av RAR-systemet vid Huddinge polismyndighet. Redovisning 1993:1 REB

10 Riksskatteverket

- a) "Kvalitetsutveckling inom skatteförvaltningen". Av Klas Ki och Lena Åndell.

11 SIDA

- a) Aid and Macroeconomics: An evaluation of Swedish In Support to Guinea-Bissau, Mozambique, Nicaragua and Viet 1993/2 By Stefan De Vylder
- b) Fungibility and Sustainability. Import Support to Tanz 1993/5. By Joshua Doriye, Howard White and Marc Wuyts

- c) Swedish support Cape Verde (an evaluation of the import support program and the Pkfund). Dec 1993. By Mats Bartling and Allan Gustafsson

12 Skogsstyrelsen

- a) "Resultat av Särskilda skogsvårdsinsatser i norra Sverige." Av P-Å Arvidsson, L. Rådström, P. Stenbeck. Forskningsstiftelsen Skogsarbeten.

13 Skolverket

- a) Den nationella utvärderingen av grundskolan 1987-1994. Synpunkter och empiri om intern debatt och lokal användning. Red: Rolf Lander, Per-Olof Thång och Ulf Torper. Rapport nr. 1994:14, Institutionen för pedagogik, Göteborgs universitet.
- b) Utvärdering av den nya fältorganisationen - en första lägesrapport (daterad 95-11-09).

14 Socialstyrelsen

- a) Rapport till Soc. 8 spt 1995. Intervju med tandläkare, läkare, sjuksköterskor, sjukgymnaster och psykologer om kvalitetssäkring inom vården. Av Skandinavisk opinion ab.
- b) "Pilotstude av Lex Maria anmälningar och beslut."

Bilaga 4

Förteckning över sakkunniga

I samarbete med ESO:s referensgrupp har 40 sakkunniga valts ut. Två av dessa har bedömt tre utvärderingar var. Tre av de sakkunniga har bedömt två utvärderingar var, de övriga har bedömt en utvärdering. I ett fall har en bedömning gjorts av två forskare gemensamt (dvs. tre sakkunniga har bedömt samma utvärdering). För varje utvärdering har bedömningar gjorts av två av varandra oberoende och anonyma sakkunniga.

Några av de sakkunniga (fyra stycken) har valt att vara anonyma. Övriga namn presenteras nedan.

Lars Andersson. Ekon.dr. Ekonomihögskolan. Högskolan Växjö.

Thomas Andersson. Docent. Departementsråd. Näringsdepartementet.

Mahmood Arai. Fil.dr. Institutet för social forskning, Stockholms universitet.

Roger Axelsson. Fil.dr. Nationalekonomiska institutionen. Umeå universitet.

Lars Behrenz. Fil.lic. Ekonomihögskolan. Högskolan Växjö.

Mats Bergman. Fil.dr. Nationalekonomiska institutionen. Umeå universitet.

Jan Björinge. Chef för Folkuniversitetet. Stockholm

Pontus Braunerhjelm. Fil.dr. Industriens utredningsinstitut.

Elisabeth Bylund. Bylund&Lille. Stockholm

Lennart Delander. Fil.lic. Ekonomihögskolan. Högskolan Växjö.

Per-Gunnar Edebalk. Docent. Socialhögskolan. Lund.

Anders Forslund. Docent. Nationalekonomiska institutionen. Uppsala universitet

Ing-Marie Gren. Docent. Beijer Institutet för Ekologisk Ekonomi. Stockholm.

Anette Gröjer. Fil.lic. Högskoleverket. Stockholm

Dan Hedlin. Statistiker. SCB. Stockholm

Leif Heinstedt. Statistiker. SCB. Örebro.

Hans von Hofer. Docent. Kriminologiska inst. Stockholms universitet.
Lars Hultkrantz. Professor. CTS. Högskolan Dalarna.
Lennart Hjalmarsson. Professor. Nationalekonomiska institutionen.
Göteborgs universitet
Mats Kinnwall. Fil.dr. Handelsbanken.
Thomas Küchen. Fil.kand. Revisionsdirektör. Riksdagens Revisorer.
Elin Landell. Civ.ekon. Industriförbundet. Stockholm
Lars Lundberg. Docent. FIEF. Stockholm.
Per Lundborg. Professor. Industriens Utredningsinstitut. Stockholm
Magnus Medelberg. Fil.kand. Avdelningsdirektör. Statskontoret.
Karl-Markus Modén. Fil.dr. Industriens utredningsinstitut. Stockholm.
Ingrid Munck. Docent. SCB. Stockholm.
Mats Nilsson. Statistiker. Västerbottens Läns Landsting
Carl Norrbom. Universitetslektor. Institutet för Kommunal Ekonomi.
Stockholms universitet
Clas Olsson. Ek.lic. Sektionschef. Kommunförbundet. Stockholm.
Roger Pyddoke. Fil.lic. Avdelningsdirektör. Statskontoret
Eva Rabinowicz. Docent. Statens Lantbruksuniversitet.
Svante Sjöberg. Planeringschef. Rådet för Arbetslivsforskning.
Stockholm
Per Skedinger. Fil.dr. Industriens Utredningsinstitut. Stockholm
Eva Wallberg. Fil.kand. Revisionsdirektör. Statskontoret.
Elisabeth Wilhelm. Statistiker. Linköping

Bilaga 5

Sakkunnigbedömningar

Innehållsförteckning

1	Starta eget-bidrag (Arbetsmarknadsstyrelsen)	102
	Bedömning 1	102
	Bedömning 2	106
2	Rekryteringsstödet (Arbetsmarknadsstyrelsen)	115
	Bedömning 1	115
	Bedömning 2	117
3	Bosnier hjälper bosnier – En utvärdering av självhjälpsverksamhet (Invandrarverket)	123
	Bedömning 1	123
	Bedömning 2	127
4	Tusentals kontakter - nätverk för framtiden (Invandrarverket)	131
	Bedömning 1	131
	Bedömning 2	133
5	Minskade hälso- och miljörisker vid användning av bekämpningsmedel – resultatet av handlingsprogrammet (Jordbruksverket)	140
	Bedömning 1	140
	Bedömning 2	145
6	Varudistribution – Uppföljning av resultat (Konsumentverket)	148
	Bedömning 1	148
	Bedömning 2	152

7	Värderad service - Utvärdering av Konsument-	
	verkets stöd till den kommunala konsumentverk-	
	samheten , Rapport 1994/95:16 (Konsumentverket)	157
	Bedömning 1	157
	Bedömning 2	160
8	Rattfällan - Utvärdering av kurs för fängelsedömda	
	rattfyllerister (2 rapporter)	
	(Kriminalvårdsstyrelsen)	170
	Bedömning 1	170
	Bedömning 2	173
9	Goda råd om fisk – Utvärdering av Livsmedelsverkets	
	information till gravida och ammande	
	(Livsmedelsverket)	183
	Bedömning 1	183
	Bedömning 2	187
10	Utvärdering av "Startpaket Nyckelhål" – Intervjuer	
	med restauranger och vidareinformatörer av kost-	
	och hälsoinformation (Livsmedelsverket)	190
	Bedömning 1	190
	Bedömning 2	194
11	Värdet av regionalpolitiken , RA 1993:43	
	(Huvudrapport inkl. delstudier)	
	(Närings- och teknikutvecklingsverket)	199
	Bedömning 1	199
	Bedömning 2	205
12	Försöksverksamhet inom rehabiliteringsområdet	
	i Malmöhus, Norrbottens och Blekinge län –	
	en utvärdering (Riksförsäkringsverket)	209
	Bedömning 1	209
	Bedömning 2	212

13 En studie av rehabiliteringsarbetet på försäkringskassan 1994:5 (Riksförsäkringsverket)	217
Bedömning 1	217
Bedömning 2	221
14 Utvärdering av RAR-systemet (Rikspolisstyrelsen)	229
Bedömning 1	229
Bedömning 2	236
15 Kvalitetsutveckling inom skatteförvaltningen (Riksskatteverket)	243
Bedömning 1	243
Bedömning 2	250
16 Aid and Macroeconomics: An Evaluation of Swedish Import Support to Guinea-Bissau, Mozambique, Nicaragua and Vietnam (SIDA)*	255
17 Fungibility and Sustainability – Import Support to Tanzania (SIDA)*	255
18 Swedish Support to Cape Verde (SIDA)*	255
Bedömning 1	255
Bedömning 2	264
19 Resultat av särskilda skogsvårdsinsatser i norra Sverige (Skogsstyrelsen)	273
Bedömning 1	273
Bedömning 2	279
20 Utvärdering av den nya fältorganisationen – en första lägesrapport (Skolverket)	283
Bedömning 1	283
Bedömning 2	285

21 Den nationella utvärderingen av grundskolan 1987-1994 – Synpunkter och empiri om intern debatt och lokal användning (Skolverket)	287
Bedömning 1	287
Bedömning 2	294
22 Pilotstudie av Lex Maria-anmälningar och beslut (Socialstyrelsen)	301
Bedömning 1	301
Bedömning 2	303
23 Rapport till Socialstyrelsen, 8 september 1995 (Socialstyrelsen)	308
Bedömning 1	308
Bedömning 2	314

* SIDA:s tre utvärderingar har bedömts samlat av de sakkunniga

1 Starta eget-bidrag (Arbetsmarknadsstyrelsen)

Bedömning 1

Rapporten beskriver situationen för de personer och därigenom de företag som beviljats starta-eget bidrag (SEB). Den syftar till att kartlägga dessa personers arbetsmarknadssituation och deras uppfattningar om bidragshanteringen efter avslutad bidragsperiod. SEB syftar till att ge den arbetslöse möjlighet att starta egen näringsverksamhet genom att bidrag utbetalas under uppstartsperioden. För att bli berättigad till bidraget, måste individen vara eller riskera bli arbetslös, eller vara bosatt i ett stödområde. SEB kan då uppfattas som en sysselsättnings-skapande åtgärd.

A Dokumentation

Dokumentationen är bristfällig när det gäller att beskriva åtgärden, dess mål och målgruppen. Det framgår inte heller hur stor andel av de sökande som beviljades SEB. Det redovisas heller inte någon information om hur SEB-tagare skiljer sig från de eventuella bidragssökande som inte fick bidraget och de arbetslösa som inte sökte bidraget.

B Data

Undersökningen kallas totalundersökning vilket ger en felaktig bild eftersom materialet är ett urval av den underliggande populationen på 16.500 personer och urvalskriteriet är en månad, dvs de 519 personer som avslutade sin bidragsperiod under perioden oktober 1993. Detta urvalskriterium diskuteras inte. Huruvida dessa är representativa för den underliggande populationen diskuteras inte heller. Man borde redovisa enkla statistika för denna grupp och populationen för att kunna ge en bild av hur detta urval är relaterat till populationen. Vidare finns det risk att bortfallet är systematiskt korrelerat med SEB-mottagarens arbetsmarknadsaktivitet.

Man kan även ha synpunkter på frågorna i enkäten. Till exempel, de som anger att de skulle ändå starta eget (ca 50 %) borde ha svarat på en följdfråga om SEB kan ha påverkat omfattningen av verksamheten med avseende på sysselsättning eller överlevnadsmöjligheter. Detta vore intressant att veta eftersom 70 % har angett att SEB inte gett dem konkurrensfördelar.

Slutligen kan påpekas att inga andra, i och för sig lättillgängliga, data används för att kunna göra en jämförelse av egenskaperna hos urvalsindividerna, populationen i fråga och de övriga.

C Metod

Den enklast möjliga statistiska metoden har använts. Genomgående används medelvärden i olika grupperingar efter kön, ålder, födelseland och näringsgren. Det framgår inte varför dessa uppdelningar är de enda intressanta. Inga samband har undersökts på ett tillfredsställande sätt. Kort sagt är rapporten ett preliminärt explorativt första steg och inte på något sätt en utvärdering.

D Slutsatser

På sista sidan i slutsatserna skriver författaren: "Vart fjärde företag har mer än fem anställda. Eftersom antalet anställda i företagen överstiger antalet personer som ej längre bedriver verksamhet har SEB

därmed resulterat i ett sysselsättningstillskott som överstiger antalet beviljade SEB."

Denna slutsats saknar grund och förutsätter följande:

1. Att alla (eller en stor del av) SEB-mottagare skulle vara arbetslösa om de inte hade fått bidrag. Detta är i strid mot enkätsvaren att ca 50 % ändå skulle starta företaget utan bidrag och att 70 % anser att SEB inte har gett dem konkurrensfördelar. En enkel korrelation skulle vara på sin plats för att veta hur stor andel av de 50 % som skulle starta verksamheten ändå som finns bland de 70 % som tycker att SEB inte har gett dem konkurrensfördelar. Det ligger också nära till hands att korrelera andelen som utgörs av de som skulle starta utan bidrag och de som vid undersökningen var sysselsatta. Det är inte orimligt att tro att en inte oväsentlig andel av bidragstagarna ändå skulle vara sysselsatta utan SEB.
2. Vidare förutsätter slutsatsen ovan att även en tillräckligt stor andel av de anställda i dessa företag skulle vara arbetslösa utan SEB. Om detta vet vi absolut ingenting.

När det gäller kostnaderna sägs på sidan 1 att bidragskostnaderna är samma när alternativet utan bidrag då SEB motsvarar vad individerna skulle få som arbetslösa. Detta förutsätter att alla skulle fortsatt vara arbetslösa vilket är orimligt.

E Kostnadseffektivitet

På basis av mina synpunkter under punkterna A-D och sammanfattande kommentarerna nedan, vill jag påstå att det är möjligt att åstadkomma bättre resultat med de resurser som rapporten tycks ha använt.

F Objektivitet

Risken för avsaknad av objektivitet är minimal eftersom rapporten enbart är beskrivande. Trots detta finns det exempel på brist på objekti-

vitet. Avsnittet om sysselsättningseffekter saknar saklig grund (se punkt D ovan).

Sammanfattande kommentarer

Denna rapport är ingen utvärdering, varken med avseende på myndighetens verksamhet eller politikens effekter. Som rapportens författare själv uppger är syftet att ge en bild av verksamheten.

En utvärdering bör analysera effekterna av åtgärden mot bakgrund av målet och implementeringen av politiken. Man bör då skilja på utvärdering av politik, dvs SEB som den har formulerats och den berörda myndighetens implementering av politiken.

En central brist i rapporten är att jämförelsegrupper saknas. Det är centralt att ha alternativet helt klart för sig innan man kan ta ställning till åtgärdens effekter. På basis av rapporten kan man omöjligen bilda sig någon som helst uppfattning om hur det hade gått för denna grupp om de inte hade fått SEB. Att definiera en jämförelsegrupp för att relatera situationen för personer i åtgärd till vad som kan inträffa utan åtgärd är inte oproblematiskt eftersom personer i åtgärd skiljer sig i många avseenden utöver det faktum att de till skillnad från andra personer är i åtgärd. Detta problem kallas för selektionsproblemet och diskuteras i all seriös utvärdering.

Om man håller sig till det som har redovisats i rapporten har många för mig något förvånande enkätsvar lämnats oanalyserade. Till exempel att 86 % inte följdes upp av arbetsförmedlingen, att 66 % inte deltagit i nyföretagarutbildning trots att, som det står i rapporten: "Enligt den antagna strategin för starta eget-bidragen skall i princip alla deltaga i en nyföretagarutbildning omfattande två veckor." Förklaringen kan vara att denna strategi antogs av AMS styrelse hösten 1993 (sidan 1). Det vore intressant att veta vad som gällde innan. Om få har fått utbildning eller handledning, finns det risk för att starta eget-bidragen har varit rena transfereringar för en stor grupp, vilket inte har varit målet med politiken.

Avslutningsvis vill jag hänvisa till en god utvärdering av SEB i Norge av Jon Reiersen, "Arbetsledige etablerere" (Stiftelsen for samfunns- og næringslivsforskning i Norge, 1995) vilken jag finner relevant att jämföra med rapporten från AMS.

Bedömning 2

Undersökningens syften uppges vara att (1) "ge en bild av de personer och därigenom företag som beviljats bidraget", (2) "följa upp hur verksamheten utvecklats efter avslutad bidragsperiod" samt (3) "studera synpunkter på bidragshandlingen ur bidragstagarnas synvinkel".

De uppgifter som rapporten bygger på har inhämtats med hjälp av en enkät till samtliga personer som erhållit starta eget-bidrag och som avslutade sin bidragsperiod i oktober 1993.

Vi kommer först att kort kommentera rapporten mot bakgrund av vart och ett av de tre angivna syftena. Därefter lämnas avslutningsvis omdömen i ett antal avseenden i punktform.

1 En bild av de personer och företag som beviljats bidraget

I den här delen av rapporten presenteras enkätsvar fördelad på två eller flera svarsalternativ för ett stort antal undersökningsvariabler med anknytning till det första av de angivna syftena. Med hjälp av korstabeller redovisas för flertalet undersökningsvariabler svaren på respektive svarsalternativ uppdelade först efter kön, sedan efter ålder osv. enligt den schematiska översikten i tabell 1 nedan.

En avsikt med denna del av studien har uppenbarligen varit att undersöka om det finns skillnader vad gäller svarens fördelning med avseende på sådana bakgrundsvariabler som kön, ålder, ursprung och verksamhetstyp (för ett par undersökningsvariabler finns uppdelningar på ytterligare ett par bakgrundsvariabler). *Det framgår däremot inte varför man har ansett det vara intressant att kartlägga dessa skillnader.*

Vi menar att de insamlade uppgifterna kunde ha kommit till från utvärderingssynpunkt betydelsefull användning, om de hade utnyttjats på ett annat sätt. Till detta återkommer vi i avsnitt 2 nedan.

I texten efter respektive tabell i rapporten finns utsagor om samvariationer mellan bakgrunds- och undersökningsvariabler, som endast tycks vara baserade på en okulärbesiktning av tabellerna. Med lite fantasi kan emellertid de som besvarat enkäten betraktas som ett

slumpmässigt stickprov, eftersom det rör sig om personer som fått starta eget-bidrag under en viss tidsperiod (personer som avslutade sin bidragsperiod i oktober 1993). Med det sätt att utnyttja informationen som man har valt, borde man därför ha använt chi2-test för att undersöka om skillnader i svar med avseende på bakgrundsvariabler kan bedömas vara slumpmässigt uppkomna, eller om de tyder på verkliga olikheter.

Tabell 1

Undersökningsvariabler med två eller flera svarsalternativ	Bakgrundsvariabler			Verksamhetstyp
	Kön	Ålder	Ursprung ¹	
Sysselsättning före SE ²	*	*	*	
Arbetslöshetstid före SE	*	*	*	
Vidareutbildning	*	*	*	
Vidareutbildningens längd				
Utbildningens betydelse för inriktningen på företaget	*	*	*	*
Praktisk erfarenhet från verksamhetsområdet	*	*	*	*
Verksamhetstyp	*	*	*	*
Företagsform	*	*	*	*
Ensam/kompanjon		*		*
Drivit företag tidigare	*	*	*	*
Deltagit i nyföretagarutbildn.	*	*	*	*
Utbildningens längd				
Upplevd nytta av utbildningen	*	*	*	*
Kombinerat SEB med ALU	*		*	
Kontakt med konsult före SE	*	*	*	*
Vilken slags hjälp av konsult				

¹ Svensk/Ej svensk

² Starta eget

2 Verksamheten efter avslutad bidragsperiod

Den övergripande frågan som man sökt svar på i den här delen av studien kan formuleras så här: Hur har det gått för de personer som erhöll starta eget-bidrag och för de verksamheter som startades?

Med lite god vilja kan denna del därför betraktas som en *måluppfyllelseutvärdering*. Det är visserligen inte så att observerade resultat i relevanta fall jämförs med uppställda mål (kanske har det inte funnits några sådana), men man har otvivelaktigt ställt målrelaterade frågor. Av den anledningen menar vi att detta avsnitt borde ha inletts

med något som saknas helt och hållet, nämligen en analys av målen för den studerade verksamheten.

Metoden är densamma som beskrevs i föregående avsnitt: korstabeller med försök att på basis av enkätsvarens fördelning spåra samband mellan bakgrundsvariabler och "hur det har gått". En schematisk översikt finns i tabell 2 nedan. *Liksom i den föregående delen av studien saknar vi formella hypotestester. Sådana framstår här som extra viktiga, eftersom man har sökt svar på en från utvärderingssynpunkt mycket betydelsefull fråga, nämligen om olikheter i hur det har gått kan förklaras av skillnader vad gäller bakgrundsvariabler.*

Studiens resultat i det här avseendet skulle ha kunnat användas för ställningstaganden eller åtminstone en diskussion rörande kriterier för beslut rörande beviljande av starta eget-bidrag för att bidragen skall bli så verkningsfulla som möjligt. *Den sortens analys lyser emellertid med sin frånvaro.*

Tabell 2

Resultatvariabler (Beroende variabler)	Bakgrundsvariabler (Oberoende variabler)							
	Kön	Ålder	Ursprung	Verksam- hetstyp	Tid. er- farenhet från omr.	Drivit företag tidigare	Deltagit i nyföre- tagarutb.	Konsult- kontakt före start
Verksamhet i företaget idag?	*	*	*	*	*	*	*	*
Arbetar heltid/deltid	*	*	*	*	*			
Deltidens omfattning								
Övrig sysselsättning								
Företagets omsättning i nuläget								
D:o för heltidsföretag som startats av en person	*	*	*	*	*			
Orsak till verksamhetens upphörande	*	*	*		*			
Sysselsättning idag	*	*	*					
Antal anställda vid företagsstarten	*							
Antal anställda i dagsläget	*							

Vi menar, kanske något tillspetsat, att man med fördel kunde ha avstått från det mesta av de beskrivningar (tabeller med tillhörande text) som svarar mot de delar av studien som vi har beskrivit i detta och föregående avsnitt. *I stället kunde man, som ett huvudsyfte för studien, ha tagit sikte på att med hjälp av multivariat analys studera samband mellan utfall vad gäller resultat i olika avseenden (beroende variabler) och ett antal förklaringsvariabler (oberoende variabler).*

Aspiranter på förklaringsvariabler i de modeller, som då skulle ha specificerats, är förstås bakgrundsvariablerna i tabell 2 ovan, men också ytterligare några av undersökningsvariablerna i tabell 1 (en del av dessa är bakgrundsvariabler i tabell 2). För att ta ett exempel, så kunde man ha skattat hur värden på förklaringsvariabler har påverkat sannolikheten att vara verksam i det nystartade företaget efter ett år (den första resultatvariabeln i tabell 2).

Med analyser av antytt slag kunde studien ha gett underlag för ställningstaganden och beslut rörande hur starta eget-bidraget bäst kan komma till användning. Analyser med syfte att ta fram sådant underlag skulle ha varit mycket värdefulla, bland annat med tanke på en av observationerna i den tredje delen av studien.

3 Synpunkter på bidragshanteringen

Den tredje delen av studien kan, med blygsamma anspråk, betecknas som en *implementeringsutvärdering*, dvs. en undersökning som gäller genomförandeprocessen – hanteringen av starta eget-bidraget. Vi tycker emellertid att den har getts en alldeles för snäv inriktning. För det första har synpunkter endast inhämtats från bidragstagare. För det andra har dessa bara fått lämna två uppgifter som har att göra med arbetsförmedlingens handläggning av bidraget; den ena gällde förmedlingens uppföljning av den nystartade verksamhetens utveckling, medan den andra var en betygssättning av förmedlingens handläggning.

Det hade varit motiverat med en betydligt bredare och djupare kartläggning av problem och svårigheter i samband med handläggningen på förmedlingskontoren och, naturligtvis, av de förhållanden och orsakssamband som ligger bakom olika problem och svårigheter. Det är viktigt att en implementerings- eller genomförandeutvärdering inriktas just mot att beskriva och förklara orsaker till *problem*. Det gäller inte

minst i det här fallet, eftersom handläggningen av beslut om starta eget-bidrag ju har varit föremål för debatt inom Arbetsmarknadsverket och för kritik från utomstående. *Sammanfattningsvis menar vi att det hade varit angeläget att hämta in åsikter från olika aktörer om hur verksamheten fungerat och om hur den skulle kunna förbättras.*

Två frågor ställdes om bidragstagarnas synpunkter på bidraget. Den ena gällde om man ansåg att bidraget hade gett några konkurrensfördelar. Hälften av de tillfrågade menade att så inte var fallet. *Det är den enda kunskap om denna inte helt oväsentliga fråga som rapporten förmedlar. Frågan är väl om just bidragstagarnas åsikt kan tillmätas så stor betydelse.*

Den andra frågan innebär att det finns ett visserligen litet, men betydelsefullt inslag av *effektutvärdering* i studien. Med den här frågan har man, vilket är nödvändigt för den typen av utvärdering, jämfört starta eget-bidraget med ett annat handlingsalternativ, som i det här fallet är "inget bidrag". Den fråga som man har sökt svar på är i hur många fall företag har startats som en *effekt*, ett *resultat* av att det har funnits ett bidrag att tillgå. Utvärderarens slutsats är: "Nära hälften av dem som beviljats starta eget-bidrag skulle ha startat företaget utan det stöd som bidraget ger." Tyvärr finns det ingen jämförelse mellan hur det har gått för dessa företag och för dem där bidraget har haft betydelse.

Den nyss citerade meningen är det enda som sägs i rapporten som kommentar till svaren på den aktuella frågan. Det tycker vi är mildt sagt överraskande. Det är alltså så att av enkätsvaren att döma har starta eget-bidraget varit förknippat med en betydande så kallad dödviktseffekt – ett stort antal av de aktuella företagen skulle ha startats även utan bidrag.

Enligt rapporten räknade man för budgetåret 95/96 (18 månader) med en total kostnad för bidraget på drygt 1 300 miljoner kr. Enligt studien kan man förvänta sig att ungefär halva det beloppet inte ger någon effekt alls på nystartandet av företag. Det ska observeras att för de personer det gäller, som alltså skulle ha startat egna företag iallafall, finns det ingen motsvarande offentlig-finansiell besparing i form av minskade utgifter för arbetslöshetsersättning.

Av studien att döma kan man därför tala om ett "politikmisslyckande". Starta eget-bidraget är ett svårhanterligt ekonomiskt politiskt medel. Rimliga krav på kostnadseffektivitet i tillämpningen förutsätter in-

gående analyser av varje övervägt projekt. Prövningen av frågan om ett bidrag ska utbetalas eller inte förutsätter en förhandling där givare och mottagare har tillgång till ungefär samma information – samma bedömningsunderlag. Utbetalning av bidrag i onödan medför en omotiverad belastning för de offentliga finanserna. I en ekonomi som präglas av höga skatter och ett statligt budgetunderskott måste det ses som ett allvarligt politikmisslyckande.

Med tanke på problemets betydelse anser vi, att det i rapporten borde ha varit föremål för en ingående diskussion inkluderande tankar kring vad som kan göras för att minska omfattningen av onödiga utbetalningar, dvs. minska dödviktseffekten. Som underlag borde man även i det här fallet ha använt en multivariat metod för att analysera de insamlade uppgifterna.

Eventuellt kunde en mer analytisk hållen del av rapporten ha tagit sin utgångspunkt i just det förhållandet att man tycker sig ha funnit belägg för en betydande dödviktseffekt. Detta kunde ha använts som motivering för en noggrann analys, med tillförlitlig statistisk metod, av det slag som antydde i avsnitt 2 ovan. Det framstår ju nämligen som utomordentligt viktigt att försöka klarlägga vad som påverkar sannolikheten för att utbetalning av bidrag ska ge det resultat som eftersträvas. Den sortens kunskap är nödvändig för en kostnadseffektiv användning av starta eget-bidraget.

4 Några omdömen i punktform

A Dokumentation m.m.

Studiens syfte har uppenbarligen varit att *beskriva* (1) de bidragstagare och de nystartade företag som omfattades av undersökningen, (2) situationen ett år efter bidragsperiodens slut samt (3) bidragstagarnas synpunkter på bidraget och arbetsförmedlingen. De data som insamlats och sättet att presentera dem ger rimligt tillförlitliga *beskrivningar*.

B Data

Insamlade uppgifter skulle, om de hade utnyttjats på andra sätt, ha kunnat ligga till grund för ställningstaganden till och bedömningar av hur *verkningsfullt* bidraget har varit. Undersökningens uppläggning medger däremot inte någon uppskattning av starta eget-bidragets *effekter*. Vi återkommer till detta under punkt D nedan.

Den information som inhämtats för det under punkt A angivna tredje syftet är helt otillräcklig för en utvärdering av besluts- och genomförandeprocesserna (implementeringsutvärdering).

C Metod

Även om beskrivningarna enligt punkt A är tillförlitliga, så är de slutsatser som man drar om *samband* (t.ex. mellan bakgrunds- och undersökningsvariabler i tabell 1 ovan) inte tillräckligt väl underbyggda. Eftersom de bidragstagare som tillfrågats kan betraktas som ett slumpmässigt urval, borde det ha gjorts statistiska tester gällande förekomsten av samband.

D Slutsatserna

I stället för att basera slutsatser om samband på inspektion av tabeller, kunde man med fördel ha använt multivariat metod för att undersöka hur flera variabler *tillsammans* (bland annat bakgrundsvariabler i tabell 2 ovan) t.ex. har påverkat resultatet efter ett år. Studien tycks egentligen knappast ha tagit sikte på att dra några slutsatser rörande bidragets inriktning, hur väl det fungerar som ett arbetsmarknadspolitiskt instrument m.m. Trots detta talas det om t.ex. sysselsättningseffekter vilket är tvetydigt och lätt kan leda till förhastade slutsatser om *orsakssambanden* bakom observerade utfall. Eftersom man har studerat ett och endast ett förlopp, nämligen det som är förknippat med ut-

porten hade vunnit mycket på om man åtminstone hade fört en del diskussioner baserade på ett strikt alternativtänkande.

E Kostnadseffektivitet

För att kunna göra en sådan implementeringsutvärdering som antydde i början av avsnitt 3 ovan, hade det krävts mera tid och andra resurser än vad som gått åt för just den delen av den aktuella studien. Undersökningen i övrigt borde ha kunnat ge för beslutsfattare intressantare information med ungefär samma resurser som man har haft till förfogande.

F Objektivitet

Frånvaron av en kritisk granskning av starta eget-bidraget och analyser baserade på viktiga frågeställningar beror, enligt vår bedömning, på andra faktorer än bristande objektivitet hos utvärderaren.

2 Rekryteringsstödet resultat (Arbetsmarknadsstyrelsen)

Bedömning 1

Rapporten "Rekryteringsstödet resultat", syftar till att utvärdera rekryteringsstödet om än syftet inte är helt klart i rapporten, som i första hand är en beskrivning av rekryteringsstödet snarare än en utvärdering av detta stöd.

Det är svårt att utvärdera en åtgärd som till synes kan *lämnas för en arbetsökande som är inskriven vid arbetsförmedlingen och som anvisas till en anställning av förmedlingen om den sökandes sysselsättningssituation inte kan lösas på något annat sätt (förordningen (1986:414) om rekryteringsstöd)*.

Frågan här är vilka grupper som åtgärden riktar sig till? Hur många skall omfattas (behöver) denna typ av åtgärd? Vad kostar åtgärden? Kan vi dessutom verkligen på ett vetenskapligt sätt ställa olika arbetsmarknadsåtgärder mot varandra?

Rapportförfattarna ser det som att det är i första hand de långtidsarbetslösa som åtgärden riktar sig till. Av studien att döma emellertid är det just de som varit långtidsarbetslösa före åtgärden som i större andel än andra grupper är tillbaka till arbetsförmedlingen efter åtgärden. Därtill är inte de långtidsarbetslösa (enligt traditionell definition) en dominerande grupp bland de som fått rekryteringsstöd. Om man till den traditionella definitionen lägger till alla de som är del av rundgången i systemet blir bilden något annorlunda. De som klarat sig bäst på arbetsmarknaden efter åtgärden är individer som arbetsgivaren hade anställt oberoende av rekryteringsstödet. Ändock drar författarna slut-

satsen att rekryteringsstöd fungerar väl. Slutsatsen bygger på en jämförelse med andra åtgärdstyper. Problemet med denna jämförelse är att populationerna inte är lika.

Datansamlingen och bearbetningen är till synes ambitiös och bygger på uppgifter från en registerstudie samt uppgifter från en enkät (utförd av SCB), riktad till ett stort antal arbetsgivare, med ett urval omfattande anställda med rekryteringsstöd som också återfinns i registerstudien. Otvetydigt finns här en stor detaljrikedom och ett utomordentligt underlag. Dessutom torde materialet vara tillförlitligt. Den beskrivande delen (i princip sid 1-30) ger också en god bild av rekryteringsstödet och dess dimensioner.

Problemet med rapporten är resultatbedömningen. Eftersom tyngdpunkten i rapporten ligger på beskrivning är själva resultatbedömningen tunn, varför det inte heller går att dra några direkta slutsatser av denna rapport.

Författarna tycks till del vara medvetna om problemen eftersom de själva avslutar studien med en del "intryck".

Tyvärr går det inte heller utifrån studien att säga något om dess relation till tidigare utvärderingar eftersom tidigare studier och dessas resultat enbart redovisas i termer av referenshänvisningar.

Intentionerna har säkert varit mycket goda när det gäller denna rapport. Den har förmodligen kostat en hel del med tanke på SCB's engagemang. Rapporten har emellertid utmynnat i en beskrivning med svaga slutsatser, ändock drar författarna slutsatsen att "utvärderingen visar på förhållandevis goda resultat i form av arbete efter rekryteringsstödet". Då har man ju missat en hel del. T.ex. har man nått de som avsågs över huvud taget? Här har man misslyckats. Är tillvägagångssättet för jämförelsen mellan olika åtgärder korrekt. Författarna tar lite för lättvindigt på faktum att det är olika populationer man jämför. Vad har rekryteringsstödet kostat och är det kostnadseffektivt relativt andra åtgärder? De senare problemen har helt försumrats. Listan på frågetecken kan göras lång.

Sammanfattningsvis kan dock sägas att rapporten är i huvudsak en beskrivning av rekryteringsstödet och som sådan är den god. Som en resultatbedömning är rapporten svag och likaså är de resultat som erhålls svaga.

En utvärdering borde innehålla en noggrannare diskussion av tidigare utvärderingar inom området och vad denna utvärdering har att tillföra.

En beskrivning av rekryteringsstödet. Denna bit är bra.

En beskrivning av den metod som används och valet av densamma. Rapporten har inte motiverat något i detta hänseende. T.ex. varför har man utelämnat frågor som rör rekryteringsstödet kostnader? Är det inte viktigt att stödet är kostnadseffektivt? Eller är kostnaden oväsentlig?

De olika bitarna är inte väl sammanfogade i denna rapport. Rapporten består av tre separata delar som den är.

Huvudproblemet är emellertid resultatbedömningen. Den mest svaga punkten är att de populationer man jämför är olika och att man i sådana fall borde justera för selektion.

Bedömning 2

1 Utvärderingens syfte

Syftet med föreliggande utvärdering, utförd av AMS utredningsenhet, är att bedöma effekterna av det rekryteringsstöd som i något varierande former ingått i AMS åtgärdsarsenal sedan 1984. Stödet är en form av lönesubvention till företag som anställer arbetssökande där sysselsättningssituationen för dessa inte kan lösas på annat sätt (långtidsarbetslösa enligt förordningen 1986:414 om rekryteringsstöd). Enligt uppgift i utvärderingen har arbetsmarknadsverket under de fem senaste budgetåren spenderat 4,5 miljarder kronor på åtgärden. Utvärderingen avser effekterna på en grupp personer (hädanefter kallad försöksgruppen) som fått anställning med rekryteringsstöd och där stödet upphörde under andra halvåret 1993.

Utvärderingen för ingen explicit diskussion om lämpliga mått på effekterna, men den effekt som diskuteras i rapporten huvudsakligen handlar om försöksgruppens arbetsmarknadssituation ungefär ett år efter avslutad åtgärd.

2 Utvärderingens metod

Rapportens analys faller metodmässigt inom tre huvudområden (som motsvaras av tre avsnitt i rapporten). I det första avsnittet utnyttjas registerdata från arbetsmarknadsverkets bidragsregister och den händelsedatabas (HÄNDEL) som byggts upp vid AMS under 1990-talet för att beskriva försöksgruppen, de företag som anställt dem och deras arbetsmarknadsstatus tidigt 1995. Detta avsnitt ger en utförlig presentation av åtgärden och försöksgruppen, men innehåller inget material som med bästa vilja i världen kan kallas en utvärdering.

I det andra avsnittet presenteras resultaten från en enkät till de företag som anställt ett stratifierat urval (knappt en tredjedel av den totala populationen, eller 3083 personer) av försöksgruppen. Två frågor i enkäten (som syftar till att fastställa om företaget skulle rekryterat någon, och i så fall summa person, även utan rekryteringsstöd) är utan tvekan relevanta för att bedöma åtgärdens s.k. dödviktskostnader, dvs i vilken utsträckning subventionen enbart resulterat i en transferering till berörda företag utan reala effekter.

Den tredje huvuddelen av rapporten innehåller jämförelser mellan försöksgruppen och två kontrollgrupper, en bestående av personer som lämnat andra typer av åtgärder ungefär samtidigt som försöksgruppen, en som erhållit anställning utan rekryteringsstöd under den period personerna i försöksgruppen anställdes med rekryteringsstöd. Jämförelserna avser i bägge fallen gruppernas arbetsmarknadsstatus tidigt 1995 och utnyttjar data från HÄNDEL. Jämförelsen med andra åtgärdsdeltagare genomförs som en jämförelse mellan de två grupperna rakt av – inga försök görs för att kontrollera för andra faktorer. Jämförelsen med personerna som anställdes utan rekryteringsstöd genomförs genom skattning av logit-modeller, och kontroll görs för ett antal observerbara personliga karaktäristika.

3 Bedömning av utvärderingen

A Dokumentation

Den dokumentation som presenteras i rapporten är fullt tillräcklig för att läsaren skall kunna bilda sig en (väl grundad) uppfattning om ut-

värderingens tillförlitlighet. Möjligen saknar jag en diskussion om resultatens av tidigare utvärderingar (referens ges emellertid) och (framför allt) en fylligare motivering av de valda metoderna.

B Data

Den databas, HÄNDEL, som ligger till grund för det mesta av utvärderingen är en förnämlig databas, som AMS bör få en eloge för att ha byggt upp. Det stickprov som dragits för enkätundersökningen bör vara tillräckligt stort för att ge en tillförlitlig bild av de undersökta frågeställningarna.

C Metod

De två delar av rapporten som redovisar utvärderande material använder sig av skilda metoder (se *avsnitt 2*). Förutsatt att företagen besvarar de frågor som ställs i enkäten någorlunda sanningsenligt, bör den ge en god bild av de *dödviktskostnader* som är förknippade med rekryteringsstöd. Det finns förvisso andra tänkbara tillvägagångssätt (se t.ex. genomgången i Anxo, 1994) för att belysa detta problem, men den valda metoden är i högsta grad rimlig. Det faktum att 36 % av företagen uppger att de ändå skulle ha rekryterat, och att 56 % av dessa företag i så fall skulle ha rekryterat samma person, antyder att dödviktskostnaderna inte är försumbara. Det verkar också rimligt att tro att man här skatter en undre gräns för problemets omfattning: om något, kan företagen förväntas vara benägna att (t.ex. av strategiska skäl) ge intryck av att deras beteende faktiskt påverkats av stödet.

De två delundersökningar där försöksgruppen jämförs med olika kontrollgrupper är även de av så pass olika slag att de bör diskuteras var och en för sig. Den första delundersökningen avser, som redan noterats, en jämförelse med personer som lämnat en annan arbetsmarknadspolitisk åtgärd under den period försöksgruppens rekryteringsstöd löpte ut. Jag har två huvudsynpunkter på denna undersökning: för det första är det oklart varför man valt att inte kontrollera för individkaraktäristika genom skattning av t.ex. en logitmodell av det slag som skattas när man jämför med den kontrollgrupp som fått

arbete utan rekryteringsstöd. För det andra innebär valet av kontrollgrupp att ingen jämförelse görs med dem som varit arbetslösa utan att delta i någon åtgärd. Det är naturligtvis intressant att jämföra utfallet av olika typer av åtgärder, men det vore också intressant att jämföra utfallet för en kontrollgrupp utan åtgärdsdeltagande.

Den andra delundersökningen, som alltså är en jämförelse med en kontrollgrupp som anställts utan rekryteringsstöd, verkar vara utförd enligt konstens alla regler. Jag har därför inga kritiska synpunkter på metoden i denna att redovisa.

D Slutsatserna

De slutsatser som dras i rapporten har genomgående stöd i den redovisade analysen. Vad jag möjligen tycker mig sakna är en utförligare diskussion av resultaten. För att ta ett exempel: i jämförelsen med dem som anställts utan rekryteringsstöd och som varit arbetslösa innan, konstateras att sannolikheten att ha ett arbete i början av 1995 är högre för dem som anställts med rekryteringsstöd. Detta kan tolkas som att stödet haft en god effekt. En fråga man som läsare ställer sig är emellertid om detta kan återspegla att de som anställts med rekryteringsstöd i högre grad än andra erhållit tillsvidareanställningar (vi vet från uppgifter i AKU att en betydande del av nettotillskottet i selsättning under denna period avsett tillfälliga eller tidsbegränsade anställningar). Det skulle sålunda ha varit intressant med en diskussion om i vilken grad utfallet återspeglar att rekryteringsstöd i högre grad riktats mot tillsvidareanställningar.

E Kostnadseffektivitet

Det är naturligtvis mycket svårt att som utomstående avgöra om en bättre rapport kunnat skrivas med samma resursinsats. Möjligen har jag en känsla av att en större del av utrymme (och tid) skulle ägnats åt den del av utvärderingen som faktiskt är en utvärdering (delarna två och tre som jag kallat dem ovan) och mindre utrymme åt den inledande deskriptiva delen. Denna innehåller förvisso intressant information om hur rekryteringsstödet använts, men informationen är svår

att använda som en grund för en utvärdering av verksamheten. Förmodligen är kostnaden för att ta fram uppgifterna i denna del låg, eftersom den är baserad på statistik som ändå tas fram vid AMS. Jag skulle emellertid, som läsare, ha föredragit att snabbare "komma till skott" vad gäller utvärderingen av verksamheten. Detta skulle dessutom ha gjort det möjligt att ägna mer utrymme åt en diskussion av resultaten inom ramen för en rapport av samma omfång som den föreliggande.

F Objektivitet

Rapportens utformning ger som helhet uttryck för en godtagbar objektivitet hos utvärderaren. Jag har svårt att se att uppläggning eller diskussion av resultaten tyder på en medveten ambition att förtiga uppgifter som skulle få verksamheten att framstå i en sämre dager.

4 Sammanfattande bedömning

I det stora hela tycker jag att rapporten håller en god standard. I den mån jag är kritisk, gäller detta huvudsakligen följande punkter:

1. Jag saknar en diskussion där de valda metoderna relateras till vad som egentligen är undersökningens syfte: Vad skall avses med effekter? Hur skulle man kunna gå tillväga för att undersöka om sådana föreligger? Vilka kompromisser innebär tillgången till data? Vad skulle man kunna göra för att komma till rätta med sådana problem? En sådan diskussion borde också mer explicit referera till den litteratur kring utvärderingar av arbetsmarknadspolitiska åtgärder som faktiskt finns.
2. En alltför stor del av rapporten ägnas åt ren deskription och en alltför liten åt utvärdering.

Dessa kritiska synpunkter får dock inte undanskymma det faktum att rapporten definitivt bidrar till att öka våra kunskaper om rekryteringsstödet effekter.

Referenser

Anxo, D. (1994), *Erfarenheter av lönesubventioner i Europa*, EFA-rapport nr 31.

3 Bosnier hjälper bosnier — nätverk för framtiden (Invandrarverket)

Bedömning 1

1 Bakgrund till utvärderingen

Projektet Bosnier Hjälper Bosnier

I Sverige verkar Bosnien-Hercegovina riksförbund sedan år 1992. Förbundet har ca 12 000 medlemmar i ett 50-tal lokala föreningar. På initiativ av föreningen startade hösten 1993 projektet Bosnier Hjälper Bosnier med syfte att låta bosnierna själva få ta ansvar för sin tillvaro i Sverige.

Statens invandrarverk (SIV) har stött projektet från början dels med operativa medel dels genom att placera en av SIV anställd projektledare vid Bosnien-Hercegovinariksförbunds kansli. Vid tidpunkten för utvärderingen hade stödet funnits i 20 månader.

Projektet består av fem delprojekt, Sjukvårdsprojektet, Skolprojektet, Kulturprojektet, Informationsprojektet och Återvandringsprojektet. För varje delprojekt finns särskilda målsättningar angivna. Utvärderingen är begränsad till de fyra första delprojekten.

Totalt har SIV fram till tidpunkten för utvärderingen (aug. -95) bidragit med 1 115 000 kr. Uppdelning på delprojekt har inte angetts. Det framgår heller inte av utvärderingen om projektet finns kvar eller om stödet kvarstår.

SIV: syften med projektet

För projektet anges (rapporten s. 5) att "det rimmar väl med den svenska invandrapolitikens ambitioner att engagera asylsökande, flyktingar och alla goda krafter i samhället och att lägga verksamheten på entre-(eller intra-)prenad. Det ligger också i linje med grundprinciperna i svensk invandrapolitik, som vanligen sammanfattas i orden Jämlikhet, Valfrihet och Samverkan".

Projektets mål har varit att "tillvarata den kompetens som finns bland de bosniska flyktingarna i Sverige, dels för att underlätta deras integration, dels för att möjliggöra senare återvändande till och uppbyggnad av hemlandet."

2 Utvärderingens syfte

Utvärderingen, som utförts av en tjänsteman inom SIV, syftar enligt rapportens egna formuleringar till "att SIV:s GD ett underlag för beslut genom att

- ange vad SIV:s stöd betytt för verksamheten
- beskriva hur majoritetssamhället mött gruppens mobilisering och ange framgångsfaktorer och hinder för gruppinitierad verksamhet samt
- beskriva vilka förutsättningar som bör föreligga för att stöd till "etnisk mobilisering" skall vara effektiv och mobiliseringen skall vara framgångsrik"

3 Utvärderingens metod

Utvärderingen är som nämnts utförd internt inom SIV. Det framgår inte om utvärderaren själv varit involverad i beslutet kring själva projektet eller hur han organisationsmässigt befinner sig till beslutsfatta- ren varför styrningsambitionen bakom utvärderingen inte klart framgår för en utomstående betraktare. Något uppdrag till utvärderaren finns inte redovisat. Arbetet förefaller ha bedrivits inom ett par månader och med rel. begränsade resurser. Utifrån de tidigare angivna mål-

sättningar kan man beskriva utvärderingen som en effektanalys gjord utifrån givna, icke ifrågasatta antaganden om projektidéns samklang med övergripande politiska målsättningar.

Utvärderingen har utförts utifrån 40 intervjuer med representanter för projektet, myndigheter, m.fl. De intervjuade finns förtecknade. Någon frågelista presenteras ej och svaren har inte särredovisats utan bakas in i de allmänna resonemangen kring varje delområde. Det framgår heller inte hur de intervjuade förhåller sig till de olika delprojekten.

4 Bedömning av utvärderingen

A Dokumentation

Projektet är, med hänsyn till sin rel. ringa omfattning, väldokumenterat. Själva utvärderingen är däremot inte dokumenterad vad gäller bakgrundsfakta annat än vad gäller intervjupersoner. Arbetet är som framgått ovan mer av karaktären enkel effektanalys utan närmare reflektion kring metod, urval, tillförlitlighet i olika delar m.m.

Bedömningen i denna del försvåras dock av att något direktiv eller utgångsdokument för utvärderingen inte finns tillgängligt samt att det för en utomstående är omöjligt att avgöra hur utvärderare och ursprunglig beslutsfattare förhåller sig till varandra.

B Data

Av skäl som anges under A går det inte att bedöma om de data som ligger till grund för utvärderingen är tillräckliga eller tillförlitliga, detta skulle i så fall krävt en tillgång till betydligt mer basmateriel, bl.a. intervjufrågor och svar.

C Metod

Med tanke på det ursprungliga projektets storlek förefaller själva utvärderingsinsatsen väl avvägd.

D Slutsatserna

Slutsatserna kring delprojekten är underbyggda i analysen men det saknas en direkt koppling till de tre grundfrågorna som anges som utvärderingens syften.

De generella slutsatserna är dock sämre underbyggda och saknar även de en direkt koppling till grundfrågorna. Bl.a. saknas en analys av vad som skiljer den bosniska gruppen, som i utvärderingen beskrivs som Europaorienterad och relativt välutbildad och vars vistelse i Sverige dessutom förefaller att vara av en temporär natur, från andra invandrar- och flyktinggrupper. Denna analys borde vara nödvändig för att beskriva hur "majoritetssamhället mött gruppens mobilisering" eller för att "beskriva under vilka förutsättningar stöd till "etisk" mobilisering skall vara effektiv."

E Kostnadseffektivitet

Då data saknas är det svårt att bedöma utvärderingens kostnadseffektivitet. Som en ren uppföljning av en begränsad insats förefaller ambitionsnivån lagom avvägd.

F Objektivitet

Det vore kanske för mycket att kräva att en internutredning inom SIV skall ifrågasätta grundläggande politiska målsättningar men man bör ändå inte bara utgå från att ett projekt uppfyller dessa så som skett i utvärderingen (s. 5).

Utvärderingen är, i egenskap av en internutredning också bunden till de normer och språk som gäller inom SIV. För en utomstående läsare förefaller det lite egendomligt att beskriva problemen kring en akut flyktingvåg från en krigszon i termer som "hur majoritetssamhället mött gruppens mobilisering" (s. 4) eller som "i den bosniska kulturen finns inte den terapimedvetenhet som vi har t.ex. i Sverige" (s. 19). Det senare yttrandet förefaller högst irrelevant med tanke på landets historia under det senaste seklet. Exempelen är visserligen här lös-

tryckta ur sitt sammanhang men visar ändå genom vilka glasögon man lätt kan betrakta invandrare.

5 Sammanfattande bedömning

Den granskande "utvärderingen" är ett gott exempel på en snabbt utförd relativt enkel uppföljningsinsats kring ett projekt. Utvärderingsstämpeln och de frågeställningar som bildat utgångspunkt, ställer kanske förväntningarna för högt. Det gjorda arbetet förefaller väl fylla kraven på normalt avvägd insats men svarar inte på de grundläggande frågor som ställts. För detta skulle krävs en djupare analys och delvis ett bredare material kring flera invandrargrupper.

Som en delmängd i en sådan större analys kan dock arbetet lämna ett bra underlag.

Arbetet kan därför bedömas som en gediget utförd uppföljning inom SIV av det slag som normalt bör utföras av en myndighet. På basis av det tillgängliga materialet är det svårt för en utomstående bedömare att uttala sig om undersökningens kvalitet, pålitlighet och användbarhet vilket också begränsar utvärderingens användbarhet utanför SIV. Denna situation är förmodligen annorlunda för en beslutsfattare inom SIV.

Bedömning 2

Utvärderingen avser ett s k självhjälpsprojekt, där bosniska flyktingar i Sverige i egen regi och i ett antal delprojekt sökt tillvarata den egna gruppens kompetens, dels för att underlätta deras integration och dels för att möjliggöra eventuell återutvandring. Syftena med utvärderingen är att (i) ange vad Statens Invandrarverks (SIV) stöd betytt för verksamheten samt (ii) beskriva erfarenheterna av implementeringsproblem i de olika projekten och (iii) att ange de förutsättningar som leder till en effektiv mobilisering av självhjälpsgrupper.

Den metod som använts är huvudsakligen en ren beskrivning av åtgärderna i de olika projekten samt en mycket allmänt hållen diskussion om åtgärderna fungerat bra eller ej. De olika verksamheterna be-

stod av ett "skolprojekt", "kulturprojekt", "informationsprojekt" och "stödverksamhetsprojekt". Det senare projektet, där man med hjälp av bosnisk sjukvårdspersonal sökte bearbeta den bosniska flyktinggruppens traumatiska upplevelser från inbördeskriget i hemlandet, är ett helt nytt inslag i svensk invandrapolitik. De olika projekten behandlas i separata delar, och mest utrymme ägnas av förklarliga skäl åt stödverksamhetsprojektet.

A Dokumentation

Studien innehåller en mängd sakuppgifter om vad som utträttats i de olika projekten. Uppenbarligen bygger många av uppgifterna på muntliga uppgifter. I en särskild bilaga lämnas en förteckning över intervjuade och uppgiftslämnare, men relativt få direkta referenser till dessa personer görs i texten. I viss utsträckning omnämns också skriftlig dokumentation inom de olika projekten, men hänvisningarna är inte explicita, dvs med titel och författarnamn. Inga hänvisningar görs till eventuella erfarenheter av tidigare projekt med liknande innehåll.

B Data

De kvantitativa data som ligger till grund för utvärderingen är inte särskilt omfattande. Det rör sig t ex om antalet deltagare i de olika projekten, som dock i övrigt är tämligen detaljerat beskrivna.

C Metod

Utvärderingsmetoden går främst ut på att beskriva vad som utträttats i de olika projekten. Trots det angivna syftet görs inte några försök att utröna vilken betydelse SIV:s stöd – vilket huvudsakligen tagit formen av en anställd projektledare – egentligen har haft. På flera ställen i rapporten medges också att det varit svårt att isolera denna effekt från vad som skulle ha inträffat utan något stöd. Eftersom man inte haft någon "kontrollgrupp" att jämföra med är det inte lätt att dra några slutsatser om stödets effekter.

D Slutsatserna

I huvudsak ges positiva omdömen om utfallet i projekten. Resultaten kan mycket väl ha varit positiva, men i många fall frågar man sig hur väl underbyggda dessa slutsatser är. Omdömena är ofta svepande: "Skolprojektet har lyckats infria de flesta av sina uppställda mål" (sid 11). "(D)en centrala kommittén har arbetat mycket seriöst och har haft stor framgång med mycket av vad den föresatt sig" (sid 14). "(M)an ... ändå lyckats väl med beaktande av den tröghet och den skepsis, med vilken storsamhället mötte dessa helt nya ... propåer" (sid 25). Vissa kritiska synpunkter på utfallen framförs emellertid också. Även där frågar man sig hur väl underbyggda påståendena är. Angående stödverksamhetsprojektet heter det t ex att "(psykologen) bedömer att gruppen – inte minst pga händelseutvecklingen i hemlandet – i stort nog mår sämre nu än då verksamheten kom i gång" (sid 25).

E Kostnadseffektivitet

Det hade säkerligen varit möjligt att komma med mer underbyggda påståenden inom ramen för de givna resurserna. En kontrollgruppsansats hade varit önskvärd (se nedan), vilket möjligen hade kunnat inrymmas med en annorlunda prioritering av resurserna.

F Objektivitet

Som beskrevs i punkten D är slutsatserna om åtgärdernas effektivitet tämligen svepande. Huruvida detta är uttryck för bristande objektivitet är svårt att bedöma. Även kritiska synpunkter på utfallen framförs.

Studien kan inte kallas en utvärdering i gängse vetenskaplig mening. En sådan utvärdering är nämligen en jämförelse mellan utfallen för en experimentgrupp, som utsätts för åtgärden, och en kontrollgrupp, som inte deltar. De beskrivningar av utfallen i experimentgruppen som redovisas är därmed inte ointressanta, men man måste naturligtvis vara mycket försiktig med långtgående slutsatser.

Ofta är det förenat med stora praktiska eller administrativa svårigheter att använda sig av en kontrollgrupp. I stödverksamhetsprojektet borde man dock ha kunnat använda sig av en sådan grupp, eftersom denna verksamhet inte kom till stånd på mer än två orter – Borås och Lund – och det finns andra orter med många bosnier som man ganska lätt kunnat jämföra med. Intervjuundersökningar skulle kunna ha genomförts med deltagare respektive icke-deltagare. En intressant fråga är t ex om icke-deltagare anlitade den reguljära sjukvården i större utsträckning än deltagare. I så fall skulle man ha kunnat peka på möjligheten till kostnadsbesparingar av självhjälpsprojektet. Enligt rapporten skall en intervjuundersökning genomföras, dock endast med bosnier på orter där verksamheten bedrivits. Av rapporten framgår vidare att det psykiska välbefinnandet hos de bosniska flyktingarna också tycks påverkas av faktorer utanför projektet som påverkar både deltagare och icke-deltagare, t ex utvecklingen i hemlandet, vilket ytterligare understryker vikten av att ha tillgång till en kontrollgrupp.

De kostnader för projektet som redovisas hänför sig endast till SIV (bil. 2). Denna kostnadsbild är dock ofullständig eftersom även kommunerna har ådragit sig kostnader genom att tillhandahålla lokaler m.m. Eftersom studien behandlar en verksamhet med stor potential för kostnadsbesparingar, s k självhjälpsprojekt, är det beklagligt att man inte på ett mer systematiskt sätt försökt diskutera kostnader och intäkter.

4 Tusentals kontakter – nätverk för framtiden (Invandrarverket)

Bedömning 1

3

Syftet är att dokumentera och utvärdera det s.k. Kontaktnätet som syftar till att med hjälp av frivilliga guida flyktingar in i det svenska samhället. Rapporten ska också bidra till diskussionen om metodutveckling och att försöka stimulera forskare att analysera Kontaktnätet. Syfte förefaller dock *inte* att vara att utvärdera effektivitet och kostnader för projektet.

4

Metoden är en mycket noggrann beskrivning av genomförandet av Kontaktnätet. Man redogör för olika enkätundersökningar som gjorts samt lägger in ett stort antal intervjuer med guider och med flyktingar. Man beskriver också olika myndigheters roll och uppträdande och hur olika frivilligorganisationer har engagerats.

5 A

Detta är inte en utvärdering i egentlig mening för att bedöma effektiviteten av åtgärder. Detta är en dokumentation som ska ligga till

grund för förändring och vidareutveckling av en viss bestämd verksamhet. Kostnaderna för projektet finns inte redovisade och man redovisar inte heller några mått på resultat. Detta skulle kräva data över hur många flyktingar som säger sig vilja ha kontakt med guider, hur många som verkligen fått kontakt med guider, vilka kostnaderna varit för att skapa kontaktnätet inkluderande uppskattade arbetskostnader för personer i kommuner, verk och organisationer, m.m. Hur många personer har använts på SIV för att genomföra projektet? Vad har informationskampanjerna och enkätundersökningarna kostat.

Det förefaller ha varit underförstått att i och med att projektet bygger på frivilliginsatser så behöver man inte genomföra någon kostnads kalkyl och att projektet därmed är försvarbart. Något egentligt ifrågasättande av verksamheten förekommer inte i rapporten.

B

Data är alltför få för en traditionell utvärdering.

C

Metoden är i huvudsak intervjuer med enskilda guider och flyktingar. Dessa är dock inte systematiskt genomförda. Man vet inte heller om de utfrågade är representativa. De är också alltför få till antalet.

D

Syftet är knappast att genomföra en samhällsekonomisk utvärdering. Syftet är att dokumentera och beskriva genomförandet av projektet och de inblandade organisationernas uppträdande för att dra slutsatser om hur projektet kan vidareutvecklas. Givet detta är rapporten sannolikt av stort värde för de inblandade organisationerna. Rapporten är ytterligt välskriven och välgjord.

E

Det hade sannolikt varit fullt möjligt att genomföra en utvärdering av den samhällsekonomiska effektiviteten av projektet.

F

Jag betraktar rapporten som en intern utvärdering för att organisatoriskt vidareutveckla projektet, inte för att utvärdera projektets kostnadseffektivitet.

6 Sammanfattande kommentarer

Det är rimligt att myndigheten försöker att på något sätt mäta resultaten av projektet. Hur många flyktingar har kommit i kontakt med guider? Hur många har önskat men inte fått kontakt? En systematisk analys av hur flyktingarna uppskattar värdet av verksamheten saknas. Inga kostnader redovisas. Dessa borde inkludera ett uppskattat värde av de totala arbetsinsatserna på alla organisationer som har engagerats, kampanjkostnader, enkätkostnader m.m.

Det är en imponerande mängd information som redovisas och som har sammanställts på ett överskådligt och i många avseenden förtjänstfullt sätt. Rapporten är mycket välskriven och lättillgänglig. Det förefaller dock som om det faktum att projektet baseras på frivillighet, solidaritet och medmänsklighet och då syftet bl.a. är att motverka rasism och främlingsfientlighet och främja kontakter mellan svenskar och invandrare tycker man sig kanske inte behöva bedöma om målen uppnås och vilka kostnader som egentligen uppstår.

Bedömning 2

3. **Utvärderingens syfte** är att utvärdera den kampanj som Statens Invandrарverk initierade sommaren 1993 för att tillsammans med kommuner, frivilligorganisationer och enskilda utveckla nya arbetssätt

kring flyktingarnas integration i det svenska samhället. Kampanjens huvudsakliga aktivitet skulle vara att skapa relevanta sociala lokala nätverk som skulle vara fruktbarande för integrationsprocessen. Som ett led i detta skulle man på central nivå, inom Statens Invandrarverk, upprätta en stödjande funktion med syfte att vara den informerande och samordnande funktionen. Slutligen syftade kampanjen till att väcka intresse hos forskarna, framför allt för att utveckla metoderna inom integrationsarbetet.

Utvärderingen täcker inte alla aspekter av Kontaktnätets syften. Man utvärderar om Kontaktnätet kom att skapa lokala nätverk, s.k. flyktingguider. Vidare utvärderar man om informationskampanjen nådde ut till landets medborgare, samt om kampanjen medförde att forskning kring integrationsarbetet initierades. Däremot täcker utvärderingen inte frågan om Kontaktnätet medförde att flyktingarnas integration i det svenska samhället förbättrades. I utvärderingen utgår man ifrån att skapandet av Kontaktnätet befrämjar integrationsarbetet.

4 Utvärderingens metoder

Utvärderingen innehåller följande moment:

- a) Intervjuer med 18 st guider kompletterat med 42 st enkäter till andra guider. Syftet med denna aktivitet var att få uppslag om vilka faktorer som var avgörande för att de blev guider. Det framgår inte vem som gjorde intervjuerna.
- b) Två utvärderingar av effekterna av de centrala annonskampanjerna. Telefonintervjuer av 500 strategiskt utvalda personer i maj och dec. 1994. Undersökningen genomfördes av TEMO.
- c) Kontaktnätets sekretariat skickade ut en enkät till kommunerna under perioden juni–okt 1994. 190 av 265 kommuner besvarade enkäten, som innehöll "några frågor".
- d) Enkäter skickades till de centrala nivåerna av de organisationer som deltog i kampanjen. En skickades i mars 94 och en under sommaren -95. Syftet med enkäterna var att följa upp hur återrapporteringen från organisationernas lokalavdelningar fungerade. Det framgår inte vem som skickade ut enkäterna.

- e) Kontaktnätets sekretariat har upprättat en central dokumentation över alla tidningsartiklar som skrivits om Kontaktnätet.
- f) Slutligen nämner utvärderaren att man har tagit kontakter med kommunernas flyktingsamordnare. Några närmare omständigheter kring detta framgår dock inte.

5 Min bedömning av utvärderingen i olika avseenden

A Dokumentation

Dokumentationen av de utvärderingsaktiviteter som redovisats ovan är inte tillräcklig för att läsaren skall kunna skapa sig en egen självständig bild av de slutsatser som man drar. Enkäternas frågor redovisas inte, inga svarsfrekvenser eller statistiska bearbetningar av dessa redovisas. Detta gäller både TEMO-undersökningarna och de övriga enkäterna.

Ett undantag från denna generella bild utgör redovisningen av de intervjuer (inklusive enkäterna) som gjorts med de 18 guiderna. Dessa intervjuer beskrivs levande och intresseväckande.

Det finns ingen dokumentation angående hur man undersöker vilken forskning som sker eller som initierats av kampanjen. Utvärderaren konstaterar endast att det skett någon forskning, vilket antagligen beror på att han uppfattat kampanjens målsättningen som att initiera forskning kring *Kontaktnätets* verksamhet, och inte kring integrationsfrågor eller invandrarfrågor *generellt*. Den forskning som äger rum kring invandring och integrationsfrågor i Sverige och internationellt diskuteras inte.

B Data

Eftersom utvärderingens huvudsakliga syfte är begränsat till att försöka kartlägga hur många guider eller lokala kontaktnät som skapats så blir valet av data ganska enkelt och självklart. Däremot finns det ingen egentlig insamling av data som belyser detta, utan endast en uppskattning gjord i respektive kommun. I utvärderingen finns inte heller någon beskrivning av hur dessa uppskattningar gjorts ute i kommuner-

na. Bristen på en egentlig datainsamling är en av utvärderingens huvudsakliga svagheter.

Om man uppfattar utvärderingens syfte något bredare till att omfatta vad som har hänt i kommunerna kring dessa frågor efter introducerandet av Kontaktnätet, t ex hur har tidigare arbetssätt påverkats, så är valet av data otillräckligt.

Utvärderingen syftar också på att belysa hur de centrala informationskampanjerna når ut till medborgare i landet. Här är valet av data lämpligt. Dock är resultaten av TEMO-undersökningarna dåligt dokumenterade, som nämnts ovan.

C Metod

Under denna punkt har jag de flesta och allvarligaste invändningarna. Alla försök att mäta effekter av en aktivitet förutsätter att man jämför med en situation där insatsen/kampanjen inte ägt rum. Detta kan ske på flera sätt varav det mest uppenbara är att utgå ifrån en kartläggning av tillståndet i kommunerna innan kampanjen genomfördes, och jämföra denna med situationen efter. Om det är omöjligt att göra en total kartläggning i alla kommuner så kan man åtminstone välja ut ett antal representativa kommuner där man gör djupare kartläggningar. Ett alternativt angreppssätt skulle kunna vara att göra en jämförelse mellan de kommuner som engagerat sig i projektet och de kommuner som inte gjort det.

Ingen av dessa jämförelser har gjorts. Därigenom blir det enligt min mening mycket svårt att säga något konkret om kampanjens effekter, till och med utifrån den begränsade utgångspunkt man har i utvärderingen. Låt mig konkretisera vad jag menar:

Engagemang kring flyktingarna i den form som man vill skapa genom Kontaktnätet är inget nytt i Sverige. Utredningen pekar självt på flera sådana aktiviteter som bedrivs i mer eller mindre organiserade former. De visar sig också, enligt utvärderingen, att de flesta av de guider som kom att delta i Kontaktnätet varit engagerade tidigare kring flyktingarbetet antingen genom professionen eller på frivillig basis. Vad händer med dessa aktiviteter då Kontaktnätet drar igång? Försvinner de? Kompletteras de? Utvecklas de? Är det så att Kontaktnätet endast förändrar mera spontana arbetsformer till centralstyrda

aktiviteter? Blev frivilligarbetet sämre eller bättre? Blev det flera personer som engagerades i lokala nätverk eller ej?

Utan riktiga svar på sådana frågor så måste det vara omöjligt att ha en uppfattning ens om den begränsade fråga som utvärderingen försöker beskriva, nämligen vilka effekter på de lokala frivilliga arbetsinsatser kring flyktingarna introducerandet av Kontaktnätet fick.

Det är inte ovanligt i samband med utvärderingar att utvärderaren kommer in så sent i processen att det är mycket svårt eller omöjligt att kartlägga referenspunkter i undersökningen. Detta problem har inte förelegat i denna utvärdering eftersom utvärderaren enligt boken deltog i Kontaktnätets skapande redan i ett tidigt stadium.

En annan metodfråga som jag har svårt att se hur man hanterat i utvärderingen gäller behandlingen av de data som TEMO-undersökningarna resulterat i. I utvärderingen finns flera diskussioner och slutsatser angående dessa data, till exempel vilka grupper som är mest positiva till Kontaktnätet. Det finns ingen redogörelse för vilka statistiska metoder som använts för att nå fram till dessa slutsatser.

D Slutsatserna

Av ovanstående diskussion framgår att min mening är att man inte kan dra slutsatsen att Kontaktnätets första tid varit en framgång. Skälet till detta är att:

man **vet inte hur många guider** som kommit igång (det finns endast en mycket grov uppskattning)

man **vet inte vilka effekter** som Kontaktnätet har haft på annan frivillig eller organiserad verksamhet

Till detta kan man lägga att man inte vet om frivilliga insatser av det slag som Kontaktnätet representerar förbättrar flyktingarnas integration i det svenska samhället, till exempel om integrationen på arbetsmarknaden underlättas. Det är ju ingen orimlig hypotes att detta arbete är ytterst grannlaga och att det kommer att kräva stora och samordnade insatser vad det gäller attitydpåverkan och utbildning, och att det kommer att ställa mycket stora krav på de som har dessa arbetsuppgifter. Det kan ju till och med tänkas att frivilliga insatser, ut-

förda på felaktigt sätt, försvårar integrationsarbetet och försämrar flyktingens möjligheter till framgång.

Utvärderingens slutsatser gällande informationsinsatsernas effekter är välunderbyggda och rimliga.

Utvärderingens slutsatser gällande initierad eller pågående forskning är inte välunderbyggda eller rimliga.

E Kostnadseffektivitet

Denna fråga är svår att diskutera utan en mera noggrann redovisning av de insatser som lagts ned. Såvitt jag kan bedöma skulle utvärderingen ha vunnit mycket på att försöka sig på en jämförelse över tiden i ett par utvalda kommuner, med en rejäl kartläggning initialt.

F Objektivitet

Mitt intryck är att utvärderaren följt genomförandet på ett mycket nära håll, kanske till och med deltagit aktivt i processen. Såvitt jag förstår har han inte gjort egna enkäter utan utnyttjat de som Kontaktnätets sekretariat utfört. Om utvärderaren varit anställd av detta sekretariat så blir hans diskussion vad det gäller sekretariatets framtida roll inte objektiv.

Vidare gäller att den grundläggande värdering som styr detta projekt aldrig ifrågasätts, dvs att frivilliginsatser av det slag som Kontaktnätet representerar är bra, nyttiga och effektiva för integrationen av flyktingar. Invändningar som uttalas gäller detaljer i kampanjutformningen och inte mera fundamentala invändningar gällande till exempel avvägningar mellan frivillighet mot kontinuitet eller amatörism mot professionalism. Det är svårt att frigöra sig från misstanken att utvärderingen representerar en ganska okritisk beskrivning av den förhärskande föreställningen om vilken metod man skall använda sig av för att minska problemen kring integrationen av flyktinginvandrarna i det svenska samhället. *Denna utvärdering* har inte visat att ökade inslag av frivilliga insatser och föreningsmedverkan är rätt metod.

6 Sammanfattande bedömning och ytterligare kommentarer

Utvärderingens *styrka* är att det finns en intresseväckande beskrivning av genomförandet av ett spännande projekt, en beskrivning som görs levande genom intervjuer och kommentarer av de som deltagit i projektet.

Utvärderingens *svaghet* är att man kring de centrala frågeställningarna saknar en riktig datainsamling och en genomtänkt metod för att göra jämförelser. Därtill kan läggas att utvärderaren har en okritiskt positiv hållning till hela projektet, vilket resulterar i att slutsatserna är lite distanslösa.

5 Minskade hälso- och miljörisker vid användning av bekämpningsmedel – resultat av handlingsprogrammet (Jordbruksverket)

Bedömning 1

1 Utvärderingens syfte

Syftet är att utvärdera ett handlingsprogram gemensamt för Jordbruksverket, Kemikalieinspektionen och Statens naturvårdsverk. Målet för handlingsprogrammet är att "minska hälso- och miljörisker vid användning av bekämpningsmedel i jordbruket" (sid. 4). I avsikt att uppnå detta övergripande mål omfattar handlingsprogrammet åtgärder som kan delas in i tre kategorier:

- substitution från bekämpningsmedel med höga risker till sådana med lägre risker
- information och utbildning om hantering av bekämpningsmedel och kontroll av bekämpningsmedelsrester i livsmedel
- minskad total användning av bekämpningsmedel

2 Utvärderingens metod

Handlingsprogrammets yttersta mål om reduktion av risker "är svårt att värdera och kvantifiera och utvärdering måste därför främst ske genom beskrivningarna av åtgärderna" (sid. 8). Man har alltså valt att utvärdera handlingsprogrammet genom en genomgång av samtliga åtgärder under de tre olika kategorierna. För en typ av åtgärd, minskad

användning av bekämpningsmedel, har man kunnat följa upp effekterna av olika styrmedel – förbud och avgifter – genom att undersöka försäljningsstatistik för bekämpningsmedel.

3 Bedömning av utvärderingen

I det följande kommer jag att granska utvärderingen utifrån en mall för en "idealisk" utvärdering. Det ska påpekas att det sällan eller aldrig är möjligt att genomföra en sådan idealisk utvärdering. Mallen ska därför i första hand betraktas som ett tänkesätt för en utvärdering. Nedan följer en kortfattad beskrivning av mallen och därefter en bedömning av utvärderingen.

3.1 Principer för utvärdering av projekt

En vägledande princip för en utvärdering av ett projekt är att undersöka om nyttan av projektet överstiger uppostringarna (kostnaderna) för att genomföra det. Ett projekt kan då anses vara befogat när dess nytta överstiger kostnaderna. För att kunna genomföra en sådan analys måste man

1. Identifiera alla nytto- och kostnadsposter
2. Kvantifiera dessa poster
3. Värdera (i monetära termer) och aggregera posterna

Överför vi detta tänkesätt på utvärderingens mål om en minskning av hälso- och miljörisker vid användning av bekämpningsmedel kan vi tänka oss följande på de tre olika punkterna.

1 Identifiering av nytto- och kostnadsposter

Nyttan av ett handlingsprogram består av; i) lägre hälsorisker vid jordbrukarnas användning av bekämpningsmedel och vid intag av livsmedel (inklusive dricksvatten) och ii) minskade miljörisker. De minskade miljöriskerna innebär framför allt lägre tillförsel av bekämp-

ningsmedel till de ekosystem som omger den besprutade jordbruksmarken. Detta innebär i sin tur en högre funktionsförmåga hos dessa ekosystem vilket bl.a. bidrar till högre biodiversitet.

Kostnaderna för handlingsprogrammet uppstår hos jordbrukarna i form av minskad skörd orsakad av en reducerad användning av bekämpningsmedel och eventuell ökad tidsåtgång för hantering av bekämpningsmedel. Därutöver tillkommer administrativa kostnader för genomförande och uppföljning av regleringar m.m. och kostnader för tillsyn av att regleringarna efterlevs. Det ska noteras att efterlevnad av regleringar eller att jordbrukarna betalar de avgifter de ska kan vara ett problem eftersom det är relativt enkelt att "smuggla" förbjudna eller avgiftsbelagda bekämpningsmedel från andra länder.

2 Kvantifiering av nytto- och kostnadsposter

För att kunna kvantifiera hälsorisker behöver man känna till samband mellan olika halter av bekämpningsmedel och risker för olika typer av sjukdomar. Efter vad jag förstår finns det sådana riskanalyser fastställda. Det framgår emellertid inte av rapporten hur dessa risker är kvantifierade. För bedömning av miljörisker behöver man känna till transporter av bekämpningsmedel från spridningspunkten till omgivande ekosystem och motsvarande biologiska och ekologiska effekter. Som påpekas i rapporten är det mycket svårt att få fram sådana samband, men det pågår för närvarande forskning kring transporter och ekologiska effekter av bekämpningsmedel.

För att kunna beräkna kostnader för minskade risker behöver vi också känna till samband mellan åtgärder – t.ex. minskad användning av bekämpningsmedel eller förändrad sammansättning av bekämpningsmedel – och resulterande förändringar av hälso- och miljörisker. Det är relativt enkelt att beräkna de olika ovan angivna kostnadsposterna som uppstår hos jordbrukare och myndigheter. För samtliga kostnadsberäkningar finns det standardiserade metoder. Däremot är det idag mycket svårt att fastställa samband mellan åtgärder och risk-effekter.

3 Värdering av nytto- och kostnadsposter

Det skall påpekas att även om vi lyckas kvantifiera de ovannämnda samband mellan bekämpningsmedel och hälso- och miljörisker, uppstår det svårigheter när dessa förändringar ska mätas i monetära termer. När det gäller monetär värdering av förändrade hälsorisker finns det emellertid i dag standardiserade metoder, vilka har utvecklats under ca 30 år. Motsvarande finns inte när det gäller miljörisker. Det pågår en intensiv debatt mellan forskare om vad som kännetecknar en ändamålsenlig monetär värdering av miljöförändringar. Som sades ovan, är det betydligt enklare att skatta de olika kostnadsposterna.

En slutsats av denna mycket korta genomgång principer för en utvärdering av projekt om minskning av miljö- och hälsorisker är alltså att man tvingas vara mindre ambitiös när det gäller utvärdering av nyttoeffekter. En alternativ princip är då att undersöka om man uppnått ett eller flera mål om miljö- och hälsorisker till lägsta kostnad, dvs en utvärdering baserad på kostnadseffektivitet. Principen för en sådan utvärdering är den ovan presenterade trestegs-analysen där man enbart undersöker kostnadsposterna.

3.2 Bedömning

Min bedömning av utvärderingen kommer att ske utifrån, i huvudsak, om målet eller målen för handlingsprogrammet genomförts på "bästa" sätt där jag definierar bästa sätt som att målen uppnåtts till lägsta kostnad. Enligt mallen ovan innehåller en sådan utvärdering en genomgång av alla tänkbara åtgärder och styrmedel och beräkning av deras kostnader och effekter på miljö- och hälsorisker.

Utvärderingen har lyckats väl när det gäller en beskrivning av de åtgärder och styrmedel man valt att genomföra. Dessa åtgärder innebär i princip information och utbildning av jordbrukare, förbud mot användning av bekämpningsmedel med höga risker och ekonomiska styrmedel för att minska totala användningen av bekämpningsmedel. Dessutom bidrar man till forskning om bl.a. betydelsen av kantzoner för spridning av bekämpningsmedel, mätning av låga halter av bekämpningsmedel i dricksvatten och livsmedel och transporter av bekämpningsmedel i mark och vatten.

För att kunna bedöma om dessa åtgärder är kostnadseffektiva eller ej måste man emellertid också känna till vilka alternativa åtgärder som är möjliga. I rapporten diskuterar man t.ex. införande av avgifter som relateras till risker av olika bekämpningsmedel. Utöver detta kan man också tänka sig en höjning av den redan introducerade avgiften för att minska totala användningen av bekämpningsmedel. Ett annat alternativ utgörs av en regional anpassning av avgifter så att dessa är högre i regioner där markförhållandena är sådana att risker för spridning av bekämpningsmedel är större eller där vi har särskilt känsliga ekosystem. Det medges att sådana differentierade system ställer högre krav på tillsyn av efterlevnad. En möjlighet att minska den faktiska användningen av bekämpningsmedel är då att, istället för ytterligare regleringar eller andra styrmedel, öka insatserna för tillsyn. Uppenbarligen finns det flera möjliga åtgärder än vad som inkluderats i rapporten.

Min slutsats när det gäller punkt 1 ovan, dvs identifiering av åtgärder och kostnadsposter, är att rapporten innehåller många åtgärder men man har inte gjort en systematisk genomgång av samtliga tänkbara åtgärder. Det är möjligt att det funnits flera typer av åtgärder och styrmedel på förslag men att dessa av olika skäl inte inkluderats. Som läsare är det emellertid omöjligt att bedöma valet av åtgärder eftersom man inte får information om möjliga åtgärder utan enbart om faktiska.

I rapporten är man väl medveten om behovet av att fastställa samband mellan åtgärder och effekter på miljö- och hälsorisker, dvs punkt 2 ovan. Som påpekas är nuvarande kunskaper om transporter och effekter av bekämpningsmedel otillfredsställande. Forskning kring detta stimuleras därför. Den enda uppföljning man därför anser sig kan göra är att titta på förändringar i försäljningsstatistik. Jag håller i stort sett med om detta. Min invändning gäller emellertid vad man vet om eventuell "smuggling" av bekämpningsmedel. Detta kan vara ett problem eftersom det är ganska enkelt att föra in bekämpningsmedel från andra länder.

Ett annat krav för att kunna bedöma kostnadseffektivitet är att man anger beräknade kostnader för olika åtgärder, punkt 3. Kostnaderna uppstår, som sagt, hos jordbrukarna och tillsynsmyndigheter och för administration. Rapporten innehåller inga som helst kostnadsberäkningar. Detta är förvånansvärt eftersom det finns relativt enkla och

standardiserade metoder för beräkning av de flesta kostnader. Dessutom har kostnader för skördebortfall mm beräknats bl.a. på Sveriges lantbruksuniversitet. Det hade inte inneburit någon större ansträngning att ta del av det forskningsmaterial som finns tillgängligt och inkluderat och diskuterat detta i rapporten. Uppgifter om kostnader för administration av olika åtgärder och styrmedel torde finnas tillgängliga hos de myndigheter som utformat rapporten.

4 Sammanfattning

Givet att utvärderingen baseras på hur man uppfyller målen för handlingsprogrammet, är det tveksamt om man skall kalla rapporten för utvärdering. Som begreppet antyder är det då nödvändigt med någon form av värdering. Den mall för värdering som jag valt är om man har lyckats uppnå de mål man stipulerar till en så låg kostnad som möjligt. Emellertid är det omöjligt att bedöma utvärderingen utifrån detta kriterium eftersom dokumentationen är så ofullständig. Man kan inte göra en sådan ändamålsenlig utvärdering baserad enbart på en beskrivning av de åtgärder man valt att genomföra. Ett minikrav är att också kostnadsberäkningar finns med. Enligt min uppfattning uppfyller alltså inte rapporten minikraven för vad man skall kräva av en utvärdering.

Självfallet kan man diskutera om mitt val av bedömningsgrund är ändamålsenligt. Tyvärr har jag stora svårigheter att se hur man kan applicera andra bedömningsgrunder. En möjlighet vore att bortse från målet med handlingsprogrammet, dvs att minska hälso- och miljörisker förknippade med bekämpningsmedel, och istället undersöka hur man genomfört de faktiska åtgärderna. Även vid en sådan begränsad utvärdering behöver man emellertid kostnadsuppgifter.

Bedömning 2

3. Syftet är att rapportera om hur arbetet med att minska användandet av bekämpningsmedel har utvecklats.

4. Rapporten innehåller en presentation av de åtgärder som vidtagits samt en redovisning av hur användandet av bekämpningsmedel har utvecklats mellan 1981–1993. Även kostnaderna för hur man gått till väga för att minska användningen av bekämpningsmedel är redovisade i rapporten.

5 a

Ja. Syftet är här att belysa om att kvantitativt mål har uppnåtts. Data för detta är ganska enkelt att ta fram och läsaren får en klar bild av hur användningen av bekämpningsmedel, totalt och uppdelat på olika typer av medel, har utvecklats. Även kostnaderna för programmet redovisas klart.

b

Såvitt jag kan bedöma är data relevanta. (Detta kan dock bättre bedömas av en kemist.)

c

Metoden är helt enkelt en redovisning av hur mycket användningen av bekämpningsmedel har minskat.

d

Såvitt jag kan bedöma är slutsatserna väl underbyggda och tolkningarna rimliga.

e

I detta fallet gäller det att belysa om ett kvantitativt mål har uppfyllts eller ej och till vilka kostnader. Det är tämligen lätt att konstatera att det kvantitativa målet har uppfyllts och vilka kostnader som har uppstått. Kostnaderna förefaller vara mycket låga. Jag har svårt att se att man skulle göra en utvärdering på något annat sätt.

f

Ja. Jag har ingen anledning att ifrågasätta objektiviteten hos utvärderaren. P.g.a. att det s.k. halveringsprogrammet är kvantitativt uttryckt finns små möjligheter att tolka resultaten på andra sätt.

6 Sammanfattande kommentarer

Det kvantitativa målet som myndigheterna har fått i uppdrag att uppnå synes vara lätt att belägga i en rapport. Även kostnaderna framstår som lätta att kartlägga och redovisa. Förändringen av användningen av bekämpningsmedel är klart och tydligt redovisade, liksom kostnaderna.

6 Varudistribution – uppföljning av resultat (Konsumentverket)

Bedömning 1

Utvärderingens syfte

Inget explicit syfte anges i denna resultatuppföljning. Det är dock inte omöjligt att ett sådant angetts i den särskilda rapporten till vilken resultatuppföljningen är en bilaga.

Bilagans skrivning tyder på att avsikten också varit att den även ska kunna fungera som en fristående utredning på området. I sådana fall hade återgivning av direktiven (uppdraget) och/eller syfte varit önskvärt.

Utvärderingens metod

Utvärderingens metod är inte heller angiven. Rapporten är dock en deskriptiv studie, en skrivbordsprodukt så tillvida att den bygger på tidigare genomförda studier.

A Dokumentation

De stödformer som redovisas i resultatredovisningen är uppenbarligen komplexa så tillvida att det dels förekommer olika stödformer dels olika aktörer på olika nivåer. Redovisningen omsluter dessutom nästan

två decennier varvid både stöden och aktörerna förändrats. Konsumentverkets roll och relation till övriga aktörer har också ändrats över tiden. Detta sammantaget ställer höga krav på struktur i studien. I föreliggande rapport återkommer dessa olika delar i alla huvudkapitel. Om dessa olika delar hållits åtskilda hade rapporten kanske blivit mer lättillgänglig och förståelsen ökat.

Validitet eller reliabilitet diskuteras ej.

B Data

Studien bygger på tidigare genomförda studier av olika organ inklusive Konsumentverkets egna:

"Analysen omfattar åren 1973/74–1990/91 med tonvikten lagd på åtgärder under de senaste fem budgetåren. Den bygger dels på en rad tidigare utförda centrala, regionala och lokala uppföljningar/utvärderingar samt ett antal effektmätningar Konsumentverket genomfört på senare tid En avstämning har gjorts med länsstyrelsernas handläggare 1991 Vidare har avstämningar av resultat gjorts med ett antal kommuner samt företrädare för glesbygds-handeln. Gjorda analyser och bedömningar av resultat har fortlöpande stämts av med Glesbygdsnämnden." (sid 4).

Tonvikten ska enligt uppgift vara "de senaste fem budgetåren". Även om ambitionen varit denna så har man inte riktigt nått fram. Beskrivningarna och redovisningarna i alla studiens delar är jämt fördelade över alla de år som studien omfattar. Denna ständiga tillbakablick medför att läsaren hela tiden måste vara observant på vilket decennium och år som avses. Av de redovisade källhänvisningarna är 14 från budgetåren 87–91. Totalt redovisas ca 61 källor i resultatredovisningsdelen.

Inga nya data för just denna studie förefaller ha insamlats.

Uppgifter om slutmottagarna saknas (se vidare under metod). I till exempel *Del II Dagligvaruhandelns utveckling* (sid 15) redovisas uppgifter om ålder på innehavarna av de privata butikerna. Motsvarande uppgifter om slutmottagarna redovisas ej. Demografiska data och and-

ra uppgifter om slutmottagarna borde ha införskaffats och sammanställts. Vilka är dom? Vilken ålder? Varför behöver de stöd?

C Metod

Eftersom inget syfte med studien redovisas är det i princip omöjligt att diskutera om använda data, metod och tillvägagångssätt är de lämpligaste. Förutsatt att studien ska vara deskriptiv är den tillfyllest med de reservationer som gjorts under de andra punkterna.

Problematisering saknas. Varför behövs stöd av dessa slag? Vem är mottagare? Just den senare frågan är mycket intressant. Det förhållande att stödet ges till en annan (t ex butiksinnehavare) än slutmottagaren (konsumenten) är av stor vikt.

D Slutsatserna

Även slutsatserna påverkas av det faktum att stödet ges till annan än slutmottagarna.

I sammanfattningen sägs att *"Avgörande för uppnådda resultat har varit det breda spektra av kompletterande åtgärder som efter hand utvecklats. Det har skett i nära samverkan med främst länsstyrelser, kommuner och handeln."* (sid 2). Den senare meningen är ett exempel på att andra aktörer än slutmottagarna är de som framför allt uppmärksammas i Konsumentverkets resultatuppföljning. Beskrivning av de olika aktörerna och deras inbördes relation finns visserligen men beskrivningen är splittrad och ibland svåråtkomlig. Dessa aktörers relation till slutmottagarna kunde också beskrivits på ett mer samlat och lättillgängligt sätt. Slutmottagarna och deras uppfattning om servicen borde ha uppmärksammats mer.

Slutsatsen att *"stödet givit mycket goda resultat till relativt blygsamma kostnader"* kan tänkas ha varit en annan om Konsumentverket tagit reda på slutmottagarnas uppfattning om stödets effekter. Dessutom andas kostnadsaspekten och den ibland detaljerade redovisningen av kostnader att verket kanske tänkt sig en slags produktivitetmätning (förhållandet mellan insatta resurser och fullgjord prestation) även om detta inte är uttalat.

E Kostnadseffektivitet

En skrivbordsstudie i form av sammanställning från redan tidigare genomförda rapporter och utredningar är naturligtvis ett lågbudgetalternativ. Data finns redan och pengar behöver inte satsas på insamlande av primärdata. Tidsmässigt går en sådan sammanställning givetvis snabbare än om alla data måste insamlas i egen regi eller om komplettering av befintliga data krävs.

Att som i det här fallet enbart använda sekundärdata kan knappast vara vanligt förekommande. Med ett sådant utgångsläge är det risk för att studiens uppläggning anpassas till befintliga data och intressanta uppgifter riskerar att utelämnas på grund av att de ej är insamlade (se vidare under objektivitet). Kompletteringar i form av primärdata hade visserligen fördyrat undersökningen men inte i sådan omfattning att det varit ogörligt. Kanske är det kort om tid som påverkat avsaknaden av primärdata.

F Objektivitet

Objektivitet kan användas i flera olika betydelser. Objektivitet som opartiskhet är en sådan. Ett sådant krav är i detta fall omöjligt att uppnå eftersom Konsumentverket getts i uppdrag att utreda ett stöd som verket självt till vissa delar handhar. Konsumentverket är därmed part i målet. Denna problematik med att myndigheterna utvärderar sin egen verksamhet är inbyggd i det nya budgetsystemet.

En närbesläktad betydelse är objektivitet som förutsättningslöshet. Härmed menas också förenklat uttryckt att utredningen bör genomföras utan några förutfattade meningar. Även i detta fall kan samma invändningar som i det tidigare fallet företas, dvs Konsumentverket är part i målet.

Den tidigare refererade slutsatsen att "*stödet givit mycket goda resultat till relativt blygsamma kostnader*" kan tyda på slutsatsen relateras till just kostnader istället för till målet med stödet eller det behov av service som slutmottagarna anser sig ha. Kostnadsaspekten kan vara en förutfattad mening eller förutsättning som borde redovisats öppet.

Objektivitet kan också ha betydelsen fullständighet. Med detta menas förenklat uttryckt att det som studeras beskrivs och förklaras på ett fullständigt sätt. Inga fakta får förtigas eller vinklas så att det som beskrivs inte överensstämmer med verkligheten. Det räcker alltså inte med att beskrivningen är sann den ska också vara fullständig och ingen information får utelämnas. I det här fallet har Konsumentverket brutit i enlighet med diskussionen om att slutmottagarnas åsikter utelämnats. Å andra sidan bör objektivitet också medföra att fakta och värderingar skiljs åt. Om slutmottagarnas åsikter om stödet redovisats kan det diskuteras om åsikter och attityder är fakta eller vad som den tillfrågade tycker just i det ögonblicket.

Sammanfattande bedömning

Studien har både goda och mindre goda företräden. Lovvärt är den beskrivning av hur stödet växt fram och förändrats över tiden, även om den kunde varit mer strukturerad. Konsumentverket har också genomfört en rejäl genomgång av vidtagna undersökningar på området.

Mindre gott är avsaknaden av diskussion och problematisering av det förhållandet att stöden är indirekta, dvs stödet ges till en annan än den som det avser att förbättra för. Mindre gott är också att uppgifter om slutmottagarna saknas och dessas behov av stöden. Inte heller framgår med någon större tydlighet hur de uppfattar nyttan.

Trots detta kan det anses att Konsumentverket kommit långt och att det finns goda förutsättningar för att ytterligare utveckla framtida redovisningar.

Bedömning 2

Syfte

Ett explicit angivande av syfte har ej hittats. Däremot kan man slutleda sig till att syftet är att undersöka huruvida Konsumentverkets verksamhet har givit goda resultat i förhållande till kostnaderna för verksamheten beträffande varudistributionen i glesbygd.

Den behandlade perioden 1973–1990.

Metod – dokumentation

Rapporten innehåller åtskillig dokumentation: utförlig beskrivning av Konsumentverkets verksamhet, angivande av andra utredningar och utvärderingar samt kvantitativ information. Emellertid ger ej denna dokumentation läsaren möjlighet att bedöma tillförlitligheten i slutsatserna.

Notera dock att detta omdöme gäller det som är skrivet i den rubricerade rapporten. Omdömet kan mycket väl bli annorlunda om angivna referenser granskats. Detta har dock ej gjorts enär uppdraget har uppfattats avse en bedömning av just den rubricerade rapporten. Notera också vad som sägs nedan om angivna källor (se "data").

Metod – data

Rapportens väsentligaste del bör utan tvekan vara momentet "EFFEKTER". Avsnittets syfte är att visa effekterna av stödinsatserna. Om detta kan först sägas, att slutsatserna till stor del baseras på "gammal" information (de utredningar som texten refererar till är från 1978, 1979, 1980 samt två från 1981; i dokumentationsförteckningen finns dessutom aktuella källor: från 1982, 1986, 1987 samt 1990).

Vidare bör läsaren notera att momentet "EFFEKTER" endast avser stödinsatser. I rapportens avsnitt "RESULTATREDOVISNING" behandlas ytterligare sex verksamhetsområden (se nedan under slutsatser).

Metodval

I inledningen till nyssnämnda avsnitt "RESULTATREDOVISNING" skrivs att beskrivningarna avgränsas "till ett urval aktiviteter där inriktning av verksamheten och uppnådda resultat i betydande grad påverkats av Konsumentverkets insatser".

Om detta måste sägas: vid utvärdering av en verksamhet får man inte begränsa sig till de områden där man uppnått goda resultat.

Metod – slutsatser

Rapporten behandlar som nämnts sju verksamhetsområden:

- a) stödinsatser
- b) varuförsörjningsplanering
- c) konsumentinflytande
- d) nya distributionsformer
- e) samordning av offentlig och kommersiell service
- f) vidgat serviceutbud
- g) förbättrad tillgänglighet i butikerna

I arbetet med dessa verksamhetsområden samarbetar Konsumentverket med flera andra organisationer, förutom den aktuella detaljhandlaren. Utfallet inom dessa områden är således ett resultat av dels de verksamma aktörernas beteende, dels utvecklingen i stort i samhället.

Därav följer att utfallet inte kan hänföras till just Konsumentverkets insatser, vare sig i positiv eller negativ mening.

I rapporten hävdas att Konsumentverkets insatser lett till goda resultat i följande verksamhetsområden: a, b, c, e, f. Beträffande område d säger rapporten att Konsumentverket inte ansåg sig ha resurser för insatser. Beträffande område g har ett projekt (En affär för alla) igångsatts 1981 men ej utvärderats.

Förutom den nyss framförda synpunkten att det med nödvändighet är svårt (omöjligt?) att hänföra utfall till Konsumentverkets arbete, betonas härmed att rapporten i sig inte ger läsaren möjlighet att bedöma om slutsatserna är korrekta.

Metod – kostnadseffektivitet

Denna frågeställning kan ej bedömas av mig. I yttrandets avslutning kommer dock att föreslås en alternativ uppläggning som förslag till fortsatt utvärdering.

Metod – objektivitet

I två avseende brister rapporten i objektivitet: dels ifråga om urvalet av verksamhetsområden, där rapporten – visserligen öppet redovisat – begränsar sig till de områden där Konsumentverket anser sig ha goda resultat.

Dels ifråga om rapportens upprepade slutsatser att insatserna lett till goda resultat utan att just denna rapport tydligt visar att så är fallet.

Sammanfattande bedömning

Rapporten är synnerligen rik på information. Ett övervägande antal rapporter i referensförteckningarna är dock av årgång 1979–81. Om man beaktar att det empiriska underlaget i denna typ av rapporter brukar vara ca två år gammalt, framgår att behovet är stort av förnyad och systematisk utvärdering av Konsumentverkets verksamhet.

Denna synpunkt förstärks av att rapportens slutsatser inte i (rapporten) är väl underbyggda.

Förslag till frågeställningar att utvärdera

1. Hur inverkar Konsumentverkets insatser – tillsammans med medaktörernas – på **konkurrensförhållandet** mellan olika detaljhandlare?
2. Hur inverkar Konsumentverkets insatser – tillsammans med medaktörernas – på **insatsmottagarens effektivitet**? Uppstår en lägre **motivation** för effektivisering när stöd erhålls?
3. Förekomsten av **fusk**? På andra områden där bidrag förekommer fuskas det ju en hel del.

4. Konsumentverket samarbetar ju med ett flertal **andra organisationer**. Vad anser dessa om samarbetet?

7 Värderad service – Utvärdering av Konsumentverkets stöd till den kommunala konsumentverksamheten (Konsumentverket)

Bedömning 1

3 Syftet med utvärderingen är att:

- ta reda på om verkets stöd till de kommunala konsumentvägledarna tillgodoser den lokala verksamhetens behov
- följa upp verkets servicegrad över tiden.

4 Utvärderingens metod

Samtliga kommuner med konsumentverksamhet har tillställts ett enkelt frågeformulär. Frågorna omfattar fem olika områden som Konsumentverket ansvarar för. Hur rutinerna fungerat inom dessa områden har Konsumentverket fångat upp genom att uppgiftslämnarna besvarat ett antal frågor med fasta svarsalternativ. Det har i huvudsak handlat om att ta ställning till om arbetet fungerat bra eller dåligt. Om något har fungerat dåligt har uppgiftslämnaren fått motivera varför.

De insamlade svaren har sammanställts i tabeller. Varje fråga har redovisats i två tabeller, dels fördelade efter storleken på kommunerna (antal invånare), dels efter kommunvägledarnas tjänstgöringstimmar.

Resultatet av de sammanställda tabellerna har legat till grund för analys och slutsatser.

5 A Dokumentation

De data som används, metoden och resultaten är beskrivna på ett sätt så att läsaren kan skapa sig en rimlig bild av utvärderingens resultat. Det man kan anmärka på är att värdena i tabellerna bara är angivna i procent och att det saknas redovisning av totaler för kommunerna. På flera ställen i texten redovisas uppgifterna som procent av procent. T.ex. "totalt 60 procent av dem som besvarat enkäten har dataprogrammet budgetrådgivning. Hela 96 procent av de som har budgetrådprogrammet tycker att det är bra". För att förenkla och för att undvika missförstånd är det ibland nödvändigt att komplettera andelar med absoluta tal.

5 B Data

Vad jag kan bedöma är de data som ligger till grund för utvärderingen tillräckligt tillförlitliga. Önskvärt är dock att tabellerna innehåller absoluta tal och kolumner för totaler. I texten skall man undvika att uttrycka frekvenser som andelar av andelar.

5 C Metod

Metoden är lämplig men det går att ifrågasätta om den är helt konsekvent tillämpad.

Slutsatserna grundar sig huvudsakligen på sammanställningarna av svaren på de relativt få fasta svarsalternativen. Det är ett enkelt tillvägagångssätt i jämförelse med att försöka sammanställa och analysera synpunkter som lämnats i form av öppna svar. Att öppna svar kräver större arbetsinsatser får inte leda till att sådan information får en underordnad betydelse. Allra helst som den, i den här aktuella studien, huvudsakligen omfattar argument och motiveringar till varför vissa saker har fungerat mindre bra. Dessutom är mängden lämnade motiveringar och synpunkter mycket omfattande.

5 D Slutsatserna

De slutsatser som dras är att Konsumentverket sköter sina åtaganden mycket bra. Undantag är området samarbetsformer, där "knappt två av fem anser att referensgruppen fungerar väl". Uttryckssättet att försöka tolka svaren så positivt som möjligt är kännetecknande för studien. Samma resultat skulle kunna uttryckas som att, "60 procent av vägledarna har ingen uppfattning om referensgruppen eller också tycker de att den fungerar dåligt". Andra exempel är att, "närmare hälften av vägledarna tycker att hämtfax underlättar deras arbete". Det skulle kunna uttryckas som, "mer än hälften av vägledarna tycker inte att hämtfax underlättar deras arbete". En slutsats som Konsumentverket inte borde vara helt nöjd med.

Slutsatserna grundar sig huvudsakligen på sammanställningarna av frågorna med de fasta svarsalternativen. Det rör sig om ett tjugotal frågor (har det fungerat bra eller dåligt?) om Konsumentverkets olika former för stöd till vägledarna. En genomgång och en analys av svaren på de öppna frågorna saknas däremot. Tillsammans utgör de en omfattande mängd information, nästan 19 fullskrivna A4-sidor, trots att endast ett svar tagits med om flera verksamheter svarat på ett likartat sätt. Huvuddelen av svaren är argument och motiveringar till varför vissa saker fungerat dåligt. En mer noggrann analys av dem hade förmodligen påverkat de slutsatser som Konsumentverket kommit fram till.

5 E Kostnadseffektivitet

Jag tror inte att det, med samma resursinsats, hade gått att göra en bättre utvärdering med avseende på hur uppgifterna samlats in och sammanstälts. Däremot kunde skrivandet, analysen ha gjorts bättre. Om hänsyn hade tagits till en systematisk genomgång av de 19 A4-sidorna med öppna svar hade stödet till den lokala verksamheten kanske inte framstått så positivt som utvärderingen hävdar.

5 F Objektivitet

Det är oftast svårt att vara helt opartisk. Än svårare blir det när någon skall granska/utvärdera den verksamhet som han/hon själv är ansvarig för. Det är den situationen som föreligger när Konsumentverket utvärderar sin stödverksamhet till de kommunala konsumentvägledarna.

Resultatet, Värderad service, uppfattas av läsaren som om de positiva omdömena har lyfts fram mer än de kritiska synpunkter som finns bland enkäternas öppna svar. En mer riktig beskrivning är kanske att olika problemställningar har tonats ned omedvetet.

6 Sammanfattande bedömning

Uppläggningsen – frågeformulering, datainsamling och sammanställning av insamlade uppgifter – finns i stort sett ingenting att anmärka mot. Däremot är det tveksamt om slutsatserna/resultatet kan betraktas som helt objektivt. Kritiska synpunkter har inte getts samma tyngd som positiva omdömen. En annan inriktning på utvärderingen – hur kan Konsumentverket förbättra stödet till konsumentvägledarna? – ha-de säkert varit mer givande för Konsumentverket.

Ett av syftena med studien var att följa upp Konsumentverkets servicegrad över tiden. Detta syftet har inte uppfylls. Det finns ingenting i utvärderingen som belyser eller behandlar hur verksamheten utvecklats.

Bedömning 2

3 Utvärderingens syfte

Enligt förordet i Konsumentverkets rapport är syftet att "Syftet med undersökningen är att se om verkets stöd och service till de kommunala konsumentvägledarna tillgodoser den lokala verksamhetens behov. Avsikten är också att kunna följa upp verkets servicegrad över

tiden." Detta är två syften som inte nödvändigtvis behöver utvärderas med samma metod.

4 Utvärderingens metod

Att med hjälp av en skriftlig enkät fråga samtliga konsumentvägledare vad de anser om det stöd som Konsumentverket ger konsumentvägledarna i deras arbete.

5 Bedömning av utvärderingen i olika avseenden

A Dokumentation

Den dokumentation som finns är en enkel resultatrapport där resultaten presenteras i kortfattade texter som verbalt redogör för resultaten samt tabeller och diagram. Dessutom finns en sammanfattning av resultaten. I rapporten finns också en enkel teknisk beskrivning av hur undersökningen genomförts. Av denna framgår att datainsamlingen genomförts med hjälp av postenkät till samtliga konsumentvägledare, dvs en totalundersökning och att andelen svar var 84 procent. I bilaga finns resultattabeller och frågeformulär. Dokumentationen gör att man kan skapa sig en bild av tillförlitligheten i materialet. Det som saknas är på vilket sätt och i vilket sammanhang som resultaten kommer att användas för att utgöra ett kunskapsunderlag i Konsumentverkets verksamhet. Det finns t ex inga slutsatser om hur väl Konsumentverkets stöd tillgodoser det stöd som den lokala verksamheten behöver. Detta tillsammans med det dubbla syftet gör att det svårt att bedöma valet av datagenereringsmetod. Förmodligen har den del av syftet som handlar om att följa verkets servicegrad över tiden haft avgörande betydelse för valet av metod.

B Data

De data som är underlag för utvärderingen (även om själva värderingen av verksamheten saknas i det studerade underlaget) utgår från Kon-

summentverkets perspektiv i och med att man utgår från det stöd som man för närvarande ger. Det framgår inte av underlaget om detta stöd överensstämmer med de krav och önskemål som konsumentvägledarna har på stöd från Konsumentverket. Det hade varit bra om en sådan utgångspunkt hade varit avstamp för utvärderingen. Man har alltså enbart studerat hur det stöd man ger idag fungerar och inte om det finns synpunkter på annat eller annorlunda stöd.

C Metod

I och för sig är valet av metod bra för att tillgodose syftet att följa hur verksamheten fungerar över tiden. Det kan även fungera för det andra syftet förutsatt att frågorna fångar in det väsentliga i verksamheten. För detta syfte bör man dock reflektera även kring andra metoder. Den valda metoden har sina fördelar då alla konsumentvägledarna har möjlighet att komma med sina svar. Det framgår inte hur det praktiska fältarbetet har gått till, men då hela 84 procent svarat utgår jag från att det har gjorts på ett tillförlitligt sätt.

Det hade varit bra att, oberoende av dagens stöd, man på ett mer förutsättningslöst sätt först samlat in de krav och önskemål som finns på stödet från Konsumentverket. Man kunde t ex göra detta via gruppintervjuer med referensgrupperna eller andra typer av gruppintervjuer/djupintervjuer med konsumentvägledare.

Det finns en del brister i undersökning av vad avser frågeformulering. Frågorna kunde ha formulerats så att de tillåter mer nyanserade svar. Så som frågorna är formulerade nu är det troligt att svaren ger en mer positiv bild än vad som är fallet i verkligheten. Metodstudier har det visat sig att det är lättare att avge ett positivt svar än ett negativt svar, särskilt om det bara finns två svarsalternativ och eller om vet ej alternativ saknas. Det är också tveksamt om de positiva svaren verkligen avspeglar kärntjänsten i det stöd som Konsumentverket skall ge konsumentvägledarna då tonvikten i stödet som efterfrågas ligger på hur rutiner fungerar i samband med materialutskick och tekniskt hjälpstöd i form av dataprogram och fax. Jag hade gärna sett att frågebatteriet utökats ifråga om rådgivningsstöd, utbildningsinsatser och samarbetsformer. Då svaren på dessa frågor tyder på att i stödet inte fungerar lika bra som de mer rutinmässiga stöd som ges. Hade dessa

frågeområden utvecklats mer så hade det troligen gett en mindre positiv bild av hur stödet fungerar.

D Slutsatserna

Tyvärr redovisas inga slutsatser i den dokumenterade utvärderingen.

E Kostnadseffektivitet

Värdet av utvärderingen hade också ökat om synpunkter och önskemål från konsumentvägledarna hade beaktats då val av frågeområden gjordes. Detta kan ha beaktats men det framgår inte av rapporten. Det jag föreslår är att man i studiens första fas gör någon form av kvalitativ datagenerering i form av t ex några gruppintervjuer med några konsumentvägledare, möjligen kan de olika referensgrupperna användas i detta syfte för att inte kostnaderna för detta skall bli för höga. Detta medför en extra kostnad jämfört med nuvarande uppläggnings, men jag anser att man vinner i trovärdighet och användbarhet av resultaten med en sådan uppläggnings. Resultaten från dessa samtal bör sedan utgöra stommen i det frågeformulär som tas fram. En frågeteknisk översyn av frågeblanketten kan rekommenderas för att nyanserade svarsalternativ används och att inte flera frågor ställs i samma fråga för att tolkningen skall underlättas. Slutligen bör tonvikten i frågorna läggas på Konsumentverkets kärntjänster. Särskilt om utvärderingarna är tänkta att bli ett löpande utvärderingsinstrument av måluppfyllelsen inom stödet tror jag att denna investering i mätinstrument lönar sig. Uppläggnings så som den är gjord är bra, då samtliga konsumentvägledare ingår i utvärderingen. Kvaliteten på datainsamlingen är hög i och med att andelen svar är 84 procent. Det är heller ingen anledning att genomföra en urvalsundersökning till en så liten målgrupp. Extrakostnaderna för att fråga alla är försumbar i det här sammanhanget. Fördelen är också att man då slipper den statistiska osäkerheten som alltid finns i en urvalsundersökning och de extrakostnader det innebär med urvalsdragning och viktberäkningar. Även om frågeformuläret ökar i omfattning till 8-10 sidor tror jag inte det påverkar svarsandelen då frågorna gäller stödet i det egna arbetet för konsumentvägledarna. Det har visat sig att ju mer angeläget som frågeom-

rådet känns för det tillfrågade desto mer benägen är man att svara. Om utvärderingen skall upprepas är det även angeläget att tala om i vilket sammanhang och på vilket sätt Konsumentverket dragit nytta av svaren.

F Objektivitet

Tillvägagångssättet att genomföra en postenkät och fråga alla leder till att man ej styr att enbart positiva får besvara frågorna, vilket garanterar en viss objektivitet. Samtidigt så är valet av frågor gjort av Konsumentverket, vilket gör att man kanske betonar frågeställningar där man är bra eller att man ställer frågor om sådant som är lätt att formulera frågor om. Jag misstänker att det kan ha gått till så eftersom frågeområden som jag uppfattar mer tillhör kärntjänsterna inte är särskilt grundliga, t ex personligt stöd, utbildning och samarbete. På några av dessa områden t ex fortbildningskurser och är inte konsumentvägledarna riktigt lika nöjda som med övriga frågor. Det indikerar att man borde se över dessa frågeavsnitt mer. Jag anser också att man bör byta perspektiv i frågeställningen och klart utgå från konsumentvägledarnas perspektiv när en sådan här utvärdering skall genomföras.

6 Sammanfattande bedömning

Först några reflektioner och allmänna frågor kring Konsumentverkets utvärdering. Jag undrar om denna utvärdering avspeglar huvudinriktningen för Konsumentverkets verksamhet eller görs även andra utvärderingar för att följa huvudmålen? Det vore också intressant att koppla kundvärdet av verksamheten och ställa det i relation till nedlagda resurser, såväl interna som externa. Några sådana resultat framkommer inte i den studerade dokumentationen.

Utvärderingar kan göras på många olika sätt och för många olika syften. Det är främst två aspekter som är viktiga att tänka på då en utvärdering skall göras. Det är

- Hur skall resultaten användas?
- Vilket djup på informationen efterfrågas?

Syftet med utvärderingen kan vara

- Resultaten skall ge övergripande information om verksamheten
- Resultaten skall ge ledningen information så att man kan följa verksamheten
- Resultaten skall ge de ansvariga för verksamheten information om hur verksamheten kan förbättras

Jag antar att Konsumentverkets utvärdering hör hemma under den typ av utvärdering som har till syfte att ge Konsumentverkets ledning information som gör att man kan följa den verksamhet som riktas till konsumentrådgivarna. Om detta är huvudsyftet tycker jag att metodvalet är riktigt, dvs att man väljer en metod som gör att man kan följa verksamheten över tiden och göra jämförelser. Då man gör en upprepad utvärdering är det viktigt att använda en metod som tillåter jämförelser och att också mätinstrumentet (frågeformuläret) är detsamma i de olika mätningarna så att jämförelser är möjliga. Jag skulle dock rekommendera att Konsumentverket såg över nuvarande mätinstrument så att det verkligen speglade kärntjänsterna och att frågorna formulerades så att trovärdigheten ökade.

Om syftet däremot är att använda resultaten i ett förändringsarbete, bör nog andra metoder övervägas, där ett större informationsdjup och konkretiseringsgrad i svaren är möjlig att nå. Då behöver heller inte objektiviteten vara någon helig ko som eftersträvas, utan då kan förändringar göras utifrån egen kompetens och erfarenhet parat med vetliga och genomförbara förbättringar som konsumentvägledarna efterfrågar. Hela tiden måste dock en realism finnas mellan resursåtgång och de resultat som uppnås av förändringarna.

I följande avsnitt ges en något fylligare beskrivning av de olika syften som ovan presenteras i punktform.

Övergripande information – image

Syftet med att kartlägga en organisations image är att ta reda på kundernas eller medborgarnas allmänna inställning till verksamheten. Image kan sägas vara den föreställning om t ex en tjänst som har vuxit fram genom erfarenheter, hörsägen, information eller påverkan från massmedia. Frågorna i en imageundersökning är allmänt ställda. Det går för det mesta bra att ställa ungefär samma frågor om så vitt skilda områden som skolan, sjukvården, skatteförvaltningen, försäkringsverksamhet och bilverkstäder. Tycker man att verksamheten är nyttig, att organisationen är framåt, att utförandet av tjänsterna sker obyråkratiskt och smidigt, etc, etc. Eftersom frågorna skall belysa en allmän attityd går det inte att dra slutsatser om konkreta brister i tjänsterna eller vad som format attityderna. Den information som man får från imagekartläggningar går därmed inte att använda som underlag för konkreta förbättringsåtgärder. Den tas vanligen fram för att kunna avläsa effekter av profilerade informations- och reklamkampanjer. Det är än så länge inte särskilt vanligt att sådana kampanjer görs inom den offentliga verksamheten. Den här typen av frågor blir mer intressanta om tjänsterna produceras på en konkurrensutsatt marknad. Vilket börjar bli verklighet för delar av den offentliga verksamheten. Det gäller ju att inte ha sämre image än andra som t ex erbjuder likartad vård, då kan ju patienterna välja dem som har en mer fördelaktig image.

Resultaten från imagekartläggningar blir mer intressanta då man har en tidsserie. Det är därför viktigt att genomföra upprepade kartläggningar med samma metoder och samma frågor. Eventuella förändringar i imagen skall bero på ändrade attityder och inte på förändringar i metod eller frågeformulering. Det innebär att imageundersökningar bör genomföras med metoder som ger resultat som kan jämföras över tiden. Vanliga och användbara metoder som tillgodoser dessa krav är postenkäter eller telefonintervjuer med ett korrekt draget slumpmässigt urval.

Imageundersökningar är en typ av undersökningar som blev vanliga när intresset för kundernas inställning började vakna inom den offentliga verksamheten. Detta kan vara en förklaring till att imageundersökningar blivit så populära trots att de ofta ger kunskaper som är mest matnyttiga för en informatör. Att alltför ensidigt påverka förväntningarna med information kan leda till problem. Det viktiga är ju

att förväntningarna stämmer överens med de erhållna tjänsterna. Ett annat skäl till imageundersökningarnas popularitet är att det var förhållandevis enkelt att föra över ett redan utvecklat koncept till den offentliga verksamheten. Flera av de genomförda imageundersökningarna visade sig sedan inte räckta till när mer konkret information efterfrågades.

Information till ledningen – uppföljningsinriktade

Informationen i de uppföljningsinriktade studierna är mer konkret och verksamhetsinriktad. Syftet med dessa kartläggningar är att ge de som har ett övergripande ansvar för verksamheten en bild av hur viktiga delar fungerar. Informationen bör enbart omfatta det som är av primärt intresse för denna ledning. Ofta försöker dessa kartläggningar studera måluppfyllelsen. Något som naturligtvis förutsätter att det finns mätbara mål. Det är därför viktigt att de mål som mäts verkligen pekar ut det viktigaste i verksamheten. Målen får inte formuleras utifrån det som går att mäta.

Det finns en tendens att studera det som kan studeras även om detta inte är det mest centrala för verksamheten. Faran är att det som studeras sedan pekas ut som mest väsentligt för verksamheten. I nästa led anpassas verksamheten till det uppmätta. Därmed inte sagt att man måste avstå från mätningar om det inte går att mäta allt som är viktigt, men det blir då särskilt viktigt uppmärksamma viktiga områden som inte mäts. Kunddialoger inom den offentliga sektorn handlar till alltför stor del om tillgänglighet och bemötande. Det borde finnas mer att fråga om kring själva kärntjänsten.

Det är viktigt att de uppföljningsinriktade undersökningarna mäter rätt saker på ett relevant sätt. Det skall ju ligga till grund för hur de ansvariga uppfattar verksamheten. De kan t o m vara underlag för sanktioner och resurstilldelning. Grunden för sanktionerna dvs den information som fås via kunddialogen måste uppfattas som relevant. Om inte de som berörs av sanktionerna uppfattar det som mäts som det mest väsentliga i verksamheten kan undersökningen ställa till med mer skada än nytta.

De metoder som väljs måste vara stabila över tiden och tillåta jämförelser mellan olika organisationer. Även för detta ändamål är det

vanligt att använda enkäter och telefonintervjuer med slumpmässigt urval. Utformningen av frågorna ställer särskilt höga krav i dessa undersökningar. De måste på ett tillförlitligt sätt mäta det som de är avsedda att mäta. Det är mycket viktigt att det inte uppstår någon diskussion om tillförlitligheten i resultaten. Delvis kan samma kvantitativa metoder som redovisas i föregående avsnitt användas även för detta ändamål. Nackdelen är att det är utfrågaren som bestämmer vad som är viktigt. Detta behöver inte överensstämma med kundernas preferenser. För att ta reda på det krävs andra metoder.

Inte heller resultaten från denna typ av undersökningar ger mer än i undantagsfall underlag för konkreta åtgärder.

För att ledningen skall kunna ställa rätt frågor i dessa uppföljningsinriktade undersökningar krävs att den har mycket god kunskap om kundernas förväntningar och behov. Om ledningen har felaktiga uppfattningar styrs verksamheten mot felaktiva mål. Kunskaper om detta skaffar sig ledningen med fördel genom att åtminstone då och då vara ute i verksamheten och föra löpande diskussioner med kunderna. Här finns mycket ogjort inom den offentliga verksamheten.

Information till de verksamhetsansvariga – åtgärdsinriktade

Ännu mer konkret och verksamhetsinriktad blir informationen från åtgärdsinriktade undersökningar. Dessa kartläggningar har till uppgift att leda till förbättringar av verksamheten. Informationen tas fram för de som arbetar i den operativa verksamheten. Det innebär att den här typen av kartläggning arbetar med de mest konkreta frågorna. Här gäller det att i detalj fånga upp de brister som kunderna ser i tjänsterna. Kunskaper om dessa brister skall sedan ligga till grund för att förbättra verksamheten. Information som erhålls är alltför detaljerad för de som har ett övergripande ansvar.

Åtgärdsinriktade kartläggningar måste anpassas och utformas efter lokala förhållanden. Frågorna måste vara så konkreta att svaren kan omsättas i direkta åtgärder. Det innebär att de frågeställningar som är aktuella inom ett område är helt ointressanta inom ett annat. Centralt utformade undersökningar utan utrymme för lokala frågor ger lätt otillräcklig information.

Här är frihetsgraderna när det gäller metodvalet betydligt större än i de föregående undersökningarna. De statistiska kraven är t ex inte lika höga. Resultaten skall vara tillräckligt säkra för att motivera ett förändringsarbete. Det är egentligen bara den egna fantasin som begränsar tillvägagångssättet för att få den efterfrågade informationen. Särskilt stor blir friheten om dessa undersökningar slipper de krav som hör hemma under image- och de uppföljningsinriktade undersökningarna. Det är vanligt att dessa ur metodsynpunkt mer krävande undersökningar och dess informationsbehov styr metodvalet med följd att den mer åtgärdsinriktade informationen blir för ytlig. Det är förmodligen en av de allvarligaste bristerna i de kundundersökningar som genomförs idag. Det går i och för sig att använda de statistiska undersökningar som nämnts tidigare men även andra metoder kan nu med fördel användas. Exempel på andra metoder är olika typer av gruppsamtal, observationer, dagbok, djupintervjuer mm. En mycket viktig informationskälla är de dagliga kontakterna med kunderna. Ibland kan det dock finnas ett behov av att låta de som deltar i undersökningen vara anonym då det kan finnas ett beroendeförhållande som gör att man vid en personlig kontakt inte säger allt man tycker för att slippa framtida tråkigheter. Ett sätt kan vara att byta personal med varandra. Någon från en annan avdelning kan genomföra djupintervjuer på den egna avdelningen och vice versa. Anonymiteten kan också förbättras genom metodvalet. De mer statistiskt utformade undersökningarna med slumpmässiga urval kan genomföras med rutiner som garanterar anonymiteten.

8 Rattfällan – Utvärdering av kurs för fängelsedömda rattfyllerister (Kriminalvårdsstyrelsen)

Bedömning 1

Utvärderingens syfte

Att utvärdera en kurs för fängelsedömda rattfyllerister (RATT-FÄLLAN).

Utvärderingens metod

Dels en registerstudie (analys av i polisregistret registrerade återfall i trafiknykterhetsbrott), dels en enkätundersökning, där alkoholvanor samt attityder till och kunskaper om alkohol och alkohol/trafik analyserades. Registerstudien följdes upp 1993 (4 års uppföljning).

Undersökningarna har utförts av Jan Törnros (VTI) på uppdrag av Trafiksäkerhetsverket och Kriminalvårdsstyrelsen (KVS). Vid uppläggningsen av undersökningarna har samarbete skett med personal på KVS.

Bedömningen

Dokumentation

Utvärderingen är redovisade på ett professionellt sätt. Redovisningen innehåller även sammanfattningar och litteraturhänvisningar, vilket ökar rapporternas läsbarhet.

Data och metod

Syftet med *Rattfällan* anges ha varit att genom framför allt kunskapsförmedling och attitydpåverkan försöka minska, och helst helt förhindra, återfall i trafiknykterhetsbrott. Utifrån denna utgångspunkt undersöktes eventuella kunskaps-, attityds- och beteendeförändringar. Enkät- och registerstudier är i sammanhanget vedertagna instrument.

Registerstudien bygger på uppdagad trafikbrottslighet, alltså inte den faktiska brottsligheten. Genom att välja en kontrollgruppsansats samt att utgå från det rimliga antagande att försöks- och kontrollgruppen inte skiljer sig med hänseende på upptäcktsrisk, är det faktiskt möjligt att (åtminstone delvis) studera eventuella faktiska beteendeförändringar efter genomgången kurs. Såsom det påpekas i rapporten ger detta tillvägagångssätt dock ingen möjlighet att besvara kausalfrågan, dvs. i vilken mån eventuellt funna skillnader beror på själva kursdeltagandet. Detta hade krävt en experimentell ansats. Jag bedömer det som försvarbart att utvärderingen inte lagts upp som ett experiment.

Som matchningsvariabler valdes ålder, intagningsmånad, kön, tidigare kriminalitet samt huvudbrott i aktuell dom. Speciellt ålder, kön och tidigare kriminalitet har i många kriminologiska undersökningar visat sig vara variabler av stor betydelse för variationer i återfall. Valet av dessa variabler får därför bedömas som helt tillräckliga för att standardisera jämförelsegrupperna på ett tillfredsställande sätt.

För 84 procent i "rattfällgruppen" hittades en perfekt matchningspartner. En bortfallsanalys har utförts. Erhållna skillnader mellan försöks- och kontrollgruppen har statistiskt testats med chi²-testet. Inga invändningar kan riktas mot detta tillvägagångssätt.

Värdet av 4-års uppföljningen kan diskuteras, men den förefaller ändå berättigad med tanke på den låga upptäcktsrisken för rattfylleri. Ef-

ter den långa uppföljningsperioden visade det sig att det inte bara finns statistiskt signifikanta, utan uppenbarligen också *relevanta* skillnader mellan undersöknings- och kontrollgruppen.

Enkätundersökningens syften angavs vara att utvärdera huruvida kursdeltagare förändrades på ett mer gynnsamt sätt vad gäller (1) attityder till alkohol i trafiken och (2) alkoholvanor, jämfört med andra fängelsedömda rattfyllerister. Dessutom avsågs att studera i vad mån kursdeltagarna erhöll förbättrade kunskaper om alkohol och alkohol/trafik. Man ville dessutom inhämta kursdeltagarnas åsikter om genomgången kurs. Enkätundersökningens uppläggning var ambitiös: kursdeltagarna fick besvara enkäter vid tre tillfällen, kontrollgruppen vid två tillfällen. Icke oväntat ökar bortfallen vid de senare undersökningstillfällena. Bortfallsanalyser har genomförts.

Personligen intar jag en något skeptisk hållning gentemot sådana undersökningar, eftersom sambandet mellan kunskap, attityder och avvikande beteende är oklart – såväl teoretiskt som empiriskt. Enkätundersökningar är dock accepterade som "normalvetenskap". Därför vill jag inte resa några invändningar bortsett från det redan gjorda påpekandet, att själva kausalfrågan inte kan besvaras med den valda undersökningsdesignen.

Slutsatserna

Nationella och internationella utvärderingar av behandlingsforskning på brotts- och missbruksområdet ger vid handen, att positiva effekter av olika behandlingsåtgärder i form av minskad brottslighet eller missbruk hör till undantagen. Föreliggande resultat tycks höra till de positiva undantagen: utfallet mellan försöks- och kontrollgrupperna går i många avseenden i undersökningsgruppernas favör. Som utredaren dock själv påpekar förblir slutsatserna osäkra på grund av den icke-experimentella undersökningsdesignen (se ovan).

Kostnadseffektivitet

Undersökningarnas kostnad framgår inte av bedömningsunderlaget. Därför kan jag inte ta ställning till frågan om kostnadseffektivitet.

Med tanke på att slutsatserna i framför allt enkätundersökningarna måste förbli osäkra skulle det kunna diskuteras om enkätundersökningarnas utformning inte var något överdimensionerad. Enkätundersökningarna brukar dessutom vara relativt kostsamma.

Objektivitet

Tillvägagångssättet i utvärderingarna, valet av frågeställningar, tolkningen av resultat m.m. ger – trots mina tidigare redovisade invändningar – inte anledning att ifrågasätta objektiviteten hos utvärderaren.

Sammanfattande bedömning

Det är mycket positivt, att utvärderingen utförts av ett utomstående professionellt utredningsinstitut. Även om utvärderingarna inte kunnat visa att det positiva utfallet för kursdeltagarna faktiskt berodde på genomgången kurs, tillåter utvärderingarna ändå slutsatsen, att kursen med största sannolikhet inte *försämrat* situationen för kursdeltagarna. Detta tillsammans med att kursdeltagarna (och förmodligen även personalen) *upplevt* kursen som meningsfull torde vara skäl nog att fortsätta med denna verksamhet.

Bedömning 2

1 Utvärderingarnas syfte

Sedan hösten 1984 bedrivs inom kriminalvården en kursverksamhet för fängelsedömda rattfyllerister (främst förstagångsdömda) som kallas Rattfällan. Till en början bedrevs verksamheten vid enbart två öppna lokalanstalter. Vid tiden för den första utvärderingens genomförande hade verksamheten införts vid ytterligare fem öppna lokalanstalter och tre öppna rixanstalter.

Deltagandet i Rattfällan är frivilligt och erbjuds personer som har rattfylleri i domen. De flesta fängelsedomda rattfyllerister får inget erbjudande att gå kursen. När den första utvärderingen genomfördes deltog mindre än var femte dömd rattfyllerist i Rattfällan.

Rekryteringen till kursen sker dels före anstaltsvistelsen, dels efter att den dömd har anlänt dit. Uppläggningsen av kursen varierar något mellan olika anstalter. Normalt genomförs kursen på halvtid (fyra timmar/dag) fem dagar i veckan under en tidsperiod om cirka en månad (tre veckor på en anstalt).

Syftet med Rattfällan är att genom främst kunskapsförmedling och attitydpåverkan till alkohol och alkohol/trafik påbörja en process avsedd att leda till en minskning av (och helst förhindra) återfall i trafiknykterhetsbrott.

Den första utvärderingen av Rattfällan består av två delstudier. Huvudsyftet med den första delstudien var att studera återfall i trafiknykterhetsbrott under en två-årsperiod efter frigivning från anstalten.

Den andra delstudien hade tre syften. För det första att undersöka i vad mån de som deltag i Rattfällan förändrades på ett mer gynnsamt sätt än vad andra fängelsedomda rattfyllerister gjorde avseende attityder till alkohol i trafiken samt alkoholvanor. Det andra syftet var att studera i vad mån kursdeltagarna erhöll förbättrade kunskaper om alkohol och alkohol/trafik. Det tredje syftet var att inhämta kursdeltagarnas åsikter om kursen.

Den andra utvärderingens syfte var att följa upp de individer som ingick i den första delstudien under ytterligare en två-årsperiod med avseende på återfall i trafiknykterhetsbrott.

2 Utvärderingarnas metod

Den första utvärderingen består, som tidigare nämnts, av två delstudier. Den första delstudien är en registerundersökning och grundas på de 818 personer som deltog i Rattfällan under åren 1986 och 1987. Åtta personer uteslöts på grund av felaktiga personnummer och femton personer hade avlidit vid tidpunkten för studiens genomförande. Vidare uteslöts 66 personer som inte hade trafiknykterhetsbrott i domen. Återstår alltså 729 kursdeltagare av vilka 648 hade rattfylleri som huvudbrott.

Dessa 729 kursdeltagare jämförs avseende återfall i registrerade trafiknykterhetsbrott inom en två-årsperiod efter fängelsestraffet med en lika stor jämförelsegrupp av individuellt matchade individer. Jämförelsegruppen består av individer, som i likhet med kursdeltagarna, var dömda till fängelse och där trafiknykterhetsbrott ingick i domen. Jämförelsegruppen konstruerades utifrån fem matchningsvariabler; ålder, intagningsmånad, kön, tidigare kriminalitet samt huvudbrott i aktuell dom (rattfylleri respektive annat brott).

Matchningen genomfördes i två steg. I det första steget gjordes en maskinell matchning i kriminalvårdsregistret med fyra av de fem matchningsvariablerna (undantag för tidigare kriminalitet). I matchningens första steg utgick man endast från huvudbrott i tidigare fängelse- eller skyddstillsynsdomar.

Den slutliga matchningen gjordes därefter utifrån polisregisterutdrag. Resultatet av matchningen blev att för 611 av de 729 personerna (83.8 procent) lyckades matchningen perfekt.

Skillnader mellan kursdeltagargrupp och jämförelsegrupp avseende återfall i trafiknykterhetsbrott signifikantestades därefter med hjälp av chi²-test.

Den andra delstudien är en enkätundersökning som riktades till personer med rattfylleri i domen och som gick Rattfällan under första halvåret 1989. Kursdeltagarna ombads besvara enkätformulär vid tre tidpunkter, ett vid intagningen på anstalten (67 frågor), det andra direkt efter kursens avslutning medan personerna fortfarande befann sig kvar på anstalten (55 frågor) och det tredje cirka sex månader efter frigivningen (23 frågor). Personer som hade avbrutit kursen (drygt ett tiotal) uteslöts ur undersökningen.

Den första enkäten riktades till 308 kursdeltagare. Bortfallet uppgick till nio personer. Av de 299 personer som besvarade enkäten hade 17 annat huvudbrott än rattfylleri och dessa uteslöts följaktligen. Av resterande 282 personer hade 256 män och 26 kvinnor rattfylleri som huvudbrott. Studiet av Rattfällans eventuella effekter avgränsades på det ringa antalet kvinnor, till att enbart avse män med rattfylleri som huvudbrott.

Av de 256 män med rattfylleri som huvudbrott som besvarade den första enkäten, besvarade 229 personer även den andra enkäten. Vid det andra tillfället gjordes också en kursvärdering.

I utvärderingen av Rattfällan görs en jämförelse mellan hur man har besvarat attitydfrågorna till alkohol dels innan kursen, dels efter kursen. Förändringen mellan dessa två tillfällen har testats med McNemar's Change Test.

Den tredje enkäten, sex månader efter frigivning, handlade huvudsakligen om förändring över tiden rörande attityder samt alkohol- och drogvanor. I enkäten ingick även frågor om livssituationen i stort. Kursdeltagarna gjorde också en kursvärdering. 168 personer besvarade även det tredje enkätformuläret.

För att bedöma Rattfällans eventuella effekter skapades en jämförelsegrupp bestående av rattfylleridömda personer (rattfylleri som huvudbrott) intagna under samma tidsperiod som kursdeltagarna, men som inte hade deltagit i kursen. Jämförelsegruppen besvarade två enkäter, dels vid inskrivningen på anstalten (67 frågor), dels sex månader efter frigivningen (21 frågor).

Den första enkäten till jämförelsegruppen vände sig till 731 personer. Av dessa vägrade 152 personer att delta. För ytterligare 14 personer saknas av andra anledningar ett ifyllt första enkätformulär. Följaktligen besvarade 565 personer i jämförelsegruppen det första enkätformuläret. Av dessa besvarade 416 personer det andra enkätformuläret.

Individuell matchning företogs mellan de 168 manliga kursdeltagare med rattfylleri som huvudbrott som även hade besvarat den sista enkäten och ett lika stort antal personer ur jämförelsegruppen som likaledes hade besvarat den sista enkäten. Matchningsvariablerna var ålder, kön och tidigare kriminalitet. Svaren på enkätfrågorna, som personerna i de två grupperna har besvarat, har chi²-testats för att se om det föreligger signifikanta skillnader mellan deltagargrupp och jämförelsegrupp.

I den andra utvärderingen av Rattfällan följdes de 729 personer som ingick i registerstudien upp under ytterligare en tvåårsperiod. Återfall i trafiknykterhetsbrott relaterades till den matchade jämförelsegruppen uppgående till 611 personer. Även denna gång analyserades resultaten med chi²-test.

3 Bedömning

3.1 Dokumentation

Den första utvärderingen är väl dokumenterad. Rapporten består av 104 sidor exklusive tre bilagor. Bilaga 1 innehåller de frågeformulär som riktades till de personer som ingick i enkätstudien; tre formulär till kursdeltagarna samt två formulär till personer i jämförelsegruppen. I Bilaga 2 redovisas kursdeltagarnas svar på öppna frågor i den kursvärdering som genomfördes direkt efter avslutad kurs. Bilaga 3 redovisar de matchade personernas (168 personer i varje grupp) svar sex månader efter anstaltsvistelsen på en öppen fråga om personernas livsföring hade förändrats sedan anstaltsvistelsen.

Rapportens överskådlighet dras dock ner av den rika förekomsten av tabeller; i stort sett redovisas svaren på varje enkätfråga i tabellform. Många tabeller blir det I rapportens 93 första sidor förekommer 114 tabeller.

Den andra rapporten – uppföljning av återfall i rattfylleri inom en fyraårsperiod – består av 15 sidor. Även här redovisas resultaten genomgående i tabellform (12 tabeller).

3.2 Data

Analysen av de personer som ingick i registerstudien, såväl uppföljningen inom en två-årsperiod som uppföljningen inom en fyra-årsperiod, baseras på två datakällor. Uppgifter om ålder, intagningsmånad, kön, tidigare kriminalitet under de senaste fem åren före den aktuella straffverkställighetens början samt huvudbrott i aktuell dom har hämtats från kriminalvårdsregister. Återfallsdata har inhämtats från polisregistret (s k fullständiga polisregisterutdrag). Dessa två datakällor torde kunna betraktas som mycket tillförlitliga.

Enkätstudien bygger, vilket framgår av namnet, på material som har insamlats via enkäter. I den del av enkätstudien som utnyttjar ett jämförelsegruppsförfarande var bortfallet stort. Av dem som avsågs att utgöra underlag för den matchade jämförelsegruppen vägrade 22 procent att medverka vid det första enkättillfället. Stora bortfall förekom, såväl för deltagargrupp som jämförelsegrupp, även vid uppföljningen

sex månader efter frigivning; var tredje bland kursdeltagarna jämfört med var fjärde av jämförelsepersonerna vägrade besvara den sista enkäten.

Enkätdata materialet kan, på grund av det betydande bortfallet, inte betraktas som helt tillförlitligt. I den första rapporten påpekas också att erhållna skillnader mellan kursdeltagargrupp och jämförelsegrupp bör tolkas med försiktighet.

3.3 Metod

Huvudsyftet med att utvärdera ett program är att analysera programets effekter på deltagarna, eller annorlunda uttryckt att uppskatta om deltagarna i något avseende har fått det "bättre" på grund av programdeltagandet än de skulle ha fått utan programmet. I Rattfällans registerstudie definieras utfallet i termer av återfall i rattfylleri. Den frågeställning som Rattfällan försöker att besvara är alltså "hur har deltagandet i programmet påverkat antalet återfall i rattfylleri jämfört med vad som skulle ha varit fallet utan programmet".

För varje programdeltagare existerar det ett hypotetiskt tillstånd att vara en icke-programdeltagare. Detta är själva kärnan i utvärderingsproblemet eftersom det inte är möjligt att observera samma person i två olika tillstånd vid en och samma tidpunkt. Det grundläggande utvärderingsproblemet blir då att identifiera tillståndet utan programmet. Detta kan endast göras med hjälp av personer som inte deltar i programmet men som har liknande egenskaper som programdeltagarna. Programmets effekt erhålls genom att jämföra skillnader i utfall mellan de som har deltagit respektive inte deltagit i programmet. Metodproblemet består i att göra denna jämförelse på ett adekvat sätt.

Två ansatser har dominerat utvärderingsforskningen. Den ena ansatsen är klassiska experiment. Denna ansats innebär att presumptiva programdeltagare slumpmässigt fördelas i en experimentgrupp, som får delta i programmet, och i en kontrollgrupp som inte får delta i programmet. Utfallet av programmet jämförs sedan mellan de två grupperna och den eventuella skillnad som kan uppmätas betraktas som effekten av programmet. Om man på detta sätt skatta effekterna av ett individinriktat program korrekt krävs att icke-deltagarna inte heller påverkas av programmet.

Den andra ansatsen är den icke-experimentella. Den innebär att man utgår från data över deltagare respektive icke-deltagare givet den selektionsprocess som avgör deltagandet. Selektionsprocessen kan vara ett resultat av antingen någon programadministratörers beslut och/eller den enskilde individen. När man i en icke-experimentell ansats jämför dem som har deltagit i programmet med dem som inte har gjort detta kan man inte utesluta att grupperna skiljer sig åt på ett systematiskt sätt även i andra avseenden än själva deltagandet i programmet. Det är därför inte en acceptabel utvärderingsmetod att enbart jämföra utfallet (effekten) mellan deltagare respektive icke-deltagare en tid efter genomgången program.

Givetvis förekommer det skillnader mellan deltagare och potentiella icke-deltagare såväl avseende observerbara som icke observerbara egenskaper. Om det endast föreligger skillnader i observerbara egenskaper är det tämligen enkelt att konstruera en jämförelsegrupp av icke-deltagare genom någon form av matchningsförfarande. Metoden innebär att man utväljer en grupp av icke programdeltagare som i de observerbara egenskaperna överensstämmer med programdeltagarnas. Om det däremot föreligger skillnader mellan grupperna avseende icke observerbara egenskaper, t ex skillnader i motivation, har vi ett s k selektionsproblem.

Den icke-experimentella ansatsen kan inte lösa detta selektionsproblem med hjälp av hur jämförelsegruppen konstrueras. En klassisk experimentansats löser problemet genom att slumpmässigt fördela individerna antingen som programdeltagare eller som icke programdeltagare.

Fördelen med experimentansatsen är att utfallet av ett program kan estimeras förväntningsriktigt och med en känd statistisk precisionsgrad. Den icke-experimentella ansatsen kännetecknas alltid av en okänd grad av icke-förväntningsriktighet beroende på korrelation mellan programdeltagande, utfall och andra icke observerbara egenskaper som påverkar utfallet.

När Rattfällan utvärderas används, förutom en före–efter jämförelse, en kvasi-experimentell ansats enligt principen individuell matchning, dvs till varje kursdeltagare utvaldes en jämförelseperson som var "identisk" med kursdeltagaren med avseende på vissa observerbara matchningsvariabler. Denna ansats ger ingen som helst information om selektion. Man antar implicit att selektionsproblemet eliminerats

genom matchning på ett fåtal (observerbara) variabler. Matchning på icke observerbara variabler är omöjligt och matchning på många variabler är tekniskt svårt.

Jag är inte övertygad om att ett matchningsförfarande löser alla problem med självselektion. Om de individuella egenskaper, som avgör om en person väljer att delta i ett program respektive inte gör det, är direkt observerbara är det dock i princip möjligt att särskilja effekten av programmet från effekten som beror på att deltagar- och icke-deltagargrupp skiljer sig åt. Emellertid är inte flertalet betydelsefulla individrelaterade egenskaper observerbara. I Rattfällans fall kan det vara sådana egenskaper som kunskaper, ambition, attityder, vilja att bearbeta sitt alkoholmissbruk m m. Om det skulle förhålla sig så att de som själva väljer att genomgå Rattfälle-kursen t ex är mer motiverade att bearbeta sina alkoholproblem än jämförelsegruppen så skulle troligen de förstnämnda ha klarat sig bättre (i förhållande till jämförelsegruppen), även om de inte hade genomgått kursen. I fall man inte tar hänsyn till detta, t ex med någon statistisk metodansats, kommer förmodligen en jämförelse mellan de två grupperna att överskatta kursens effekter.

Ett sätt att undvika detta problem är som sagts att slumpmässigt fördela individerna på experiment- respektive kontrollgrupp. Argumentet mot detta är främst av etisk karaktär. Det skulle vara oetiskt att utesluta vissa individer – kontrollgruppen – från att delta i ett program. Denna etiska invändning är dock inte lika stark beträffande Rattfällekursen; de flesta fängelsedömda rattfyllerister får överhuvudtaget inget erbjudande att gå kursen och kursen finns dessutom inte vid alla anstalter (åtminstone inte vid tidpunkten för utvärderingen).

Givet att man trots allt – som i utvärderingen av Rattfällekursen – valt en kvasi-experimentell ansats finns det olika statistiska metoder som beaktar att beslutet att delta i ett program har fattats av deltagarna själva. Någon sådan metod borde ha använts för att uppskatta om Rattfällekursen har haft någon effekt på antalet återfall i trafiknykterhetsbrott.

Utvärderingen av Rattfällekursen demonstrerar också en i praktiken ofta förekommande konsekvens av matchning: En stor förlust av observationer, vilket kan medföra svårigheter att generalisera resultaten till den ursprungliga populationen.

3.4 Slutsatserna

Givet den kvasi-experimentella ansatsen och givet den analysmetod som valts (i huvudsak signifikansprövningar av skillnader mellan deltagargrupp och jämförelsegrupp med hjälp av chi2-test) kan slutsatserna som dras i utvärderingarna som dras i utvärderingarna betraktas som underbyggda.

3.5 Kostnadseffektivitet

Som framgår ovan anser jag att utvärderingen av Rattfällan, vad avser de delar där jämförelsegrupper har använts, avsevärt skulle ha vunnit på om en annan ansats än den kvasi-experimentella hade tillämpats. En klassisk experimentansats är inte ekonomiskt kostsammare jämfört med en kvasi-experimentell ansats. Till skillnad från vad som är fallet vid utvärderingar av många andra individrelaterade program torde en utvärdering av Rattfällan med hjälp av en klassisk experimentansats kunna betraktas som etiskt acceptabelt.

Själva bearbetningen och resultatredovisningen skulle, också inom en oförändrad kostnadsram, ha vunnit på ett stramare tillvägagångssätt. I stället för redovisning av enkätsvar i tabellform, i stort sett fråga för fråga, borde i stor utsträckning någon form av sammanfattning av resultaten med t ex regressionsanalys ha gjorts.

3.6 Objektivitet

Utvärderingen av Rattfällan visar prov på godtagbar objektivitet. I den första rapporten diskuteras t ex med en kvasi-experimentell ansats där matchning görs utifrån observerbara egenskaper. Utvärderaren konstaterar att det dock är omöjligt att uttala sig om i vilken utsträckning t ex icke observerbara egenskaper har påverkat erhållna skillnader mellan kursdeltagare och jämförelsegrupp vad avser såväl återfall som attitydförändringar.

Även bortfallsanalys har genomförts. Utvärderaren sticker inte heller under stol med att de stora bortfallen kan ha påverkat de uppmätta

skillnaderna mellan kursdeltagare och jämförelsegrupp, varför de erhållna resultaten bör tolkas med försiktighet.

4 Ytterligare synpunkter

Personligen anser jag att det hade varit intressant om utvärderingen av Rattfällan hade utvidgats till att även beakta kostnader och intäkter av den bedrivna verksamheten. Detta borde inte minst vara intressant ur ett myndighetsperspektiv; Kriminalvårdsstyrelsen är tillsammans med Trafiksäkerhetsverket uppdragsgivare för den första utvärderingen och Kriminalvårdsstyrelsen är ensam uppdragsgivare för den andra utvärderingen.

En utvärdering av en myndighets verksamhet inom något område bör idealt ligga till grund för en diskussion om huruvida den bedrivna verksamheten skall omprövas i något avseende eller ej. Utan att beakta kostnader och intäkter, som i utvärderingarna av Rattfällan, blir beslutsunderlaget tunt.

9 Goda råd om fisk – Utvärdering av Livsmedelsverkets information till gravida och ammande (Livsmedelsverket)

Bedömning 1

Utvärderingens syfte

Det angivna syftet är brett – att belysa om Livsmedelsverkets information i en viss fråga – kostråd om fisk till gravida och ammande kvinnor – nått ut till målgruppen, varit begriplig och lett till ändrade kostvanor.

Det faktiska syfte som väglett studien tycks mig ha varit mer begränsat, nämligen att undersöka de *direkta* effekterna (effekterna på de som fått broschyren i sin hand) av en *broschyr*. Jag diskuterar problemen med denna avgränsning nedan.

Metod

Verket har låtit SCB genomföra en brevenkät till kvinnor i målgruppen i två regioner. Svaren analyseras med korstabeller. Konfidensintervall anges.

Min bedömning av utvärderingen

Övergripande bedömning

Enkätundersökningen är väl utförd. Den innehåller några enkla och tydliga frågor. Eftersom svarsfrekvensen är hög är resultaten lätta att tolka. Trots detta är jag tveksam till undersökningens nytta. En sådan nytta kan studien ha om den ingår som en del i bredare analys. Möjligen har en sådan mer omfattande analys gjorts, men jag väljer att tolka det faktum att något ytterligare material inte har bifogats utredningen som att detta saknas. Med denna förutsättning blir min övergripande slutsats att studien är så avgränsad att den inte ger något meningsfullt svar på frågan om den verksamhet som den utvärderar varit framgångsrik. Jag utvecklar här denna kritik innan jag går genom checklistan.

1 Kostnadsnyttoanalys

Yttersta syftet med en utvärdering av detta slag är rimligen en kostnadsnyttoanalys: var åtgärden värd sin kostnad? Som i många andra fall är det här knappast möjligt att exakt värdera nyttan. Det finns emellertid ingen anledning att avstå från att genomföra de delar av en kostnadsnyttoanalys som lätt kan göras. Några kostnadskomponenter kan enkelt uppskattas och något kan sägas om nyttan, se nedan. Dessa steg i analysen kan föra mycket långt.

Det är förmodligen förenat med betydande kostnader att genomföra denna kostrådgivning till landets alla blivande eller nyblivna mödrar. En kostnadspost är Livsmedelsverkets direkta kostnader för framställning och spridning av information. En annan viktig kostnad uppkommer vid informationen på mödravårdscentralerna (MVC). Tidsåtgången kan tänkas leda till ökat personalbehov på MVC och/eller att andra angelägna rådgivningsfrågor får mindre tid och uppmärksamhet. Slutligen kan det även finnas kostnader för mottagarna av informationen, de blivande eller nyblivna mammorna (och indirekt deras barn), t ex därför att några av dem blir nedstämda av att få ytterligare ett problem att oroa sig för. Av dessa kostnadsposter ger studien enbart vissa bidrag till information om den sistnämnda. Kostnaden för broschyren ger studien enbart vissa bidrag till information om den sistnämnda.

Kostnaden för broschyren borde emellertid vara enkel att beräkna. Med ett litet antal telefonsamtal till några MVC borde man kunna få en god bild av hur krävande personalen på MVC anser att informationsarbetet varit/är.

Nytan består främst i en förväntad positiv hälsoeffekt på spädbarnen. Därutöver finns kanske även indirekta effekter som att det bidrar till en ökad trovärdighet för såväl mödravården som Livsmedelsverket att man ger seriösa besked i en fråga som oroar många mammor. Nyttan är emellertid svår att fastställa. Dos-responssambanden är mycket osäkra. Verkets rekommendation bygger sannolikt på försiktighetsprinciper. Det är emellertid ändå viktigt att kunna göra även en mycket grov bedömning av nyttoeffekten. Världen är full av möjliga projekt som har liten men positiv nytta, prioriteringar måste göras. MVC har begränsade resurser och mammorna har begränsad förmåga att ta till sig all information. Frågan om hur mycket krut som skall spillas på fiskfrågan måste ses i ett större sammanhang. Kanske bör mödravårdens insats ännu mer fokuseras på t ex information om rökning?

Vissa insikter i dos-responssambanden skulle kunna underlätta användningen av enkätundersökningen. En viktig fråga är om sambandet är linjärt. Förmodligen kan hälsoeffekten (sannolikhet för hälsoskador) förväntas öka mer än linjärt med dosen. I så fall är nyttoeffekten i särskilt hög grad beroende av att informationen förmår påverka kostvanorna bland storkonsumenter. Utan en kompletterande information om detta är det svårt att värdera enkätens resultat. Eftersom Verket rekommenderar ett nollintag visar studien att det finns ett problem eftersom många kvinnor enligt enkäten brukade äta den hälsoskadliga fisken. Är detta ett stort problem, som motiverar höga åtgärdskostnader, givet att enkäten visar att mindre än 1 procent har brukat äta sådan fisk mer än en gång i veckan (och cirka 90 procent högst en gång i månaden)?

Dessa synpunkter är naturligtvis ingen kritik mot den redovisade studien i sig, utan pekar på den kompletterande information som behövs för att studiens resultat skall kunna användas i en utvärdering.

2 Informationssammanhanget

Studien tycks mig utgå från en mycket enkel syn på såväl mammornas tillgång till och förmåga att utnyttja olika informationskanaler. Jag själv, som bara har skolbarn, känner till Livsmedelsverkets rekommendation till blivande och ammande mammor, för det har jag läst om i tidningen. En koll med närmaste nyblivna tvillingmamma visar att hon inte sett någon broschyr men fått utförlig muntlig information på MVC. Informationen från Livsmedelsverket kan således spridas på fler sätt än genom direkt läsning av en broschyr. Många av dem som nås av broschyren har redan nåtts av informationen. Detta diskuteras inte i studien som enbart mäter a) hur många som tagit del av broschyren och b) ändrat sina kostvanor *efter* att de tagit del av broschyren. Detta innebär förmodligen en betydande underskattning av effekten av Livsmedelsverkets information. Jag förstår inte varför man har gjort en så snäv avgränsning av enkätundersökningen. Med några få kompletterande frågor borde man väl ha kunnat få en god bild av hur många som fått del av informationen på andra vägar.

A Dokumentation

Enkätundersökningen är dokumenterad på ett utmärkt sätt. Detta är utfört av SCB och innehåller såvitt jag kan se allt man behöver veta om population, urval, olika typer av bortfall och fel, uppräkningsförfaranden osv. Enkäten återges inte, men det förefaller som om enkätfrågornas exakta ordalydelse finns i resultattabellerna, vilket jag tycker är OK.

B Data

Studien insamlar data som är lämpliga och tillförlitliga, men begränsade.

C Metod

Metoden är OK. Mer avancerad statistisk analys är möjlig (t ex en logit eller probit-modell för att studera hur sannolikheten att påverkas av broschyren förklaras av olika faktorer), men jag ser inte någon poäng med det här.

D Slutsatser

OK, men begränsade.

E Kostnadseffektivitet

De kompletteringar jag pekat på ovan skulle inte kräva någon stor resursinsats (särskilt inte i förhållande till kostnaden för att genomföra en enkätstudie av detta slag).

F Objektivitet

Jag ser inga problem på denna punkt.

Bedömning 2

Livsmedelsverket gav våren 1992 ut broschyren "Goda råd om fisk" riktad till gravida och ammande kvinnor. Broschyren distribuerades via samtliga mödravårdscentraler. Syftet med utvärderingen var att belysa följande frågor:

- Har verket nått ut med informationen till målgruppen gravida och ammande kvinnor?

- Är informationen begriplig?
Har kostvanorna ändrats i och med att informationen givits?

Utvärderingen gjordes av en extern konsult, Statistiska centralbyrån (SCB) i samarbete med Livsmedelsverket som har hjälpt till med att utforma frågeblanketten. Undersökningen inriktades på barn i åldern 3 till 5 månader och utformades som en urvalsundersökning från en population i riksområdena (definierade som Kronobergs, Jönköpings, Kalmar, Örebro, Värmlands, Kopparbergs, Gävleborgs, Västernorrlands län samt Västerbottens landskap) samt i Stockholm. Ett representativt sannolikhetsurval drogs, varav 700 i riksområden.

- A. Den statistiska metodologin som användes i undersökningen beskrivs mycket detaljerat i avsnitt 5 i rapporten, under titel Teknisk rapport. Detta avsnitt upptar en stor del av hela utvärderingsrapporten.
- B. Data som används är lämpliga och tillräckligt tillförlitliga.
- C. Utvärderingsmetoden är lämplig dock anges inte någon närmare motivering varför även Stockholm ingår i undersökningen och vad jämförelserna med Stockholm egentligen syftar till.
- D. Den vägda svarsprocenten blev 92 i riksområden och 83 i Stockholm, vilket är ett bra resultat. Därmed får slutsatserna anses vara väl underbyggda.
- E. Kostnaden för undersökningen torde motsvara den externa konsultens (SCB) normala taxa för dylika undersökningar. Huruvida SCB bedriver sina urvalsundersökningar kostnadseffektivt är det omöjligt att avgöra inom ramen för denna bedömning. Det som kan kritiseras ur kostnadssynpunkt är att man borde försöka, eftersom urvalsundersökningar är dyra, att utvinna mer av relevant information ur undersökningen, se nedan.
- F. Objektiviteten i utvärderingen är godtagbar. Rapporten har dessutom skrivits av den externa konsulten (Birgitta Melin Skeppstedt, SCB) som knappast har en anledning att söka framställa Livsmedelsverkets verksamhet i en alltför positiv eller negativ dager.

Sammanfattande bedömning: Utvärderingen är, i sig, gjort på ett korrekt och objektiv sätt. Samtidigt kan dock den analys som presenteras

beskrivas som för snäv och bakåtriktad. Det övergripande syftet med att utvärdera redan genomföra informationskampanjer måste rimligtvis vara att förbättra effektiviteten inför framtiden. Ammande och gravida kvinnor är en viktig målgrupp för Livsmedelsverkets framtida verksamhet eftersom gruppen är mycket känslig för koststörningar. Ny information kommer med all sannolikhet att även i fortsättningen behöva riktas till gruppen i fråga. Mot denna bakgrund kan det påpekas att rapporten innehåller för mycket tabeller och för lite analys. Vissa tabeller är helt ovidkommande, t.ex. jämförelserna mellan riskområdena och Stockholm med avseende på antalet barn eller utbildningsnivån. (Den senare befinns, inte helt oväntat, vara högre i Stockholm.) I stället borde man på ett mer systematiskt sätt ha försökt att relatera bakgrundsvariablerna, som t.ex. utbildning, till kosthållning, benägenhet att ändra kost, sätt att söka information m.m. Det enda samband av detta slag som åberopas i rapporten är att broschyren om fisk befanns lättfattlig oavsett utbildningsnivån. Detta är naturligtvis ett gott betyg åt Livsmedelsverkets pedagogiska förmåga. Sammanfattningsvis kan det sägas att Livsmedelsverket borde i stället för att nästan helt överlåta arbete på den externa konsulten, arbetat tillsammans med SCB för att få ut mera av relevant information ur urvals-undersökningen.

10 Utvärdering av "Startpaket Nyckelhål – Intervjuer med restauranger och vidareinformatörer av kost- och hälsoinformation (Livsmedelsverket)

Bedömning 1

1 Utvärderingens syfte

Syftet med utvärderingen är en "uppföljning av kampanjen 'Startpaket Nyckelhål'" (s. 3). Kampanjen 'Startpaket Nyckelhål' är i sin tur en del av Kost- och hälsa-programmet, som syftar till att främja folkhälsan och förebygga sjuklighet och för tidigt död i olika kostrelaterade sjukdomar. Detta uppnås bl.a. genom att via information minska intaget av skadliga ämnen. Syftet med värdering av 'Startpaket Nyckelhål' kan därför uppfattas som en utvärdering av i vilken mån och hur man via den typen av information lyckats påverka kostvanor i önskvärd riktning.

2 Utvärderingens metod

'Startpaket Nyckelhål' består av information om näringsriktig och mager mat som främst riktar sig till restauranger och vidareförmedlare. För 50 kr kunde den intresserade erhålla ett paket innehållande: en broschyr med information om näringsriktig mat, skyltar, bordsryttare och dekalerna och etiketter. Symbolmärkningen ska då kunna ge kunderna information om att restaurangen i fråga erbjuder näringsrik och mager mat, främst lunch.

Metoden för utvärderingen är att intervjua ett antal av de personer som antagit erbjudandet i syfte att få information om "i vilken utsträckning man tagit del av och använt informationsmaterial; om man tillämpar symbolmärkning och i vilken omfattning; vilka erfarenheter och synpunkter det finns på kampanjen, materialet och på Livsmedelsverkets roll och funktion" (sid. 3-4).

3 Bedömning av utvärderingen

I det följande kommer jag att granska utvärderingen utifrån en mall för en "idealisk" utvärdering. Det ska påpekas att det sällan eller aldrig är möjligt att genomföra en sådan idealisk utvärdering. Mallen ska därför i första hand betraktas som ett tänkesätt för en utvärdering. Nedan följer en kortfattad beskrivning av mallen och därefter en bedömning av utvärderingen.

3.1 Principer för en utvärdering av projekt

En vägledande princip för en utvärdering av ett projekt är att undersöka om nyttan av projektet överstiger uppoffringarna (kostnaderna) för att genomföra det. Ett projekt kan då anses vara befogat när dess nytta överstiger kostnaderna. För att kunna genomföra en sådan analys måste man

1. Identifiera alla nytto- och kostnadsposter
2. Kvantifiera dessa poster
3. Värdera (i monetära termer) och aggregera posterna

Givet att vi uppfattar målet med 'Startpaket Nyckelhål' som att minska frekvensen av kostrelaterade sjukdomar, kan en sådan tre-steps-analys se ut på följande sätt.

1. Identifieringen av nytto- och kostnadsposter

Nytan av informationen består av därmed förknippade förbättringar av folkhälsan och eventuella minskningar i kostnader för att producera den näringsriktiga och magra maten. I princip uppstår det kostnader i tre olika led; administrativa kostnader hos de informationsförmedlande myndigheterna, eventuella förhöjda kostnader för ändringar i matlagning vid restauranger, och eventuella ökade kostnader för kunderna i form av prishöjningar.

2. Kvantifiering av nytto- och kostnadsposter

Nytan av kampanjen kan kvantifieras om man känner till dels dess in-verkan på intaget av näringsriktig och mager kost och dels motsvarande effekt på hälsan. Kostnaderna beräknas genom uppgifter från de olika agenterna beskrivna under 1 ovan.

3. Värdering av nytto- och kostnadsposter

Det finns för närvarande väl beprövade och standardiserade metoder för monetär värdering av förändringar i risker för olika allvarliga sjukdomar. Om man har information under 2 är det fullt möjligt att värdera nytan, dvs. förbättrad folkhälsa, i monetära termer. Detsamma gäller för skattning av kostnader i de olika leden, vilket torde vara tämligen enkelt att genomföra.

Det är däremot förenat med större svårighet att utvärdera om kampanjen är kostnadseffektiv, dvs. om det uppsatta målet uppnåtts till lägsta kostnad. För att kunna genomföra en sådan utvärdering måste man känna till alternativa möjligheter att påverka kostrelaterade sjukdomar såsom utbildning eller information direkt till restaurangkunder eller konsumenter som själva lagar mat. Ett ytterligare, lite mer ovanligt, alternativ är att avgiftsbelägga "ohälsosam" mat. För att kunna genomföra en analys om kostnadseffektivitet bör man alltså känna till samtliga möjliga åtgärder och därmed förknippade kostnader.

Med bakgrund av denna mycket kortfattade redogörelse av principer för utvärdering av ett projekt som 'Startpaket Nyckelhål' kan vi alltså

konstatera att det idag finns metoder som innebär att man skulle kunna komma långt på den redovisade tre-steps-analysen. Det är fullt möjligt att genomföra en utvärdering som innebär en monetär värdering av projektets förknippade nytta och kostnader.

3.2 Bedömning

Min utgångspunkt för bedömning av utvärderingen av 'Startpaket Nyckelhål' är de ovan redovisade möjligheterna att genomföra en tämligen fullständig nytto-kostnadsanalys. Som nämnts under kapital 2 har metoden för utvärderingen varit att man intervjuat ett antal personer som antagit erbjudandet om Startpaket Nyckelhål. Man har då ställt frågor om vilka som mottagit informationen och hur denna har använts. I detta avseende är rapporten mycket bra. På ett klart och tydligt sätt redovisar man resultaten på de olika frågorna. Resultaten redovisas efter hur personerna tillgodogjort sig informationen och om det resulterat i konkreta handlingar. Resultaten pekar på att ca 3 % av de tillfrågade har näringsberikat maten som en följd av kampanjen.

Utvärderingen har med andra ord lyckats ganska väl med att kvantifiera kampanjens effekter på restaurangers tillagning av näringsrik och mager mat. Om man däremot jämför utvärderingen med den under 3.1 skisserade mallen för en ändamålsenlig värdering, är det uppenbart att det föreligger stora brister. Man har förvisso identifierat nyttoposter (punkt 1) i form av förbättrad hälsa och eventuella kostnadsminskningar vid tillagning av mat. Däremot har man inte alls diskuterat alternativ till den genomförda kampanjen. När det gäller kvantifiering av olika poster (punkt 2) har man, som sagt, lyckats med att beräkna kampanjens effekter på restaurangers matlagning. Däremot finns det ingen diskussion om kostnaderna för kampanjen. Trots att det är möjligt att säga en hel del om värdering av kampanjens nytta (punkt 3) innehåller inte rapporten någon sådan diskussion. Inte heller finns det några som helst kostnadsberäkningar med.

Till fördel för utvärderingen ska dock anges att objektiviteten med all sannolikhet är garanterad genom att en utomstående konsult genomfört utvärderingen. Å andra sidan är denna konsult bunden av Livsmedelsverkets direktiv, vilket kan ha påverkat utvärderingen. Detta antyds genom "Huruvida det (kampanjens resultat) är bra eller då-

ligt kan bara avgöras utifrån de mål som från början ställts upp. Den intressanta frågan är vad som kan göras bättre." (sid. 20).

4 Sammanfattning

Mitt sammanfattande omdöme är att det föreligger stora brister i utvärderingen när denna bedöms efter de principer som redovisats under 3 och om målet är att utvärdera kampanjens inverkan på kostrelaterade allvarliga sjukdomar. Man kan emellertid ifrågasätta mitt val av bedömningsgrund. Om målet med utvärderingen istället är att undersöka i vilken mån Livsmedelsverket når fram med sin information, blir mitt omdöme mera positivt. Utvärderingen har då med stor noggrannhet dokumenterat resultaten. En svaghet är emellertid att man inte tagit med några som helst kostnadsberäkningar och alternativa informationsmedier och -former.

Bedömning 2

Utvärderingens syfte

Utvärderingen är en uppföljning av kampanjen "Startpaket Nyckelhål". Denna kampanj genomfördes under maj 1993. I ett säljbrev (folder) som skickades ut till 9.000 restauranger samt genom annonsering i tidskrifter (nämnda exempel är "Livstecknet" och "Vår Föda") gavs erbjudandet att köpa ett "startpaket" för 50 kronor styck. Kampanjen resulterade i beställningar av ca 1.000 "startpaket".

Syftet med utvärderingen var att belysa "i vilken utsträckning man tagit del av och använt informationsmaterialet, om man tillämpat symbolmärkningen, vilka erfarenheter och synpunkter det finns på kampanjen, materialet och Livsmedelsverkets roll och funktion".

Utvärderingens metod

Utvärderingen genomfördes våren 1994 av Långedrag's Informationskonsult AB. De restauranger och vidareförmedlare som hade beställt materialet (237 respektive 64 stycken) och vars namn fanns "tillgängliga" kontaktades för en telefonintervju. Efter korrigerig för externt bortfall (konkurser, lämnade tjänster etc) bestod urvalet av 55 vidareinformatörer och 199 restauranger. Svarsfrekvensen som erhöles var 81 respektive 88 procent.

Frågorna var en kombination av fasta och öppna svarsalternativ. Frågeformuläret var något olika utformat för de båda grupperna.

A Dokumentation

Rapporten ger en ganska tydlig bild av hur utvärderingen har gått till. Enkäterna finns med som bilaga, vilket ger möjligheter att se hur frågorna formulerades. Det hade dock varit ännu lättare att bedöma resultatens tillförlitlighet om frågorna, särskilt de öppna, hade angivits i anslutning till resultatredovisningen.

B Data

Redovisade uppgifter grundar sig enbart på intervjusvar. Minst 70 procent av de tillfrågade hade gjort sina beställningar i samband med att kampanjen genomfördes, dvs ca ett år innan uppföljningen gjordes. Det finns därmed en uppenbar risk för att de tillfrågade har glömt både hur materialet användes och vad de tyckte om detsamma. I utvärderingen kommenteras detta problem, och man framhåller att detta har medfört ett större internt bortfall, dvs att inte alla har kunnat svara på samtliga frågor. Men tidsaspekten gör också att de svar som har erhållits inte kan betraktas som särskilt tillförlitliga. Denna risk nämns inte i utvärderingen.

Svaren på öppna frågor har i stor utsträckning tolkats och kodats i efterhand. Vid vaga formuleringar finns därför risken att intervjuaren själv påverkar resultatens fördelning mellan olika (i efterhand uppgjorda) svarsalternativ.

Urvalet av intervjupersoner blev de namn på köpare som fanns "tillgängliga". I rapporten sägs att "vi inte vet i vilken grad dessa 300 personer representerar samtliga köpare". Detta är ganska kryptiskt uttryckt. Om Livsmedelsverket har en lista på de som har köpt "Startpaket Nyckelhålet", så borde väl samtliga köpare finnas med på denna? Eller har inte alla köpare registrerats? Dessa frågor får inga svar i rapporten. Det blir därmed svårt att bedöma svarens representativitet.

C Metod

Telefonintervjuer är oftast det enklaste sättet att få en hög svarsfrekvens. Frågorna bör dock vara tydligare preciserade och helst inte innehålla öppna frågor (se kommentar under B).

De resultat som erhålles jämförs aldrig med några på förhand formulerade mål med kampanjen. Det framgår inte i rapporten om några sådana mål för kampanjen (i kvantitativa eller kvalitativa termer) över huvud taget hade specificerats.

D Slutsatserna

Slutsatserna som rör svaren på hur restaurangerna och vidareförmedlarna har tagit del av och använt sig av materialet stämmer väl överens med den resultatredovisning som presenteras. Samtidigt finns den tidigare nämnda svagheten med svarens tillförlitlighet till följd av att utvärderingen gjordes så lång tid efter det att kampanjen genomfördes.

När man sedan kommer till diskussionen angående Livsmedelsverkets roll har dock slutsatserna betydligt svagare grund. Det är svårt att få en uppfattning om hur många som verkligen har svarat att de önskar mer material, kurser etc. För "vidareförmedlarna" finns en tabell 14, sid 18), där det framgår att 18 av de tillfrågade 55 personerna har uttryckt någon form av önskemål. Önskemålets art skiljer sig dock ganska mycket. Någon motsvarande uppgift finns inte för restaurangerna. Detta är en svag grund för påståendet på sid 19 "Förutsättningen för att man ska vara villig att delta i den här typen av kampanj är att verket servar med informationsmaterial — helst i större ut-

sträckning än vad som skett hittills. Många, både vidareförmedlare och restauranger, efterfrågar också mer personligt engagemang – kurser, utbildning, informationsträffar, handledning och uppföljning".

E Kostnadseffektivitet

Det som har krävt mest resurser i utvärderingsarbetet är troligen att genomföra drygt 200 telefonintervjuer. I stället för att intervjua så många som möjligt borde kraft ha lagts ned på att få ett representativt urval och att utforma frågorna mer precist (med slutna svarsalternativ). Med ett bättre upplägg av undersökningen kunde troligen mer tillförlitliga resultat ha erhållits samtidigt som antalet intervjuer kunde ha minskats betydligt (halverats). Man borde således ha låtit kvalitet gå före kvantitet i intervjuarbetet. För att också få de mer utförliga svar som möjliggörs av öppna frågor kunde telefonintervjuerna ha kompletterats med ett mindre antal besöksintervjuer.

F Objektivitet

Det är lite förvånande att man i utvärderingen inte har ställt sig frågan varför det var så få i den ursprungliga målgruppen som var intresserade av att beställa materialet. I rapportens sammanfattande figur 4 (sid 20) "Informationstratten" borde detta vara det naturliga första steget. I stället framhåller man att de flesta som ingår i undersökningen är positiva till att delta i folkhälsoarbetet samt att samarbeta med Livsmedelsverket. Åsikterna bland de som intervjuats har emellertid erhållits genom en öppen fråga, utan någon förutbestämd gradering av omdömena (att svara att något "är bra" är det enklaste om man snabbt vill avsluta intervjun).

I slutsatserna trycker man mycket på att restauranger och vidareförmedlare efterfrågar stöd från Livsmedelsverket med (gärna mer) informationsmaterial och personligt engagemang (exempelvis kurser, utbildning, informationsträffar). Jag finner emellertid inte mycket i svarsredovisning som belägger detta.

För att belysa den sista frågan i syftet, dvs erfarenheter och synpunkter på Livsmedelsverkets roll och funktion, kunde det ha varit gi-

vande att relatera denna studies resultat till andra utvärderingar av liknande aktiviteter. Om man i andra sammanhang har ställt frågan på annorlunda sätt kunde en sådan jämförelse ha givit värdefull information om resultatens värde.

Sammanfattande bedömning

När olika verksamheter utvärderas är det av vikt att resultaten på något sätt kan relateras till uppsatta mål för aktiviteten i fråga. I denna utvärdering beskrivs inga mål för kampanjen, vilket gör det omöjligt att dra någon slutsats om kampanjen var lyckad eller inte. Vad rapporten ger är snarare än beskrivning av hur olika restauranger och vidareförmedlare har agerat än en utvärdering av själva kampanjen.

Utvärderingen har många svagheter. Den gjordes lång tid efter att kampanjen genomfördes, de öppna frågorna skapar ett stort utrymme för godtycke vid tolkningen och slutsatserna angående Livsmedelsverkets roll är svagt underbyggda och granska tendentiösa. Man har heller inte på något sätt försökt besvara frågan varför så få överhuvudtaget var intresserade av att köpa "startpaketet". Till detta kommer frågetecknen om hur representativa resultaten är ens för gruppen köpare av "Startpaket Nyckelhålet".

Huvudsyftet med studien var att ge en bild av hur och i vilken utsträckning köparna har använt sig av materialet. Utvärdering ger en sådan, om än skakig till följd av ovan nämnda problem. Den bör emellertid ha kunnat öka kunskaperna om hur materialet har kommit till användning.

Utvärderingen förmedlar också en bild av att Livsmedelsverket borde göra större insatser. Frågan som snarare borde ställas är om det verkligen finns någon särskilt stor efterfrågan från restauranger och vidareförmedlare på insatser från Livsmedelsverkets sida.

11 Värdet av regionalpolitiken (Närings- och teknikutvecklings- verket)

Bedömning 1

1 Utvärderingens syfte

Utvärderingen redogör för utformningen av de olika regionalpolitiska stödformerna. Huvudsyftet är att analysera medlens enskilda och samlade effekter på region, företag och samhälle i stort. De resultat man kommit fram till angående medlens effekter ska sedan jämföras med de mål som satts upp för de regionala stödformerna. Det är också meningen att utvärderingen ska ligga till grund för framtida satsningar inom det regionalpolitiska området.

2 Utvärderingens metod

Den metod man använder sig av i huvudrapporten är att i diagram och tabeller med hjälp av statistiskt material för de olika stödformerna och sysselsättningen, aggregerat på branschnivå eller indelat efter de olika stödområdenas prioritet, visa trenderna för den regionalpolitiska inriktningen och den regionala utvecklingen. I bilagorna har man för det mesta gjort intervjustudier eller intervjustudier kombinerat med statistiskt material. De personer/organisationer som fått komma till tals i intervjustudierna är de företag som fått stöden och personer inom organisationer eller statliga/kommunala verk i stödområdena. Dessutom

har man i en bilaga utfört statistiska simuleringar i syfte att estimerade effekterna av regionalstöd på företagets betalningsförmåga på kort och lång sikt.

3 Bedömning av utvärderingens data, metod slutsatser och objektivitet

3A Dokumentation

Data, metod och slutsatser i analysen i huvudrapporten är beskrivet på ett sådant sätt att man kan göra en bedömning av utvärderingens tillförlitlighet. Vad beträffar de slutsatser som dras i sektion 4 (s. 43-65) är det ibland svårt att veta vilka resultat som hör till vilka bilagor. Några direkta hänvisningar görs inte.

I huvudrapporten (s. 75-119) sammanfattas de tio separata bilagor som ligger till grund för en del av slutsatserna i huvudrapporten. Slutsatserna beskrivs här ibland (men inte alltid) osammanhängande. Det är heller inte alltid lätt för läsaren att enbart med hjälp av denna sammanfattning få grepp om hur datan ser ut i de respektive bilagorna och framför allt gäller detta den metod man har använt sig av. I det senare fallet får man endast vaga föreställningar om hur de olika intervjustudierna har genomförts. Man måste i detta fall ibland gå direkt till de separata utvärderingarna.

3B Data och metod

I huvudrapporten används data för de olika stödförmerna och sysselsättningen – aggregerat på branschnivå eller grupperat för de olika stödområdena. Vid analysen av hur många arbetstillfällen som stöden har gett, måste det betecknas som anmärkningsvärt att man har använt sig den *utlovade* och inte den *faktiska* sysselsättningsökningen för de olika stödförmerna (s. 23). I huvudrapportens analys får vi därför aldrig reda på hur stor sysselsättning som genererats med hjälp av stöden. Detsamma gäller för investeringar; *planerade* investeringar visas istället för *faktiska investeringar*.

I bilagorna används för det mesta intervjustudier. Det är i första hand de företag som får eller har fått stöd som intervjuats, men även organisationer och kommunala/statliga verk i stödområdena får säga sin mening. *Utomstående företag/personer/organisationer utanför stödområdena får sällan eller aldrig komma till tals om de regionalpolitiska stöden.* (Ett undantag är en bilaga om utvecklingsbidragen där företag som fått avslag eller dragit tillbaka sin ansökan intervjuats.) Det är inte otroligt att de företag som fått stöd eller kommuner i stödområdena, vars sysselsättning och skatteintäkter är beroende av dessa företag, kommer att säga att stöden har varit till större hjälp än vad de egentligen har varit. Företag anger här ofta också att fusk förekommer men detta gäller bara *andra* företag (s. 94). Resultaten från intervjustudier beror på vem som ställer, respektive vem som får, frågan. Detta har endast i ringa omfattning beaktats.

I en enda bilaga till utvärderingen försöker man med kvantitativa metoder skatta effekterna av stöden. Statistiska simuleringar utförs för att estimeras effekterna av företagsstöd på företagets betalningsförmåga på kort och lång sikt (likviditet och soliditet). Man kommer fram till slutsatsen att ett positivt samband finns. Detta är dock inte förvånande, ty om man ger företag stöd i form av pengar så kommer naturligtvis deras finanser att förbättras. Frågeställningen i bilaga kan betraktas som felformulerad. Huruvida medlen ger effekter som uppfyller de mål som regionalpolitiken satt upp undersöks ej. Om medlen verkligen ger ökad sysselsättning eller fler "mjuka investeringar" i enlighet med regionalpolitikens mål för de olika stödformerna (som är beskrivna på s. 7-10) får vi med andra ord aldrig reda på i denna bilaga, eller i utvärderingen som helhet.

3C Slutsatserna

Många av slutsatserna baserar sig på intervjuer med de berörda företagen. Efter den kritik ovan beträffande vilka man har intervjuat så är frågan om man över huvud taget kan utvärdera effekterna av stöden från dessa intervjuer. Ofta dras slutsatser i huvudrapporten som är dåligt underbyggda eller resultaten tolkas till fördel för att stöden verkligen har haft den önskade effekten. Ibland dras t.ex. slutsatsen att stöden har gett långsiktigt livskraftiga bolag, för att i nästa stycke del-

vis motsägas av påståendet att sysselsättningsutvecklingen i stödbeviljade företag i ringa omfattning påverkats av att företaget erhållit stöd (s. 91). Alla bilagor är inte positiva till stöden och dess effekter, men huvudrapporten drar ofta långt mer positiva slutsatser än vad de enskilda bilagorna gör (se även stycket om objektivitet nedan).

Man ställer sällan frågan om investeringarna eller antalet anställda hade ökat oavsett om stödet hade givits eller ej. Utvecklingsbidragen, dvs. de stöd som ska användas till "mjuka investeringar", är dock ett utav två undantag (det andra exemplet är utlokaliseringen av tjänsteföretag, vilket är beskrivet nedan). Här anger de stödbeviljade företagen att 70-80 % av investeringarna hade skett även utan stöden (s. 54). I huvudrapporten menar man dock att denna siffra är överskattad – av någon anledning som inte klargörs. Normalt sett hade man kunnat förvänta sig att siffran varit underskattad, eftersom den angetts av stödmottagarna själva. Trots den stora andel av företagen som uppger att stöden egentligen är onödiga drar man i huvudrapporten slutsatserna att effekterna av utvecklingsbidragen kan betraktas som mycket positiva (s. 55). Stödets effekter på konkurrensen diskuteras heller aldrig i huvudrapporten. Att produktiva, ej stödbeviljade, företag kan ta skada av att andra företag får statliga stöd tar man ej hänsyn till.

I huvudrapporten drar man slutsatsen att den regionala obalansen har ökat under de senaste decennierna, vilket står utom allt tvivel. Men man påstår konstigt nog också att denna regionala obalans hade varit större än om stöden inte funnits, t.ex. säger man på s. 34: "Stöden har inte kunnat medverka till att skapa jobb i den omfattning som skett utan stöd i andra delar av riket. Utan stöd hade underskotten i stödområdena således varit ännu större". Man tar alltså här för givet eller antar att stöden verkligen har haft en positiv effekt på sysselsättningen i stödområdena.

Man hävdar även att stöden har haft effekter på den regionala och nationella strukturomvandlingen (s. 51). En närmare titt på den bilaga som ligger till grund för denna slutsats (*Regionalpolitiskt företagsstöd – Effekter på*) visar dock att stöden och den regionala strukturomvandlingen aldrig har kopplats samman. Data på trenderna för stöden och den regionala utvecklingen visas aggregerat i separata figurer/tabeller och några kvantitativa eller statistiska test har ej genomförts. Det kan därför lika gärna vara andra, ej observerade, faktorer som ligger bakom strukturomvandlingen.

Att slutsatserna i huvudrapporten och data och metod i hela utvärderingen är av bristande kvalitet ska dock inte förmörka att det finns många givande diskussioner och aspekter som tas upp i utvärderingen. Problem diskuteras som att stöden skapar bidragsberoende hos företagen eller förekomsten av överlappande stöd. En bilaga som undersöker stöden till utlokalisering av tjänsteföretag kan betecknas som intressant. Författarna menar att stöden egentligen inte hade behövts för att göra utlokaliseringen lönsam. Dock har stöden fungerat som incitament till utflyttning eller gjort företagen uppmärksamma på möjligheten att flytta. Fördelen med utflyttningen tycks enligt författarna vara att denna typ av företag ej missgynnas av långa avstånd (vilket är fallet för industriföretag) och att kostnaderna är betydligt lägre i stödområdena. Vidare kompletteras den traditionella strukturen i stödområdena med tyngdpunkt på jord/skog eller industri av arbetsplatser med en större andel kvinno- och tjänstejobb. Nackdelarna tar sig uttryck i att tjänsteföretag sällan skapar några indirekta effekter till bl.a. underleverantörer. Samtidigt betonas förekomsten av överlappande stöd.

3D Kostnadseffektivitet

I första hand borde bättre data för regionalpolitikens målvariabler ha samlats in, t.ex. den *faktiska* sysselsättningsökningen och de *faktiska* investeringarna som kommit till stånd tack vare stöden. Det förefaller underligt att detta inte varit obligatoriskt för samtliga företag som erhållit stöd. Tillgång till sådan data hade förmodligen möjliggjort kvantitativa och statistiska test mellan de olika stöden och målvariablerna. Om man använder sig av intervjustudier så borde även utomstående företag/organisationer/forskare ha fått säga sin mening om stödets effekter på sysselsättning och framför allt konkurrenssituationen. Dessa tillägg hade kunnat göras till en rimlig kostnad, samtidigt som man borde ha dragit ner på en del andra bilagor.

3E Objektivitet

Att det är NUTEK som analyserat effekterna av de stöd som man själv varit med och beslutat om gör att man kan ifrågasätta objektiviteten i hela utvärderingen. På flera ställen i huvudrapporten dras slutsatser att regionalstöden har haft de önskade effekterna, trots att detta är mycket dåligt underbyggt. Det hade varit mer lämpligt om utomstående personer, utan anknytning till NUTEK, hade gjort utvärderingen istället. Vissa av bilagorna har skrivits av konsulter på uppdrag av NUTEK. *Ett genomgående drag är att konsult-bilagorna tycks vara mer kritiska till stöden än vad man är i NUTEK-bilagorna.* Trots detta drar man i huvudrapporten, som skrivits av NUTEK, slutsatser från konsult-bilagorna som är mer positiva till stöden än vad slutsatserna är i dessa bilagor. En hel del medvetna eller omedvetna feltolkningar av bilagorna tycks således finnas.

4 Sammanfattande bedömning

Några empiriska belegg för att de regionalpolitiska medlen har haft de avsedda effekterna på sysselsättning och investeringar i de stödbeviljade företagen finns inte, eftersom man i utvärderingen inte har kopplat samman stöden med målvariablerna eller den regionala utvecklingen. Dessa variabler har istället analyserats var för sig och några tillförlitliga statistiska test har ej gjorts. Inte ens data för faktiska ökningarna i sysselsättning och investeringar har samlats in. Trots detta drar huvudrapporten på vissa ställen långtgående slutsatser om stödets positiva effekter på de stödbeviljade företagen och den regionala omgivningen. De bilagor som utförts av konsulter är mer kritiska till stöden än vad bilagorna gjorda av NUTEK är. I huvudrapporten har man dock dragit mer positiva slutsatser från konsult-bilagorna än vad man egentligen har gjort i dessa bilagor. Objektiviteten i huvudrapporten kan ifrågasättas eftersom det är NUTEK som utvärderar de stöd som man själv har varit med och beslutat om.

Den metod som används i utvärderingens bilagor är i första hand intervjustudier. De som fått komma till tals i dessa studier är dock främst de stödbeviljade företagen själva samt organisationer och statliga/kommunala verk i stödområdena, dvs aktörer som har ett intresse

av att stöden fortsätter. Utomstående personer/organisationer eller ej stödbeviljade företag utanför stödområdet intervjuas sällan eller aldrig. Med andra ord: *De som gett stöden har intervjuat dem som fått stöden*. Denna metod måste betecknas som otillförlitlig.

Trots bristerna i den data och den metod som används finns det flera givande och intressanta diskussioner i huvudrapporten och bilagorna, t.ex. problem med överlappande stöd eller att det finns tendenser till fortsatt bidragsberoende hos stödbeviljade företag. Även kommunikationsproblem mellan handläggare och bidragssökare tas upp, och stödets eventuella effekter på den regionala strukturen diskuteras. Det framförs också kritik från företagen själva att uppföljningen från statens sida är bristande när stöd har givits.

Det är en mycket stor brist i huvudrapporten att man aldrig diskuterar vilka effekter stöden har på konkurrensen. En snedvridning av konkurrensen kan definitivt inte uteslutas. De regionalpolitiska medlen kanske hjälper ineffektiva företag som egentligen inte har någon överlevnadsmöjlighet på lång sikt *på bekostnad* av livskraftiga och produktiva företag som inte får något stöd.

Bedömning 2

Följande kommentarer bygger på en genomläsning av huvudrapporten, speciellt kapitlen 3 och 4, samt punktvisa nedslag i bilagorna. Kommentarererna gäller alltså med viss reservation för material som kan finnas i rapportens övriga delar.

1 Utvärderingens syfte. Utredningens syfte är att utvärdera regionalpolitiken och dess medel mot bakgrund av regionalpolitiska mål, framför allt vad man kallar regional balans, dvs befolkningsutveckling, och näringsstruktur, där rikets struktur används som mall. Detta innebär att olika medel liksom politiken som helhet bedöms ur synpunkten i vad mån de bromsar en negativ befolkningsutveckling eller medverkat till en i någon mening gynnsammare näringsstruktur. Andra mål diskuteras inte, ej heller görs några försök till allmän utvärdering av välfärdseffekter.

2 Utvärderingens metod. Utredningen använder sig av ett antal olika metoder, bl.a. intervjuer. Framför allt synes bedömningarna vara baserade på den faktiska utvecklingen av sysselsättning, struktur etc. i jämförelse med riket totalt, i vissa fall också i jämförelse med "planerad" eller "utlovad" sysselsättning. I utredningen finns mycket litet av metoddiskussion, jämförelse med andra studier m.m., och inte heller någon diskussion av alternativa sätt att gå tillväga eller om för- och nackdelar med den/de valda metoderna, vilket kanske skulle ha varit av värde; här finns dock en ganska omfattande litteratur.

3. På grund av att så många metoder och källor använts är det svårt att ge något sammanfattande omdöme om datakvalitet, reliabilitet m.m. Primärmaterial synes i stor utsträckning bygga på case studies av företag, intervjuer med handläggare m.m., som ger möjlighet till viss fördjupning. Givetvis finns det felkällor även här, men man synes ha undvikit allmänna och ytligare enkäter till stödföretagen (av typen "hade ni någon nytta av stödet?") vilka kan väntas vara av mycket begränsat värde.

4. När det gäller effektbedömning borde kanske utredningen lite tydligare ha gjort klart skillnaden mellan planerad förändring, faktisk förändring och effekt, t.ex. av företagsstöd på sysselsättningen. För att beräkna effekter måste man jämföra den faktiska utvecklingen med den hypotetiska utveckling som skulle ha skett om åtgärden inte vidtagits. Att den faktiska ökningen är positiv, eller större (mindre) än motsvarande i riket eller den planerade ökningen, säger därför inte så mycket om effekten eller politikens effektivitet.

Utredningen är emellertid mycket försiktig när det gäller kvantifieringar. På basis av det material och de metoder som använts är det kanske inte heller meningsfullt att göra några sådana. Utredningens huvudslutsats synes vara att regionalpolitiken, i jämförelse med andra ekonomiska utvecklingstendenser som påverkar näringslivets lokalisering, har spelat en marginell roll, och att politiken i vart fall inte lyckats uppnå de uppställda målen om regional balans m.m.; häri bör man väl kunna instämma.

Utredningen kringgår alltså i stort sett mätproblematiken. Det går naturligtvis att med en mer ambitiöst metodval komma fram till någon form av kvantifiering av effekterna med hjälp av någon slags jäm-

viktsmodell. Frågan är emellertid om detta vore meningsfullt – regionalpolitikens effekt är troligen inte speciellt stora i jämförelse med andra drivkrafter (jfr ovan) – och om det kan göras inom samma kostnadsramar som man använt här; detta är svårt att avgöra.

5. En annan variabel som utredningen ger stor vikt är politikens effekter på näringsstrukturen. Om politiken tenderar att utjämna regionala strukturskillnader är den gynnsam. Det är emellertid inte helt självklart att regional specialisering är något negativt – här bör man kanske uttrycka sig försiktigare.

Utredningen argumenterar också för att det är gynnsamt om strukturen utvecklas i riktning mot ökade andelar för "nya" produktcykelvaror och kunskapsintensiv sektor med snabbväxande marknader. Det är också gynnsamt om stödet koncentreras till sådana områden. Detta görs delvis på basis av ett lite svårförståeligt argument om att stöd i expansiva branscher ger större nettoeffekter eftersom "bortträngningseffekten" blir mindre: rimligen blir "bortträngningen" densamma om alternativfallet innebär att en liknande produkt skulle ha tillverkats någon annan stans.

Utredningen klassar också olika stödformer med avseende på om de bromsar eller påskyndar "strukturomvandlingen", varmed avses en utveckling som närmar sig rikets struktur. Strängt taget innebär detta en välfärdsvinst endast om a. den optimala strukturen verkligt är densamma som rikets men b. omvandlingen bromsas av diverse trögheter.

Klassificeringen av olika stödformer på Ohlsson-Vinellska sektorer är i och för sig intressant men man kanske inte ska dra alltför stora växlar på den. Intressant är också utredningens bedömningar av effekterna på strukturomvandlingen av olika stödformer.

6. Vad gäller objektivitet kan man konstatera att utredningen har en viss distans till den förda politiken, i varje fall delar av den. Uppdelningen på "utvecklingsstöd" och "transfereringar" är belysande, liksom diskussionen av "bidragsberoende" och "bidrag som en rättighet" (här kunde det givetvis ha varit underhållande att göra en mer djupgående analys av den svenska regionalpolitiken ur ett public choice-perspektiv, om lobbying, rent-seeking m.m.). Andra omdömen av typen att handläggningen inom stödsystemen fungerar väl får väl tas med en nypa salt i den mån den bygger på omdömen från handlägg-

garna själva. Detta gäller möjligen också önskemålet om starkare samordning mellan stödformer och organ: sannolikt utgör väl poängen med att ha många oberoende aktörer på en riskkapitalmarknad en främmande tanke för stödbyråkratin.

Sammanfattningsvis kan man om utredningen säga att:

- a. underlaget för de slutsatser som dras i stort sett är beskrivet så att man kan följa resonemangen; däremot är det kanske svårare att bedöma reliabiliteten hos en del av primärmaterialet, speciellt enkäter/intervjuer.
- b. metodvalet har betingats av att utredningen undvikit att ge sig in på kvantifieringar; därmed kan analysen hållas på en mindre sofistikerad nivå. Givet denna utgångspunkt förefaller man ha extrahört mesta möjliga information ur sitt material.
- c. de försiktiga slutsatser som dras förefaller att vara hyggligt underbyggda av materialet och analysen, även om man kan ha invändningar på vissa punkter, t.ex. poängen med värderingen av strukturförändringarna eller värdet av de omdömen om olika stödformers effektivitet som erhålles från handläggare och företag.
- d. den typ av slutsatser som man har anledning befara i s.k. självutvärderingar, dvs myndigheter, utvärderar själva sin verksamhet, av typen "politiken är ytterst effektiv och handläggs på ett utmärkt sätt, men vi behöver mera pengar", i stort sett lyser med sin frånvaro.
- e. frågan om "värdet av regionalpolitiken" får väl strängt tagit inte något definitivt svar i utredningen (dvs om avkastningen är värd insatsen). Däremot ges en hel del konkret och intressant information om politikens utformning. Däremot ges en hel del konkret och intressant information om politikens utformning, fördelning på olika medel och deras för- och nackdelar, en information som kunde ingå i en mera övergripande bedömning. Det är ju möjligt att en metodiskt mer sofistikerad ansats skulle kunna ha gett mer inom ramen för samma budget, men det är svårt att avgöra.

12 Försöksverksamhet inom rehabiliteringsområdet i Malmöhus, Norrbottens och Blekinge län – en utvärdering (Riksförsäkringsverket)

Bedömning 1

3.

För budgetåren 1990/91 och 1991/92 anvisade regeringen medel till särskilda personalinsatser för försöksverksamhet för att utveckla försäkringskassans roll i rehabiliteringsarbetet. Två försökskassor utvaldes; härtill gavs genom särskilda projektmedel en tredje kassa möjlighet att delta. Kassorna skulle utveckla arbetsmetoder och samarbetsformer som främjar en effektiv rehabilitering och pröva nya verksamhetsformer inom området. Försöksverksamheten skulle så utformas att redovisning och utvärdering av erfarenheterna snarast möjligt kunde göras tillgänglig för övriga försäkringskassor. RFV har tillsammans med försökskassorna utarbetat metoder och mätinstrument för att analysera de nya arbetsformerna.

Försökskassorna har haft ansvar att följa upp och redovisa försöksverksamheten. Den här aktuella rapporten baseras väsentligen på försökskassornas egna slutrapporter och på statistiska analyser av det datamaterial som försökskassorna samlat in och på statistiska analyser av det datamaterial som försökskassorna samlat in och som bearbetats av RFV. Syftet har alltså varit att analysera och utvärdera de nya arbetsformerna. Ett viktigt syfte har varit att beräkna skillnader i det sk

ohälsotalet mellan försökskassornas lokalkontorsområden och övriga sådana områden i landet.

4. Utvärderingen har utgått från ett pluralistiskt metodologiskt perspektiv. Både kvalitativa och kvantitativa metoder har använts:

- a) RFV och försökskassorna har utvecklat ett system för resultatuppföljning. Med hjälp av kassornas datamaterial avseende hur längre sjukfall handläggs har kartläggningen gjorts före, under och efter försöksperioden.
- b) Statistisk analys i form av regressionsanalys har gjorts för ohälsotalet för att möjliggöra jämförelser mellan försökskassorna och övriga kassor. Data har i första hand hämtats från RVFs register och eventuella områdesskillnader avseende kön, ålder, civilstånd, inkomstfördelning mm har beaktats.
- c) Med användande bl a av resultaten från regressionsanalysen har en samhällsekonomisk kalkyl gjorts av försöksverksamheten.
- d) Vid två tidpunkter har RFV genomfört tidsstudier av försökskassorna och för riket i sin helhet.
- e) Två intervjuundersökningar har genomförts med personalen på försökskassorna och deras medaktörer (personal på arbetsförmedling, företagshälsovård osv).

5A Dokumentation

I rapporten hänvisas till tidigare rapporter som försökskassornas egna slutrapporter och RVFs tidigare uppföljningsstudier. Dessa har jag ej kunnat ta del av. Dessa skrifter finns emellertid tydligt angivna så jag menar att läsaren kan skapa sig en rimlig bild av utvärderingens tillförlitlighet.

B Data

De data som ligger till grund för utvärderingen syns lämpliga och tillförlitliga.

C Metod

Ett pluralistiskt metodologiskt perspektiv är viktigt i en utvärdering av detta slag. Tillämpningen verkar konsekvent. I rapporten redovisas vissa metodologiska frågor som t ex omöjligheten att skapa en kontrollgrupp som skulle kunna jämföras med försökskassorna eller svårigheter i samband med den samhällsekonomiska kalkylen. Enligt min mening har metodmedvetenheten varit hög. Det jag framför allt saknar vore en, åtminstone diskussion om, uppföljning av rehabiliterade personer. Vilka blev rehabiliterade, hur påverkades deras livssituation, vad hände med dem som inte fick arbete pga lågkonjunkturen osv?

D Slutsatserna

Försökskassorna förändrade sitt arbetssätt under försöksperioden, sjukfall granskades tidigt och samverkan etablerades med andra aktörer. En påtaglig positiv effekt uppmättes för ohälsotalet.

Slutsatserna är väl underbyggda i analysen. Man måste emellertid vara medveten om att det kan finnas en särskild "uppmärksamhetseffekt", som påverkar utvärderingar av här aktuellt slag. Försökskassorna hade fått extra resurser och de och deras medaktörer fick uppmärksamhet av regeringen, RFV, Försäkringskassaförbundet och utvärderingspersonerna, något som sannolikt påverkat resultatet positivt.

E Kostnadseffektivitet

Det torde knappast vara möjligt att göra en väsentligt mycket bättre utvärdering med ungefär samma resursinsats. Här bör observeras att utvärderingen enligt regeringsuppdraget genomförts under en begränsad tidsperiod och att utvärderingserfarenheter "snarast möjligt" skulle göras tillgängliga för övriga försäkringskassor.

F Objektivitet

Utvärderaren har visat godtagbar objektivitet.

6 Sammanfattande bedömning

I instruktionen till denna bedömning används uttrycket "självutvärdering". Den här aktuella rapporten är emellertid knappast en självutvärdering i gängse mening. RFV är ju tillsynsmyndighet över försäkringskassorna.

En utvärdering kan ha olika funktioner. Det kan gälla kontroll (nås uppställda mål?), legitimering (att motivera och bekräfta en verksamhet) eller kvalitetsutveckling (att direkt påverka). Den här aktuella utvärderingen har, enligt min bedömning, haft som funktion att legitimera försäkringskassornas rehabiliteringsarbete och att vara till nytta i försäkringskassornas fortsatta arbete. Sådillvida har utvärderingen fyllt sin funktion. Utvärderingen har visat upp förebilder och metoder för tidig och samordnad rehabilitering. Den har beskrivit processer och den har pekat på problemområden som exempelvis attitydpåverkan, förändringar i tjänstestruktur och behov av kompetensutveckling etc.

Som ovan nämnts har jag inga allvarliga metodologiska invändningar, givet regeringsuppdraget. På den tregradiga betygsskala som vi tillämpar enligt högskoleförordningen (underkänd, godkänd, väl godkänd) ger jag betyget väl godkänd.

Bedömning 2

3. Regeringen anvisade för budgetåren 1990/91 och 1991/92 särskilda medel för att bedriva försöksverksamheten inom rehabiliteringsområdet i två län. Medlen har varit avsedda dels för extra personalinsatser i syfte att vidareutveckla försäkringskassans roll i rehabiliteringsarbetet, dels för köp av rehabiliteringsinsatser. Medlen tilldelades Malmöhus och Norrbottens läns allmänna försäkringskassor men därutöver tilldelades även Blekinge läns allmänna försäkringskassa projektmedel i samma syfte.

Syftet med den utvärdering som här ska granskas har varit att bedöma utvecklingen vid de aktuella försäkringskassorna samt de uppnådda resultaten.

Syftet med försöksverksamheten har varit att utveckla nya arbetssätt och samarbetsformer i rehabiliteringsarbetet vilket på sikt kan leda till att människor snabbare än hittills får möjlighet att återgå i arbete efter sjukdom eller skada. Målet har varit att försäkringskassan inom 45 dagar skulle ta ställning till om rehabilitering kunde vara aktuellt i sjukfallet. I så fall skulle kassan initiera och samordna de åtgärder som kunde vara nödvändiga. Detta kräver utvecklandet av bättre samarbetsformer än tidigare med arbetsgivare, vårdcentraler, företagshälsovård och arbetsförmedling.

4 Utvärderingens metod

Den genomförda utvärderingen kan uppdelas i tre olika delar.

En första del utgörs av en beskrivning av tillskottet av extra resurser till de tre försökskassorna samt en statistik beskrivning av förändringar i vissa nyckeltal mellan 1990 och 1992 som t ex förändringar i den genomsnittliga "väntetiden" från den första sjukskrivningsdagen till dess att någon viss åtgärd vidtagits.

En andra del av utvärderingen utgörs av en omfattande beskrivning av vilka konkreta förändringar som skett i försäkringskassornas sätt att arbeta. Här beskrivs hur arbetet har organiserats i olika avseenden, hur rekryteringen skett till de nya arbetsuppgifterna, vilken självständighet eller grad av decentralisering som karaktäriserat verksamheten etc.

En tredje del av utvärderingen utgörs av en djupare vetenskaplig analys av försöksverksamheten baserad på ekonometriska standardmetoder.

5 Bedömning av utvärderingen

A Dokumentation. Utvärderingen ger en mycket god bild av den verksamhet som utvärderas. Speciellt innehåller utvärderingen en mycket omfattande beskrivning av de förändringar som skett i det praktiska arbetet med rehabilitering av sjukskrivna. Syftet med utvärderingen var också att den skulle kunna tjäna som vägledning för de försäkringskassor som inte deltagit i försöksverksamheten. Den de-

skriptiva resultatpresentationen är inte speciellt utförlig men fullt tillräcklig. När det gäller den djupare vetenskapliga analysen hänvisas till ett icke publicerat conference paper av E. Palmer och S. Olsson (1992) samt en mycket kortfattad redogörelse i en bilaga till utvärderingen. Det ligger kanske i sakens natur att en utvärderingsrapport av denna karaktär inte kan innehålla en alltför tung vetenskaplig del, men att enligt min uppfattning borde rapporten ingått som bilaga till utredningen. Det är speciellt motiverat i det här fallet eftersom den ekonometriska analysen ger en helt annan bild av projektresultatet än en ytlig beskrivning av de förändringar som skett vid försökskassorna. Det bör dock understrykas att utvärderingen innehåller tillräckligt med information för läsaren ska kunna skapa sig en rimlig bild av utvärderingens tillförlitlighet.

B Data. Tillgången till relevanta och tillförlitliga data synes ha varit imponerande god, och det är uppenbart att RFV satsat betydande resurser på att bygga upp tillförlitliga databaser som kan utnyttjas för en utvärdering av den egna verksamheten.

C Metod. Som tidigare angivits kan utvärderingen uppdelas i en statistisk deskriptiv del, en ekonometrisk analys samt en mycket utförlig konkret beskrivning av hur arbetet i olika avseenden organiserats och förändrats vid försökskassorna under projektets gång. Enligt min uppfattning är detta en mycket lämplig metod för den aktuella utvärderingen. Min kritik gäller framför allt att den ekonometriska analysen redovisats så styvmoderligt i utvärderingsrapporten. Efter att ha tagit del av den ekonometriska bakgrundsrapporten råder enligt min uppfattning inget tvivel om att den djupare vetenskapliga analysen är mycket kompetent genomförd och baserad på ett mycket relevant och tillförlitligt datamaterial.

D Slutsatsernas tillförlitlighet. De slutsatser som dras av den genomförda analysen förefaller mycket väl underbyggda. Detta gäller speciellt själva huvudresultatet vad gäller den samhällsekonomiska vinst som projektet resulterat i och som i huvudsak är baserad på den djupare ekonometriska analysen. Detta huvudresultat är baserat på en aggregerad bedömning omfattande de samlade effekterna av de tre försökskassornas förändringsarbete. Vad gäller huvuddelen av utvärde-

ringsrapporten är denna dock i huvudsak deskriptiv med få slutsatser och kritik av den verksamhet som bedrivits vid försökskassorna. Läsaren får egentligen aldrig riktigt klart för sig om någon av de tre försökskassorna framstår som mera framgångsrik än de övriga. Man får närmast intrycket av att ingen skall trampas på tårna. Att projektet som helhet uppvisar en mycket god samhällsekonomisk lönsamhet behöver inte betyda att varje del av projektet gör det. När det gäller att dra slutsatser för hela landet av försöksverksamheten måste också hänsyn tas till ett s k selectivity bias-problem. Detta består i att försökskassorna inte var slumpmässigt utvalda utan valdes ut efter en ansökningsprocedur. Det kan därför tänkas att det var de mest progressiva och förändringsbenägna kassorna som kom att ingå i projektet. En uppskalning av resultaten till nationell nivå kan därför innebära en överskattningsproblem av den samhällsekonomiska lönsamheten av att satsa pengar på rehabilitering. De samlade resultaten från projektet tyder dock på att den samhällsekonomiska lönsamheten av att utvidga projektverksamheten till samtliga försäkringskassor skulle ge en mycket hög avkastning på satsade medel.

E Kostnadseffektiviteten i utvärderingen. Som tidigare påpekats har jag inga principiella invändningar mot den utvärdering som är genomförd. I uppdraget ingick att beskriva hur verksamheten hade förändrats till ledning för övriga försäkringskassor samtidigt som den mera beskrivande genomgången har kompletterats med en ordentlig vetenskaplig analys av projektets resultat. Den senare är baserad på ett mycket tillförlitligt och relevant datamaterial, och det framstår som mycket förtjänstfullt att RFV satsat de resurser som krävs för att konstruera de databaser som utnyttjas i den aktuella utvärderingen, men också tydligen generellt i RFV's utredningsenhetens arbete.

F Objektivitet. Den ekonometriska analysen är baserad på en vetenskaplig rapport, som också presenterats på en vetenskaplig konferens. Den framstår som ett gediget hantverk och indikerar att utredningsenheten vid RFV har en betydande kompetens för mera avancerade evalueringar av den egna verksamheten. Den mera deskriptiva genomgången har också en mycket objektiv uppläggning och enligt min uppfattning alltför neutral framtoning. Jag hade gärna sett mera kritiska

kommentarer och vilka slutsatser som skulle kunna dras från en jämförelse mellan de tre försökskassorna.

6 Sammanfattande bedömning

Huvudintrycket av utvärderingen är mycket positiv. Detta beror framför allt på att utvärderingen innehåller en kärna av mera kvalificerad ekonometrisk analys baserad på en mycket relevant databas.

13 En studie av rehabiliterings- arbetet för försäkringskassan (Riksförsäkringsverket)

Bedömning 1

Inledning

Syftet med rapporten, såsom det uttrycks i det inledande kapitlet, är att presentera en allmän bild av de åtgärder som försäkringskassorna och andra rehabiliteringsaktörer har vidtagit genom rehabilitering för att begränsa sjukfallens längd. Av denna formulering kan utläsas att det huvudsakligen handlar om att beskriva. I begreppet utvärdering måste anses ligga att på något sätt försöka mäta resultatet av en verksamhet i förhållande till de mål som satts upp för verksamheten. Ett grundläggande problem med rapporten är således att den knappast är att betrakta som en utvärdering.

Det faktum att rapporten inte är en utvärdering ska naturligtvis inte läggas Riksförsäkringsverket till last om syftet aldrig varit att göra en utvärdering. I så fall bör det snarare ifrågasättas om studien hör hemma inom ramen för ESO-projektet. Om syftet från RFV:s sida var att göra en utvärdering från RFV:s sida så är det bara att konstatera att detta har misslyckats. Rapporten uppfyller inte de mest elementära kraven på en utvärdering. Den viktigaste bristen är att det i stort sett helt saknas en diskussion om målen för verksamheten och hur pass väl dessa uppfyllts.

Oklarheten om rapportens egentliga syfte försvårar naturligtvis bedömningen i övrigt. Det är dock möjligt att lämna några allmänna synpunkter på datamaterial, dokumentation m.m.

Data

Undersökningen baserar sig på material från tre källor: en intervjuundersökningen, en enkätundersökning samt uppgifter från Riksförsäkringsverkets register. Enligt uppgifter som jag kunnat inhämtat från oberoende källor torde datamaterialet genomgående vara av god kvalitet.

Dokumentation

Rapporten får anses leva upp till rimliga krav på dokumentation. Den innehåller ett fylligt tabellmaterial som gör det möjligt att verifiera de slutsatser som dras. De statistiska analysmetoderna beskrivs översiktligt, om än inte ingående. Redovisningen av datamaterialet lämnar dock en del att önska. Beskrivningen av insamlingsproblem, urvalsteknik, bortfall, enkätens respektive intervjuernas utformning m.m. är knapphändig eller saknas helt. Det saknas även en ordentlig beskrivning av det praktiska innehållet i de rehabiliteringsinsatser som studien behandlar.

Slutsatser

De slutsatser som dras i rapporten baserar sig på en grundlig genomgång av datamaterialet. Min bedömning är att slutsatserna är väl underbyggda. I de fall där det råder tveksamhet har författarna bemödat sig om att redovisa alternativa tolkningar. Problemet, sett ur utvärderingssynvinkel, är snarare att författarna avstår från att dra vissa slutsatser. Det handlar framförallt om sådana slutsatser av värderande natur som är naturliga inslag i en utvärdering (jfr nedan).

Kostnadseffektivitet

Enligt de uppgifter som jag inhämtat från oberoende källor har RFV lagt ned betydande resurser på att samla in det datamaterial som ligger till grund för studien. Det är svårt att se att de slutsatser som dras i rapporten skulle motivera ett så stort kostnadsåtagande. Det bör dock beaktas i sammanhanget att datamaterialet antagligen skulle kunna användas för andra syften än den aktuella studien. Frågan om kostnader och nytta måste därför bedömas i ett större sammanhang. Jag har bedömt att det ligger utom ramen för mitt uppdrag att göra en sådan bedömning. Det kan tilläggas i sammanhanget att de långa sjukfallen svarar för en mycket stor del av de samlade kostnaderna för den statliga sjukförsäkringen, varför det i princip kan vara motiverat att lägga ned betydande resurser på att utvärdera rehabiliteringsverksamheten. När det gäller frågan om huruvida det skulle varit möjligt att göra en väsentligt bättre utvärdering med ungefär samma resursinsats, så är svaret naturligtvis ja, eftersom den aktuella rapporten egentligen inte är en utvärdering. Det datamaterial som legat till grund för rapporten förefaller väl lämpat som underlag för en utvärdering av försäkringskassornas rehabiliteringsarbete.

Objektivitet

Mitt intryck är att rapporten i hög utsträckning präglas av ambitionen att ge en objektiv redovisning av de slutsatser som kan dras utifrån datamaterialet. Det bör dock beaktas i sammanhanget att studiens syfte inte är att utvärdera rehabiliteringsverksamheten. Kravet på objektivitet torde därmed ha varit förhållandevis lätt att uppfylla, eftersom risken för konflikter med Riksförsäkringsverkets intressen av att framställa den egna verksamheten i god dager rimligtvis har varit liten.

Sammanfattande bedömning

Mitt sammanfattande intryck är att rapporten innehåller grundlig, objektiv, intressant och förhållandevis väl dokumenterad genomgång av försäkringskassans rehabiliteringsarbete. Rapporten håller i dessa av-

seenden god standard. Problemet är snarare att själva syftet inte har varit att utvärdera verksamheten. Det som saknas är en diskussion om mål, måloppfyllelse och möjligheter till förbättring av verksamheten. Den vakne läsaren kan hitta en del information som leder i den riktningen, men det handlar mer om spridda stänk än om en systematiskt genomförd utvärdering. Dessutom lämnas det åt läsaren själv att dra slutsatserna. Ett exempel får illustrera detta. Det framhålls på flera ställen i rapporten att det är av stor betydelse att rehabiliteringsarbetet inleds tidigt under sjukskrivningen. Det framgår också av rapporten att andelen av de långa sjukfallen som blir föremål för rehabilitering är förhållandevis liten, att den sjunkit något mellan de två undersökningsperioderna samt att det i regel tar lång tid att innan rehabilitering inleds. En tänkbar (men inte självklar) slutsats av detta är att försäkringskassornas rehabiliteringsarbete fungerar mycket dåligt och att det dessutom verkar fungera sämre och sämre. Det som är symptomatisk för rapporten är att författarna inte uttryckligen berör frågan om hur väl försäkringskassorna lyckats i sitt arbete.

Det bör påpekas än en gång att ovanstående inte ska ses som en kritik av rapporten som sådan. Den innehåller i själva verket flera förtjänstfulla analyser. Det kan vara värt att framhålla ett exempel. Det gäller ett avsnitt som handlar om vilka faktorer som har betydelse för att en långtidssjukskriven person ska erhålla rehabilitering. Där framgår bl.a. att en ansökan eller prövning om förtidspension eller sjukbidrag har en tydlig negativ inverkan på sannolikheten för att erhålla rehabilitering. Detta trots att de som ansökt eller varit föremål för prövning om förtidspension eller sjukbidrag är överrepresenterade i bland de som erhållit rehabilitering. Förklaringen är att dessa personer ofta har långa sjukskrivningsperioder bakom sig, vilket i sig är en faktor som har en starkt positiv effekt på sannolikheten för att erhålla rehabilitering. Här är analysen väl genomförd.

Bedömning 2

Syfte

Det övergripande syftet med projektet "Långvarig sjukskrivning" är att studera socialförsäkringssystemets effekter för individ och samhälle. Riksförsäkringsverkets rapport har tagits fram inom ramen för detta projekt. Det övergripande syftet med Riksförsäkringsverkets utvärdering har varit att följa upp de åtgärder som försäkringskassan och andra rehabiliteringsaktörer vidtagit för att begränsa sjukfallens längd.

Utvärderingen består av två delar. Den första delen omfattar en studie av rehabiliteringsarbetet på försäkringskassan under perioden 1986–1991. Studien omfattar individer med långvariga sjukskrivningar omfattande minst 60 dagar.

I denna del syftar utvärderingen till att besvara två frågeställningar:

- * Hur ser den grupp som får rehabilitering ut?
- * Vilka faktorer ökar sannolikheten för att påbörja rehabilitering?

Utvärderingens andra del syftar till att göra en jämförelse av rehabiliteringsverksamheten mellan perioderna 1979–1985 och 1986–1991. Studien omfattar individer med långvariga sjukskrivningar omfattande minst 90 dagar.

Metod

För att besvara de två frågeställningarna inom ramen för utvärderingens första del har olika metodansatser använts. För att beskriva den grupp som får rehabilitering har en s.k. rehabiliteringsskala utvecklats. Med rehabiliteringsskalan studeras en faktor eller aspekt i taget som t.ex. kön. Rehabiliteringsskalan visar i detta fall om det finns en statistiskt säkerställd skillnad i förekomsten av rehabilitering mellan könen. Skalan tar däremot inte hänsyn till att faktorer som har betydelse för rehabiliteringen hänger samman på olika sätt.

För att förklara varför vissa grupper får rehabilitering har s.k. logistisk regression använts. Metoden tar hänsyn till flera faktorer samtidigt och urskiljer varje faktors egen betydelse för sannolikheten att få rehabilitering.

Jämförelsen av rehabiliteringsverksamheten mellan perioderna 1979-1985 och 1986-1991 är i huvudsak av deskriptiv karaktär.

Bedömning

Som framgår av instruktionen skall min bedömning göras utifrån vad jag bedömer vara god praxis inom området. Då rådande "praxis" enligt min uppfattning inte är av god kvalitet kommer mina bedömningar att vara baserade på vad jag anser bör vara godtagbar kvalitet på en utvärdering.

En annan utgångspunkt för min bedömning är rapportens målgrupp. Ett viktigt mål för rapportförfattarna har varit att utforma rapporten för författarnas viktigaste målgrupp, dvs. de anställda på försäkringskassan.

A Dokumentation

Kraven på en utvärdering, med avseende på läsarens möjligheter att kunna bedöma hur tillförlitlig utvärderingen är, beror bl.a. på vem som är mottagare – dvs. vem som förväntas läsa rapporten. För mitt eget vidkommande anser jag att rapporten skulle varit utförligare i vis-sa avseenden, men i stort är mina krav på en utvärdering tillgodosedda. Jag tror däremot att om "all personal som arbetar på försäkringskassorna" skall kunna tillgodogöra sig rapportens innehåll och kunna bedöma resultatens tillförlitlighet, måste rapporten vara mer utförlig i sin metodbeskrivning.

De brister som i mitt tycke föreligger gäller främst beskrivningen av hur urvalet av individer och bortfallet i undersökningen hanterats, men också hur de använda metoderna beskrivits. Mina kommentarer görs under resp. rubrik.

B Data

Data har samlats in via en intervjuundersökning, en enkätundersökning på försäkringskassorna och en inhämtning av uppgifter från Riksförsäkringsverkets register. Data har samlats in vid två tillfällen – 1986 resp. 1991–1992 – där samtliga tre insamlingsmetoder använts. Vid båda tillfällena gjordes speciella urval av långtidssjukskrivna. Vid den senare insamlingsfasen som omfattade ca 5 000 individer, skedde en uppdelning i två urval: ett urval med 2 000 individer som valts ut slumpmässigt i befolkningen och ett urval med 3 000 individer där samtliga varit sjukskrivna minst 60 dagar i följd under perioden 1986–1989.

I denna studie har huvudsakligen data använts som hämtats in från försäkringskassorna för de ca 3 000 personer som hade varit sjukskrivna minst 60 dagar i följd. Bortfallet uppgår till 211 personer (ca 7 procent).

Det framgår inte klart om urvalet avseende de 3 000 individerna är ett slumpmässigt urval. Bortfallet kan förefalla begränsat, men borde kommenterats. Jag saknar bl.a. en tydlig redovisning av hur urvalet av individer har genomförts och hur svarsfrekvensen och bortfallet fördelas på olika bakgrundsvariabler (t.ex. kön och ålder). Det framgår inte av rapporten om en bortfallsanalys genomförts. Bortfallet kommenteras över huvud taget inte i rapporten. Vill det sig illa kan dessa – i och för sig till antalet få – individer avvika från de övriga på ett systematiskt sätt. Därmed minskar möjligheterna att kunna dra generella slutsatser från insamlad data.

Data är baserade på antal sjukfall, inte individer. Vad innebär detta för t.ex. tolkningen av resultaten? Jag skulle gärna sett en kortare diskussion kring denna fråga i rapporten.

Studien omfattar även en jämförelse av rehabiliteringsverksamheten mellan perioderna 1979–1985 och 1986–1991. Denna jämförelse avser långvariga sjukskrivningar omfattande minst 90 dagar. Det finns ett antal frågor som jag önskat att författarna kommenterat som bl.a.; Har det någon betydelse för jämförbarheten mellan de två tidsperioderna att urvalet av sjukfall för den senare tidsperioden är baserat på sjukfall omfattande minst 60 dagar? Är valet av jämförelseperioder okomplicerat (t.ex. deras olika längd)?

Data framstår som relevanta med tanke på syftet med utvärderingen. Däremot finns en viss osäkerhet om hur urvalet är genomfört samt hur det okommenterade bortfallet är hanterat varför det är svårt att bedöma datas tillförlitlighet. En redovisning av urval och bortfall samt en diskussion om datas kvalitet kunde med fördel lagts som bilaga till rapporten.

C Metod

Som framgått av ovanstående resonemang har två olika metoder använts; en för att beskriva den grupp som får rehabilitering och en för att förklara varför vissa grupper får rehabilitering. Metoderna beskrivs kortfattat i en bilaga till rapporten. De två redovisningarna omfattar sju olika faktorer: kön, ålder, sjukskrivningsdiagnos, sjukfallslängd, sysselsättning vid sjukskrivning, förekomst av arbetsskadeanmälan samt förekomst av ansökan/prövning om förtidspension/sjukbidrag.

En s.k. rehabiliteringsskala har konstruerats för att möjliggöra en beskrivning av den grupp individer som får rehabilitering. Ambitionen har varit att på ett överskådligt sätt visa vilka grupper som dominerar i rehabiliteringsarbetet. De underliggande frågeställningarna var flera: Vad skiljer de individer som får rehabilitering från de som står utanför rehabiliteringsverksamheten? Hur väljs de sjukskrivna till rehabilitering? Kan de sjukskrivna indelas i två grupper där den ena bedöms ha en låg rehabiliteringspotential och den andra en hög?

Skalan beskriver graden av signifikans. Metoden baseras på att testa om den speciella gruppens rehabiliteringsandel signifikant avviker från den totala rehabiliteringsandelen. Metoden baseras på s.k. t-test. Rehabiliteringsskalan har indelats i nio nivåer med lägsta nivån – rehabilitering förekommer i ringa grad – och med högsta nivån – rehabilitering förekommer i stor utsträckning.

Sannolikheten av att komma i åtnjutande av rehabiliteringsåtgärder hanteras med hjälp av s.k. logistisk regression där hänsyn tas till alla faktorer samtidigt. Som ett resultat av analysen fås vilka faktorer som har störst betydelse för att påbörja en rehabilitering. Resultatet av den logistiska regressionen har omvandlats till s.k. oddskvoter för olika grupper av individer. Oddskvoten för olika grupper jämförs med en jämförelsegrupp vars oddskvot är 1,0. Oddskvoter överstigande 1,0

innebär att sannolikheten för att börja en rehabilitering är större än för jämförelsegruppen – och tvärtom då oddskvoten understiger 1,0.

De två metoder som använts förefaller relevanta med hänsyn till de frågor som författarna önskat besvara. Jag saknar dock i beskrivningen av hur en rehabiliteringsskala konstrueras bl.a. hur klassificeringen av skalvärden har genomförts.

Metoderna förfaller riktigt tillämpade. Det är dock inte möjligt att mer i detalj studera om så är fallet då rapporten inte på ett tydligt sätt redovisar vilka metoder som använts och hur dessa använts. Jag hade önskat en kortfattad, pedagogisk, beskrivning av metoderna och deras möjligheter och begränsning – gärna i bilaga – alternativt referenser till relevant litteratur.

Rapporten innehåller en omfattande tabellbilaga med intressanta iakttagelser. Resultaten av den logistiska regressionen (som redovisas i tabell 6 och 7) kommenteras i termer av statistisk tillförlitlighet. De resultat som redovisas i övriga tabeller kommenteras inte med avseende på eventuella statistiskt säkerställda skillnader. Dessa resultat förefaller endast vara av deskriptiv natur (vilket även verkar gälla för jämförelsen mellan de båda tidsperioderna). Som läsare hade jag önskat en ökad tydlighet i detta avseende.

Om målgruppen för rapporten enligt ovanstående resonemang tas för given, dvs. de anställda på försäkringskassan, måste det pedagogiska inslaget i rapporten med avseende på val av metoder ökas dramatiskt. De kortfattade metodbeskrivningarna i bilaga möjliggör inte för en lekman att tillgodogöra sig hur resultaten är framtagna, snarare tilltar förvirringen.

D Slutsatserna

Slutsatserna är i sig väl underbyggda. Däremot är slutsatserna i mitt tycke väl försiktiga och i huvudsak av skisserande natur. Dessutom är rapporten mycket omfattande, vilket gör att läsarens saknad av ett kapitel med sammanfattande slutsatser är påtaglig.

Vad jag saknar är ett avslutande kapitel där dels slutsatserna presenteras på ett sammanhängande sätt, dels där ett antal "varför" besvaras. Dessutom önskar jag att författarna hade tolkat sina resultat i termer av handlingsimplikationer/förslag till prioriteringar: Vilka åtgärder fö-

reslår t.ex. författarna för att förkorta de påtagliga väntetiderna för att komma i åtnjutande av försäkringskassans insatser? Å andra sidan saknas ett antal grundläggande data (läs kostnader och intäkter) för att det skall vara meningsfullt att dra några slutsatser med avseende på handlingsimplikationer (se nedan).

E Kostnadseffektivitet

Som framgått av mina ovanstående kommentarer är rapporten i sig väl genomförd. Resultaten är också av stort intresse. Däremot är det enligt min mening en påtaglig brist att begrepp som kostnader och intäkter/nyttor helt lyser med sin frånvaro. Rapportens värde som beslutsunderlag skulle tillta dramatiskt om kostnaderna för olika delar av den diskuterade verksamheten redovisats. Vilka är de totala kostnaderna för försäkringskassans arbete med rehabilitering? Vilka övriga statsfinansiella kostnader aktualiseras i samband med rehabiliteringsarbetet (t.ex. för förtidspension och sjukbidrag)? Vilka är samhällets kostnader för rehabiliteringsverksamheten? Hur fördelar sig försäkringskassans totala kostnader för rehabiliteringsarbetet på olika åtgärder, på olika diagnosgrupper, på olika led i rehabiliteringsprocessen etc.? Vad är marginalkostnaden för t.ex. olika rehabiliteringsåtgärder? Dessa frågor är bara några exempel på frågor som jag tycker studien hade vunnit på att besvara.

Dessutom vore det av intresse att belysa intäkterna av olika åtgärder och relatera dessa till åtgärdernas kostnader. I ovanstående resonemang nämndes sju faktorer som använts i studien. Ytterligare en faktor som hade varit intressant att ta med är en inkomstfaktor, t.ex. personens bruttoinkomst, vilken i så fall hade kunnat ligga till grund för en indikation på nyttan av t.ex. en viss rehabiliteringsåtgärd.

Under förutsättning att ett antal förhållanden i rehabiliteringsarbetet behöver förändras vore det av intresse att för olika handlingsalternativ beskriva ekonomiska konsekvenser för försäkringskassan, för staten i övrigt och för samhällsekonomin. Därmed hade förutsättningar skapats att relatera marginalkostnaden för en viss åtgärd till åtgärdens marginalintäkt eller marginalnytta.

Som framgår av rapporten (s. 50) är väntetiden inför de olika stegen i rehabiliteringsprocessen för långa sjukfall påfallande lång. För samt-

liga sjukfall är t.ex. väntetiden 73 dagar (medianvärde) från sjukfallets början till att en impuls ges till att utreda om att rehabiliteringsbehov föreligger. Därmed vore det intressant att studera vad en förkortad väntetid med t.ex. en (1) dag givit för resultat i form av lägre kostnader (för försäkringskassan och staten och i övrigt) och välfärdsvinster för samhället (minskade indirekta kostnader för skattefinansiering, eventuellt ökat produktionstillskott samt nytta för individer av att snabbare komma i åtnjutande av rehabiliteringsinsatser).

Ur ett samhällsekonomiskt perspektiv hade det även varit intressant om olika alternativ till rehabilitering diskuterats och jämförts. Bakgrundsfaktorn "sysselsättning" är indelad i kategorierna förvärvsarbetande, arbetslösa och annan sysselsättning. Det hade även varit värdefullt för ett samhällsekonomiskt resonemang om resultaten särredovisats för dessa tre sysselsättningskategorier. Jag saknar även en redovisning av hur framgångsrika olika rehabiliteringsåtgärder har varit.

Enligt min mening hade det sammanfattningsvis varit möjligt att göra en väsentligt bättre utvärdering om verksamhetens kostnader och intäkter/nyttor hade beaktats. Det hade enligt min bedömning varit möjligt utan någon större tillkommande arbetsinsats. Den tillkommande arbetsinsatsen för att fastställa olika åtgärders intäkter (alternativt marginalintäkter) är beroende av om denna typ av information lätt finns att tillgå eller inte.

F Objektivitet

Frågan om rapportförfattarnas objektivitet är mot bakgrund av ovanstående resonemang okomplicerad. Objektiviteten hos rapportförfattarna framstår som godtagbar.

Sammanfattande bedömning

Rapporten visar på ett antal intressanta förhållanden som rör försäkringskassans arbete med rehabilitering av de s.k. långa sjukfallen. Den är fylld med flera i sig iögonenfallande resultat. Däremot diskuteras resultaten knapphändigt och de är inte presenterade på ett sammanhängande sätt. Rapporten är omfattande och de slutsatser som presen-

teraser ger upphov till ett antal funderingar om varför vissa förhållanden verkar råda. Läsaren lämnas dock rådvill i fråga om hur resultaten skall kunna ligga till grund för att åstadkomma förbättringar i rehabiliteringsarbetet. Min huvudsakliga kritik mot rapporten avser dock inte det som behandlats, utan snarare de dimensioner som rapporten saknar.

14 Utvärdering RAR-systemet

Bedömning 1

3 Utvärderingens syfte

Utvärdera ADB-systemet RAR (*Rationell Anmälan Rutin*) vars huvudsakliga avsikter är att förenkla och effektivisera arbetet vid brottsanmälan och brottsutredning samt kunna erbjuda allmänheten en bättre service.

Utvärderingen skall bedöma funktionaliteten i applikationen: Om denna (1) motsvarar aktuella krav, ((2) ger förväntade resultat.

Utvärderingen skall utformas så att den ger en grund för ställningstagande om RAR-systemet skall spridas.

4 Utvärderingens metod

Utvärderingen redovisar följande arbetssätt vid genomförandet av utvärderingen:

1. Översättning av RAR-systemets kravspecifikation till *mål och delmål för systemet*.
2. Probleminventering av tidigare rutiner.
3. Studiebesök för att få kunskap om RAR i drift.
4. Intervjuer av olika "stakeholders": (1) fackliga representanter, (2) projektmedlemmar, (3) användare, (4) professionella ansva-

riga som berörts av RAR som ADB-folk, arkivarien, ansvariga för andra system etc.

5. Användarenkät, 30 frågor varav fr 1–5 *bakgrundsfrågor*, fr 6–19, fr 28 och fr 30 *relaterar RAR till det gamla manuella systemet*, fr 20 och fr 23 berör *kompetensutveckling och utbildning* samt fr 21–22, fr 24–27 och fr 29 berör *funktionaliteten i systemet*.
6. Rapportering kring följande rubriker där källorna för information ovan används med tyngdpunkt på enkätsvaren: Utrustning, Driftsmiljö – säkerhet och sårbarhet, Organisation och arbetsrutiner, Utbildning och dokumentation, Gallring och arkivering.
7. Utvärderingens resultat har rapporterats under följande rubriker: *Hur har givna mål uppfyllts, Erfarenheter från systeminförandet, Önskemål angående uppbyggnad av RAR, Analys och utvärdering, Förslag till fortsatt arbete*.

5 Bedömning av utvärderingen

Som avgörande problem för funktionaliteten av RAR-systemet framstår:

- I Dimensioneringen av datorn och
- II Samordningen mellan RAR-systemet och andra system samt kommunikationen mellan olika system.

När det gäller datorn så konstateras på s 26 att "Underdimensioneringen av datorn har medfört begränsningar på de flesta områden ..." och man drar i förslagen till fortsatt arbete upp frågan igen (s 27) "Inför en kommande spridning av systemet är det oerhört viktigt att man dimensionerar utrustningen på ett riktigt sätt redan från början".

Bristen på samordning med andra system tas upp på s 20. "Något som poängterats både i intervjuer och vid enkäten är bristen på samordning mellan RAR-systemets K-diarium och A- och B-diarierna. På s 26 sammanfattas kommunikationerna med andra system: "Kommunikationen mot framför allt DAFA när det gäller automatisk inhämtning av adressuppgifter till anmälan fungerar mycket dåligt och måste snarast ses över".

Som framgår tydligt av ovanstående citat har grundförutsättningarna för en rättvisande utvärdering av systemets funktionalitet varit mycket dåliga. Detta faktum genomsyrar alla svar och det är svårt att särskilja vilka delmål som egentligen uppfyllts. Ex. vad betyder svaren på följande fråga 25: *Frånsett att systemet varit trögt en längre tid – Hur har tillgängligheten varit till systemet?* I rapportens sammanfattning hänvisas till att tidigare utvärderingar gjorts av systemet som berört metoder och projektarbete men att denna utvärdering tar upp användarnas upplevelser av systemet och hur uppsatta målsättningar med systemet uppfyllts. Man saknar i denna utvärdering helt återkoppling till dessa studier och till hur man kunnat planera ett ADB-system som nästan från början visat sig "seg".

Systemets "tröghet" i användningen med långa svarstider är uppenbarligen ett överskuggande problem, som rycker undan förutsättningarna för utvärderingen. Utvärderarnas slutsats att problemet helt beror på underdimensionering är emellertid knappast styrkt, då det även kan röra sig om brister i systemlösning och programvara.

A Dokumentation

Tidigare manuella anmälningsrutiner samt s.k. A-, B- och K-diarierutiner (vilka är viktiga för att bedöma RAR-systemet) är ej beskrivna alls i rapporten eller angivna genom referenser. En probleminventering avseende den tidigare rutinen är utförd men ej redovisad. Man frågar sig om det är vid värderingen av de gamla rutinerna som felbedömningar har gjorts och som lett till underdimensioneringen av det nya systemet.

Rapporteringen av resultaten från de olika informationskällor som använts är bara tydlig när det gäller enkäten, som i sin tur har mycket stora brister som beslutsunderlag (se nedan under B) Data). I sammanfattningen sägs att användarna har fått möjligheter att utifrån sin arbetssituation beskriva hur man upplever verksamhetsstödet och vilka brister man upplever. Läsaren kan dock inte bilda sig en uppfattning om hur representativ den bild är som beskrivs i utvärderingens resultat, dels för att enkäten har brister i urvalsförfarandet (se nedan under (B) Data) och dels för att andra källor än enkäten är mycket otydligt refererade och dokumenterade. På sid 23 konstateras att möjligheten

att kunna besvara vissa frågor i enkäten har varit beroende på hur man arbetar med systemet. Av denna anledning borde man då ställt frågan om vilken användargrupp man tillhör och genomgående i dokumentationen gjort tydligt reaktionerna på systemet bland de olika användargrupperna och professionella som berörts av utvärderingen. Det saknas också helt en koppling mellan olika användargrupper och de s.k. informationsdelssystemen (Ids) som är byggstenarna i RAR-systemet.

Man har heller inte kontrollerat hur personer med olika utbildning i systemet har reagerat utan bara nöjt sig med att fråga om de tycker att de fått tillräcklig utbildning. Man redovisar på flera håll i rapporten, ex sid 24, att utbildningen fått högt betyg. Förmodligen skulle det ha varit mycket givande för kunskapsbildningen om systemet och dess genomförande om erfarenheterna från denna utbildning följts upp, systematiserats och redovisats uppdelat på olika användarkategorier och efter var de olika problemen och hindren för funktionaliteten och för god service kunnat lokaliseras.

B Data

Utvärderingen avgränsas till Huddinge polisdistrikt där RAR-systemet togs i drift oktober 1991 och ca 260 personer utbildats (sid 9), vilka under det första året arbetat vid ca 45 terminaler och registrerat ca 25.000 anmälningar.

Data har insamlats genom en skriftlig enkät under november 1992, ett år efter starten av systemet. Man anger att enkäten var riktad till samtliga användarkategorier inom polismyndigheten i Huddinge, dock utan att man ställt en fråga om vilken användarkategori som personer tillhör. Den besvarades av 84 personer (62 polisanmälda, 22 civilanställda), men 5 st kasserades på grund av bristfälligt innehåll. Det finns ingen uppgift om sammansättningen av hela populationen, hur de 84 personerna valts ut och om de varit anonyma. Viss bakgrundsinformation ges för den utvalda gruppen men den ställs inte mot motsvarande information i hela populationen.

Till bristerna i representativiteten i urvalet kommer problem med svårbesvarade frågor vilket yttrar sig i att ett stort antal frågor har höga andelar som svarat alt 4: Ej Svarat. Det gäller särskilt svaren på

frågorna som jämför de gamla och nya systemen där det % bortfall varit så höga som 59,5 procent för fråga 13 (om underlag till åklagare), fråga 18 (om ökad rättssäkerhet) hade 46,4 procent och fråga 17 (om mer flexibel polisorganisation) 45,2 procent, för att nämna några av de frågorna som haft högst bortfall. Utredaren kommenterar detta förhållande på sid 23 på följande sätt: "Möjligheten att kunna besvara dessa frågor (om delmål) har varit beroende av hur man arbetar med systemet. Detta innebär att svarsfrekvensen på dessa frågor varierar en del." Man måste konstatera att denna enkät inte uppfyller de elementäraste krav på enkätdata, entydiga frågor ställda så att alla användarkategorier har ett svarsalternativ att välja som även inkluderar "Ej aktuellt" eller "Vet inte", om frågan inte går att ta ställning till.

C Metod

Metoden som redovisats ovan under punkt 4. är närmast ett arbetssätt som till sin struktur innehåller önskvärda viktiga steg i en utvärdering, i detta fall en efterhandsgranskning av hur det nya ADB-systemet fungerat jämfört med det gamla manuella. Greppet man valt är att i enkätfrågor direkt fråga de inblandade användarna om de tycker att målen uppfyllts och att också ge utrymme för att fritt framföra synpunkter och förslag. Man har då inte frågat om personens erfarenhet av det gamla systemet och inte heller, som tidigare påpekats flera gånger, vilken användarkategori personen tillhör i det nya systemet. Man har säkert fångat in en hel del både positiva och negativa synpunkter men problemet med metoden är att det inte går att avgöra problemens tyngd, hur stor andel av användarna berörs, var de finns i organisationen och hur de stora partiella bortfallen skall tolkas.

D Slutsatserna

Utvärderingen sammanfattar i avsnitt 11 sid 26 utvärderingen: "Helhetsintrycket är att RAR-systemet i Huddinge av användarna upplevs som ett mycket bra verksamhetsstöd och att funktionaliteten i stort motsvarar de krav man ställer på en rutin för registrering av brottsanmälan". Som framgår av de ovan redovisade bristerna i under-

laget är det inte möjligt att instämma i dessa slutsatser och anse att de är väl underbyggda i analysen. Att man skall ha ett ADB-system för administrativa handläggningsrutiner i samband med brottsanmälan är en självklarhet i dagens IT-samhälle men att det skall utformas enligt RAR-systemet är en helt obesvarad fråga. Det som pekats ut som positiva synpunkter är att det är enklare att skriva anmälan, man vinner tid och att det är bra att söka i och att ärendena är lätt identifierbara, allt faktorer som naturligen ett ADB-system hanterar på ett bra sätt jämfört med manuell hantering. Baksidan av RAR-systemet, den underdimensionerade utrustningen, dålig integrering med andra system, ingen minskad arbetsbelastning på diarieexpeditionen talar ett annat språk och man frågar sig om inte dessa rön kan tolkas så att RAR-systemet blev dyrt och ineffektivt.

E Kostnadseffektivitet

Både planen för utvärderingen och genomförandet har stora brister som visar på att konsultinsatserna inte använts för insatser som säkerställt kvalitet i utvärderingen.

F Objektivitet

Arbets sättet i utvärderingen är systematiskt och har bredd i den meningen att alla användarkategorier ges utrymme att framföra synpunkter och förslag. Brister i genomförandet, negativa synpunkter etc redovisas, men frågan är om de beaktas nog i slutsatserna. Bilden av funktionaliteten i ADB-systemet blir ändå snedvriden på grund av att utredningen inte använder en representativ metod och att man heller inte speglar vad de olika användarkategorierna har för erfarenheter och värderingar av RAR-systemet. I sammanfattningarna under punkt 9 och 11 drar man alla över en kam och utmålar helhetsintrycket som mycket bra (se ovan under D), vilket – med tanke på bristerna i underlaget – snarast måste uppfattas som en subjektiv bedömning än en objektiv, sakligt grundad slutsats.

6 Sammanfattande bedömning

Det här är en utvärdering av ett ADB-system där tekniklösningen inte håller måttet. Förutsättningarna för effektivt lokalt ADB-stöd och snabb kommunikation mellan olika system har ju dessutom ändrats radikalt under 90-talet med PC-utvecklingen och nya IT-lösningar vilket också bidrar till att systemlösningar snabbt blir föråldrade. Detta torde vara en vanlig situation på många myndigheter och det är därför en angelägen fråga att diskutera hur ADB-system utvärderas i en snabbt föränderlig teknikvärld.

Direktiven till utvärderingen anger att syftet i första hand är att bedöma funktionaliteten i applikationen och om den motsvarar nu aktuella krav och ger förväntade resultat. Det centrala utvärderingsproblemet är svårpreciserat, då jämförelsen mellan det gamla manuella systemet och ADB-systemet, den jämförelse som den gjorda utvärderingen koncentreras till, inte kan ge hela sanningen om RAR-systemet verkligen är bra som ADB-system sett ur användarens perspektiv. Direktiven hjälper inte i detta avseende till med att spåra in utvärderingen till att hitta en effektiv och optimal ADB-lösning och inte heller med hur man skall kunna hantera att kraven på ADB-system och förväntade resultat successivt höjs när personalen blir mera erfaren och systemen mera användarvänliga.

I utvärderingens förslag till fortsatt arbete trycker man på att utrustningen måste vara rätt dimensionerad från början, då användarna både nu och ännu mer i framtiden kommer att få allt högre krav på tillgänglighet och svarstider. Frågan blir då vad som är effektivt i RAR-systemet i en sådan ny tekniskt gynnsam miljö och vad som då utgör alternativa lokala ADB-systemlösningar. Den här gjorda utvärderingen kan ge viss vägledning om vad som upplevs som problem i den nya hanteringen av brottsanmälan jämfört med gamla manuella rutiner men den ger inget egentligt underlag för *bedömning om själva ADB-systemet är optimalt* sett ur ett användarperspektiv med höga krav på att systemet hänger med i teknikutvecklingen.

Bedömning 2

Rikspolisstyrelsen (RPS) har utvärderat det s k RAR-systemet vid Huddinge polismyndighet. RAR (Rationell Anmälans Rutin) är ett lokalt ADB-stöd för brottsanmälan.

Det övergripande målen med systemet är:

- att förenkla och effektivisera polismyndighetens administrativa handläggningsrutiner i samband med brottsanmälan
- att förbättra servicen till allmänheten

Tidigare hade man vid Huddinge polisdistrikt använt en "manuell" brottsanmälningsrutin.

Utvärderingens syfte

Syftet med utvärderingen var att "den skall ligga till grund för ett ställningstagande huruvida RAR-systemet skall spridas till övriga polismyndigheter i nuvarande eller förändrad form".

Utvärderingens metod

Den metod som användes var, dels en skriftlig enkät som besvarades av 84 av de 260 personer på Huddingepolisen som utbildats på systemet, dels ett antal intervjuer med såväl användare som driftsansvariga.

I den skriftliga enkäten fick Huddingepoliserna svara på frågor som mestadels handlade om huruvida RAR-systemet, i olika avseenden, var bättre än de tidigare använda manuella rutinerna.

Dokumentation

Utvärderingen är enkel att följa och relativt väl dokumenterad. Några grova slarvfel(?) stör dock detta intryck:

- Det framgår inte om man avsett att göra en totalundersökning eller ej (se vidare under "Metod").
- Det framgår inte om enkäterna besvarats anonymt.
- Orsakerna till det "interna bortfallet" redovisas otydligt (se vidare under "Data").

Data

För att kunna avgöra om RAR är en bra investering måste både kostnads- och intäktssidan utredas bättre. I utvärderingen sägs i ett avsnitt som handlar om hur RAR administreras att "belastningen när det gäller själva RAR-systemet och dess drift beräknas till ca 1,5 tjänster". Man redovisar också hur många dagar som gått åt till att utbilda dem som ska använda systemet.

Dessa synnerligen oprecisa kostnadsmått måste sedan ställas mot ännu mer oprecisa intäktsmått. I användarenkäten preciseras tex inte användarnas tidsvinster. På påståendet "Stor minskning av administrativt arbete" får användarna välja mellan svaren "Stämmer", "Stämmer delvis" och "Stämmer inte alls".

Det borde rimligtvis vara en central fråga om användarnas samlade tidsvinst är större än de 1,5 tjänster driften av systemet tar i anspråk. Detta hade man enkelt kunnat få svar på med mer precisera frågor i enkäten.

En stor osäkerhet vid bedömningen av enkätsvaren är att det "interna bortfallet" är högt för vissa frågor. Vissa frågor har bara besvarats av ungefär hälften av de som deltagit i undersökningen. Orsakerna till det höga interna bortfallet redovisas bara sammanfattande: "I övrigt ställdes i enkäten frågor runt ett antal delmål. Möjligheten att kunna besvara dessa frågor har varit beroende av hur man arbetar med systemet. Detta har inneburit att svarsfrekvensen på dessa frågor varierar en del".

Ändå är just orsaken till bortfallet helt avgörande för om man överhuvudtaget kan dra några slutsatser av de svar som föreligger. Om det höga interna bortfallet verkligen beror på att vissa respondenter inte berörs av frågan är bortfallet mindre allvarligt. Om däremot respondenterna är berörda av frågan men, av någon anledning, inte vill ut-

trycka sin åsikt, är ett internt bortfall på 50 % förödande för möjligheterna att dra slutsatser av svaren.

Därför är det viktigt att, fråga för fråga, redovisa orsakerna till bortfallet. Ett enkelt sätt att åtgärda problemet hade varit att till varje fråga tillfoga alternativet "berörs ej av denna fråga".

Metod

Att via en skriftlig enkät fråga användarna vilka förändringar som skett sedan systemet infördes är en enkel och billig metod. Att man dessutom kompletterade enkäten med "djupintervjuer" med, dels några användare, dels några "experter" förefaller också vettigt.

Tyvärr är det otydligt huruvida man gjort en misslyckad totalundersökning eller en lyckad urvalsundersökning.

Av utvärderingen framgår att ca 260 personer vid Huddinge polismyndighet utbildats på RAR-systemet och att de flesta av dessa arbetar med systemet. Vidare framgår att 84 personer besvarat enkäten, men att 5 svar kasserats pga bristfälligt innehåll. Några ytterligare ledtrådar går ej att finna i utvärderingen.

Detta kan tolkas på flera sätt. Jag väljer här två extremfall som leder till olika slutsatser om utvärderingens tillförlitlighet.

1. Har man gjort en totalundersökning, dvs bett ca 260 personer besvara enkäten, men bara fått 79 användbara svar? En svarsfrekvens på 30 % är givetvis alldeles för lågt för att man ska kunna dra några slutsatser av resultatet.
2. Har man gjort en urvalsundersökning som riktat sig till 84 av ca 260 möjliga personer med ett externt bortfall på 6 % (de fem oanvändbara svaren)? Isåfall är svarsfrekvensen mycket tillfredsställande. Däremot kan man invända att ett urval på 84 av 260 ligger på gränsen till det acceptabla. Med en så liten population kan det vara mödan värt att göra en totalundersökning. Om inte annat för att slippa alla invändningar kring urvalsmetod och om storleken på urvalet är tillräckligt för att säkra slutsatser ska kunna dras.

Tidpunkten för utvärderingen är en annan osäker faktor. Systemet installerades i oktober 1991. Utvärderingen gjordes hösten 1992. Så långt är allt OK. Problemet är att när utvärderingen gjordes hade de datorer man använde på Huddinge polisdistrikt inte tillräcklig kapacitet för att RAR-systemet skulle fungera som det var avsett. Väntetiderna vid utskrifter och sökningar blev ibland oerhört långa. I februari 1993 installerades starkare datorer.

Man kan se på detta på två sätt. Å ena sidan hade det varit bättre att vänta med utvärderingen tills tillräckligt starka datorer införskaffats. Den nu gjorde utvärderingen ger ingen korrekt bild av RAR-systemets fulla kapacitet. Å andra sidan hade utvärderingen ett underliggande syfte att ta till vara på erfarenheter ifall systemet skulle spridas till andra distrikt. Det är naturligtvis en mycket viktig erfarenhet att veta hur starka datorer som behövs för att härbärgera systemet.

En annan invändning är att man för att utvärdera ett av RAR-systemets huvudsyften, nämligen att erbjuda allmänheten en bättre service, snarare bör rikta sig direkt till allmänheten än att, som i detta fall, fråga de anställda vid Huddingepolisen hur de tror att allmänheten uppfattar servicen.

Jag har dock förståelse för att man inte gjorde någon enkät till allmänheten. Detta hade kraftigt ökat kostnaderna för utvärderingen. Däremot undrar man om det inte var möjligt att, åtminstone i vissa avseenden, mer direkt mäta hur allmänhetens service påverkats. Av utvärderingen framgår det tex att en del av den förbättrade servicen skulle bestå i att man besvarade allmänhetens frågor snabbare. Här borde en enkel före-efter mätning av faktiska svarstider ha kunnat användas för att klargöra RAR-systemets effekter.

Slutsatserna

Det är svårt att svara på frågan om de slutsatser som dras av utvärderingen är väl underbyggda i analysen. Detta har mycket att göra med de oklarheter jag redovisat ovan.

Om det interna bortfallet endast beror på att de som inte berörs av frågan undviker att svara och om man gjort en urvalsundersökning riktad till 84 av ca 260 användare anser jag att de slutsatser som dras är relativt välbalanserade.

Är det däremot så att det interna bortfallet till stora delar beror på att användarna inte vågar uttrycka sin åsikt och om det hela egentligen är en misslyckad totalundersökning med en svarsfrekvens på 30 % går det inte att dra några slutsatser alls.

Man kan också diskutera urvalet på en högre nivå. Kan verkligen ett enda polisdistrikt, Huddinge, användas som ett mått på hur RAR-systemet kommer att fungera i andra polisdistrikt? Det är tänkbart att just brottsanmälningsrutinen är så pass "standardiserad" att de fördelar och nackdelar som RAR-systemet för med sig är desamma oavsett vilket polisdistrikt det gäller. Men man kan lika gärna tänka sig att det är stor skillnad mellan brottsanmälningsrutinen i ett förmodligen hårt brottsbelastat distrikt som Huddinge och i polisdistrikt ute i glesbygden. Det hade naturligtvis blivit dyrt att pröva systemet vid ett antal distrikt av olika typ. Däremot är det viktigt att beakta distriktens olikheter när man tar ställning till om man ska sprida systemet.

Kostnadseffektivitet

Rikspolisstyrelsen har använt en enkel och billig metod för att utvärdera RAR-systemet. Om man skall hålla sig inom samma kostnadsram skulle man ändå kunnat nå en bra bit längre med lite mer preciserade frågeställningar i enkäten:

- Användarnas tidsvinster borde ha preciserats bättre. Detta är en mycket central fråga vid bedömningen av RAR-systemets lönsamhet.
- Svartalternativet "Berörs ej av denna fråga" borde ha funnits med för att fånga upp orsakerna till det interna bortfallet och därmed möjligheterna att dra säkra slutsatser av svaren.

Begreppet kostnadseffektivitet kan dock även innebära att det kan vara värt att lägga ned lite mer resurser på utvärderingen om den extra information man därigenom får gör att beslutsunderlaget blir mycket bättre. Frågan är om inte en så stor investering som RAR-systemet skulle vara förtjänt av en fylligare utvärdering än den Rikspolisstyrelsen gjort av försöksverksamheten vid Huddinge.

Objektivitet

Utvärderingen ger i sig ett objektivitetsintryck. Avgörande för om jag ska kunna betrakta utvärderingen som objektiv är snarare hur en eventuell spridning av RAR ska finansieras. Detta framgår inte av utvärderingen.

Om RPS ska betala ADB-investeringar från den ordinarie budgeten för polisen finner jag ingen anledning att ifrågasätta objektiviteten. Isåfall drabbar en misslyckad investering myndigheten själv. De blir tvungna att avstå från någon annan investering eller skära ned på löpande utgifter.

Om däremot ABD-investeringarna finansieras med särskilda anslag från regeringen, utanför den ordinarie budgeten för polisen, finns det all anledning att se upp med objektiviteten. Isåfall har Rikspolisstyrelsen inget att förlora på att införa RAR-systemet och därför all anledning att försöka övertyga regeringen om systemets förträfflighet.

Sammanfattande bedömning

Rikspolisstyrelsen har valt en enkel och billig metod för att utvärdera RAR-systemet vid Huddinge polismyndighet. Informationsvärdet i den enkät som användarna fått besvara hade dock kunnat höjas betydligt, bl a med mer preciserade frågor kring hur stora tidsvinster användarna gjort.

Utvärderingen är lättläst och, så långt det går att bedöma, genomförd på ett objektivt sätt. Tyvärr slarvar (?) man i utvärderingen med informationen kring vissa mycket viktiga detaljer om genomförandet, vilket gör det mycket svårt för en utomstående att bedöma möjligheterna att dra säkra slutsatser av resultatet. Det framgår inte om man gjort en totalundersökning med högt bortfall eller en urvalsundersökning med lågt bortfall. Det framgår heller inte om respondenterna besvarat enkäten anonymt. Orsakerna till det höga interna bortfallet redovisas alltför summariskt.

Felaktig huvudfrågeställning?

En fråga som jag inte berört ovan är om Rikspolisstyrelsen valt rätt huvudfrågeställning för utvärderingen. Stod Rikspolisstyrelsens val verkligen mellan att sprida just RAR till övriga polismyndigheter eller att behålla de manuella rutinerna för brottsanmälan?

Ofta är det självklart att "nästan vilket ADB-stöd som helst" är bättre än att hålla fast vid manuella rutiner. Utan att känna till de närmare omständigheterna är det därför föga förvånande att Huddinge-poliser-
na, som var vana med manuella rutiner, upplevde RAR som en förbättring. Den fråga många myndigheter ställs inför i samband med införande av ADB-stöd istället för manuella rutiner brukar snarare vara vilket av ett antal alternativa ADB-system som är den mest effektivitetshöjande investeringen.

Nu går det inte att utifrån utvärderingen avgöra huruvida RPS kanske prövade alternativ till RAR på andra polismyndigheter och utvärderade dessa separat, eller om man i ett tidigare skede, på något sätt, kommit fram till att RAR var det bästa systemet om man överhuvudtaget skulle införa något ADB-stöd. Isåfall har jag inga invändningar mot huvudfrågeställningen.

15 Kvalitetstveckling inom skatteförvaltningen (Riksskatteverket)

Bedömning 1

Utvärderingens syfte

Den utvärdering som beskrivs i rapporten Kirstein, Åndell (1993) sammanställer och analyserar de attityder skattskyldiga företag uttrycker i en enkät. Attityderna gäller framför allt skatteförvaltningens kunnande och service. Syftet är att klassificera komponenter av kunnande och service efter de betyg de skattskyldiga ger och efter hur viktiga de tycker att komponenterna är.

Utvärderingens metod

I undersökningen, utförd 1993 i fem län i Sverige, uttalade sig skattskyldiga företag om hur väl de tycker att skatteförvaltningen fungerar i vissa avseenden. Två populationer avgränsades: skattskyldiga företag som utsatts för revision respektive ej utsatts för revision. Företagen fick på en femgradig skala ange om de instämmer eller inte instämmer med påståenden såsom "personalen är trevlig och hjälpsam" och "skattemyndighetens broschyrer är begripliga och skrivna på ett enkelt språk". Ett 40-tal påståenden gjordes som skattskyldiga fick ta ställning till. Svaren kan uppfattas som den skattskyldiges betyg. Med statistisk teknik fördes liknande påståenden samman i grupper, som i rapporten kallas "dimensioner". Sex dimensioner bildades för ej revi-

derade företag och sju för företag som utsatts för revision. Till dimensionen "sakkunskap" fördes åtta påståenden, bl.a. "personalen är kunnig när det gäller grundläggande skattefrågor". Den grad av betydelse som företagen tillmätte "sakkunskap" skattades med statistisk teknik. På så sätt tilldelades varje dimension dels en grad av betydelse, relativt de andra dimensionerna, dels ett vägt betyg. Resultaten visas effektivt i rapporten med tvådimensionella punktdiagram där grad av relativ betydelse läses av på y-axeln och vägt betyg på x-axeln. I andra kvadranten syns de dimensioner som har stor relativ betydelse för de skattskyldiga företagen och fått jämförelsevis låga betyg av dessa.

Tydligt utfördes undersökningen även 1992, men den gången bara i tre län. Rapporten innehåller jämförelser för de tre länen.

Bedömning av utvärderingen

Dokumentation

Det är tydligt att rapporten Kirstein, Åndell (1993) snarare är avsedd som ett underlag för en muntlig redovisning inför kunden, skatteförvaltningen, än som en rapport som ska läsas av personer som kanske inte alls är insatta i problemställningarna. Kirstein, Åndell (1993) är skriven på ett översiktligt sätt med åtskilliga tabeller och diagram som kan utan ytterligare förstoring kopieras till OH-bilder. Det finns ingen innehållsförteckning och inga referenser. Beskrivningen av använda metoder är knapphändig. Jag vill inte kritisera författarna för detta, eftersom jag inte vet vilket uppdrag de hade av kunden.

Läsaren av Kirstein, Åndell (1993) kan inte skapa sig en uppfattning av utvärderingens tillförlitlighet. Det beror framför allt på att rapporten som sagt är knapphändig, men till någon del även på att vissa moment i förfarandet inte förefaller helt tillfredsställande.

Hur metoderna beskrivs

Något mer än 6000 ej reviderade företag och ca 1600 reviderade företag valdes ut. Urvalsdesign framgår inte. I en fotnot nämns att urvalet

i Östergötland av ej reviderade företag drogs felaktigt, men felet beskrivs inte och inte heller de åtgärder som eventuellt vidtogs för att justera för felet. Varje utvalt företag har fått ett formulär, gissningsvis med post, eftersom bortfallet är högt. Hur formuläret såg ut framgår inte. Bortfallets eventuellt snedvridande effekt studerades i en särskild analys (se vidare nedan). Svaren på formuläret analyserades inledningsvis med faktoranalys och korrelationsanalys. De resultat som redovisas i rapporten togs fram med PLS och regressionsanalys. Inflödet i processen, förutom svaren på formulären, är en i förväg formulerad modell, som tydligen justerades något av den inledande analysen. Beskrivningen av förfarandet är mycket kortfattad.

Data

Det data som används kommer till den helt dominerande delen från de skattskyldigas svar på frågorna, d.v.s. deras värderingar (här kallade betyg). Datas kvalitet och på vilket sätt kvaliteten bedömts berörs inte alls i Kirstein, Åndell (1993). Det finns en egenskap hos betygen som är värd att peka på. Rapporten innehåller tabeller med genomsnittligt betyg för varje påstående uppdelat på län. Den visar att praktiskt taget alla betygsgenomsnitt ligger över 3 (bortsett från fyra påståenden under rubriken acceptans som rör snarare skattesystemet i sig än skatteförvaltningens sätt att utföra sitt uppdrag). Det kan bero på att de flesta skattskyldiga är ljumma eller något positiva till samtliga av de 40 påståenden som görs i formuläret eller på att datainsamlingen inte lyckats fånga nyanser i värderingarna. En femgradig skala användes. Hur den såg ut framgår inte. Det framgår inte heller varför valet föll på just en femgradig skala, inte heller varför formuläret använder positiva påståenden genomgående ("Skattekontoret ringer eller skriver tillbaka på utlovad tid" istället för t.ex. "Ringer eller skriver skattekontoret tillbaka på utlovad tid?"). De flesta av påståendena berör olika aspekter av direktkontakt med skattemyndigheterna. Det finns ett påstående som lyder: "man får en ursäkt när fel begåtts". Däremot finns inga frågor om fel faktiskt begåtts och om de var allvarliga för den skattskyldige.

Är de ljumma till positiva svaren en effekt av formulärens utformning? Kirstein, Åndell (1993) diskuterar inte den risken.

Hur formuleringarna i blanketten testades framgår inte. Ett vanligt fel i kundundersökningar i allmänhet är att en fråga är sammansatt av två eller fler frågor ("Är butiken rymlig och är varorna i montrarna tydligt förevisade?"). Påståendet "Skattemyndighetens broschyrer är begripliga och skrivna på ett enkelt språk" som använts i den här undersökningen innehåller två påståenden. Hur skall en skattskyldig svara som tycker att språket är enkelt men broschyren ändå är obegriplig?

Några av påståendena är relevanta bara för en del skattskyldiga. Det gäller bl.a. "personalen är kunnig när det gäller mer komplicerade arbetsgivaravgift- och källskatteärenden" som enligt en bilaga berör ca tre fjärdedelar av de svarande. Det framgår inte vilket svarsalternativ den resterande fjärdedelen kunde ange på formuläret och hur det svaret hanterades i analysen. Tre fjärdedelar av de ej reviderade företagen som svarat har varit i kontakt med skattemyndigheterna minst en gång de senaste 12 månaderna. Det framgår inte om den övriga fjärdedelen inte varit i kontakt med skattemyndigheterna eller om de inte svarat på den frågan.

Skattskyldiga får inga frågor om vad de tycker är viktigt. Istället skattas graden av betydelse, där samband mellan svar på detaljfrågor och svar på frågor om "total kvalitet" utnyttjas. Det är i princip ett lämpligt förfarande. Frågor om betydelse är svåra att ställa.

Metod

Det är svårt att i detalj förstå hur resultaten i rapporten tagits fram. Här bedöms bara bortfallsanalysen.

Bortfallsanalys

Bortfallet är högt, ca 50 % för ej reviderade företag och 60 % för de reviderade. Det är för all del vanligt att bortfallet är högt i kundundersökningar, men det blir ju inte mindre allvarligt för att även en del andra undersökningar lider av samma svårighet. Det är inte klart ut sagt i Kirstein, Åndell (1993) men hela den statistiska analysen förefaller bygga på ett antagande om att svarsbenägenheten är lika för alla

objekt i urvalet. Den bortfallsanalys som redovisas i rapporten tycks tvärtom visa att svarsbenägenheten faktiskt inte är lika för alla objekt, och dessutom att det finns ett samband mellan svarsbenägenhet och undersökningsvariabler. Enligt rapporten är p. g. a. bortfallet andelen företag med 0–5 anställda bland de svarande bara 68 % jämfört med 86 % i populationen (unionen av reviderade och ej reviderade företag). Motsvarande överrepresentation finns för större företag. Rapporten innehåller även en tabell för ej reviderade företag som återges här i layout-mässigt bearbetat skick. Den visar skattad relativ betydelse för sex dimensioner uppdelat på storleksklasser av företag. Den antyder ett samband mellan antal anställda (som i sin tur som sagt har ett samband med svarsbenägenhet) och det undersökningen syftar till att belysa. Det verkar därmed olämpligt att anta att svarsbenägenheten är lika för alla objekt i urvalet. Kirstein, Åndell (1993) argumenterar för påståendet att svarsbenägenheten skulle vara lika med en liknande tabell, som till skillnad från den här återgivna jämför gruppen företag med 0–5 anställda med gruppen samtliga företag, i vilken senare företag med 0–5 anställda utgör 68 %. I *den* tabellen framträder inte skillnaden mellan små och större företag särskilt tydligt, men det beror ju på att man jämför två grupper som till stor del är identiska.

Antal anställda	Tillgänglighet	Bemötande	Sak-kunskap	Snabbhet/handläggningstid	Kommunikation	Råd, rutiner	Summa
0–5	5	10	26	12	29	18	100
6–50	7	10	31	15	17	20	100
51–	11	14	34	6	16	19	100

Rapporten hänvisar även till en bortfallsuppföljning, som inte visade någon signifikant skillnad mellan dem som svarat i första omgången och dem som svarade i uppföljningen. Detta används som ytterligare ett argument för att de svarande kan betraktas som ett enkelt representativt urval ur populationen. Felet med det resonemanget är att man utan att diskutera styrkan på testet inte kan dra slutsatsen att det inte *föreligger* skillnad mellan bortfall och svarande. Rapporten anger inga som helst uppgifter om hur bortfallsuppföljningen gjordes, än mindre

någon diskussion av styrkan. Slutsatsen att "materialet måste därför betraktas som representativt för företagen i respektive län" är därför tvivelaktig.

Slutsatser

Kirstein, Åndell (1993) presenterar inte slutsatser på ett tydligt sätt. Snarare ges under rubriken "Analysinstruktioner" en kortfattad handledning i hur skatteförvaltningen ska tolka resultaten.

De största källorna till statistisk osäkerhet torde vara mätfel i data (se avsnitt "Data" ovan), bortfallet (se avsnitt "Bortfall" ovan) och urvalet. Det finns ingenstans i rapporten någon diskussion om urvalsosäkerheten. Kirstein, Åndell (1993) ger ingen tillfredsställande vägledning i hur slutsatsernas tillförlitlighet påverkas av dessa osäkerhetskällor.

Kostnadseffektivitet

Det går att göra en väsentligt bättre utvärdering än den som nu gjorts, men det skulle kräva mer resurser. Möjligen skulle det vara fördelaktigt att minska det totala stickprovet och i stället satsa på kvalitetsmässig förbättring av utvärderingens olika delar.

Objektivitet

Det finns inget som tyder på att skatteförvaltningen skulle brista i objektivitet i utvärderingen.

Utvärderingen i sitt sammanhang

Skatteförvaltningen utför andra utvärderingar av hur väl skatteförvaltningen sköter delar av sitt uppdrag. Det finns inga referenser till dessa i Kirstein, Åndell (1993), som behandlar en begränsad del av skatteförvaltningens verksamhet.

En utvärdering bör ingå i ett system av kvalitetsförbättring. Kirstein, Åndell (1993) säger inget om detta system, och gissningsvis ingår det heller inte i uppdraget från kunden. Det finns här några frågor som blir obesvarade, bl.a.:

- Varför är just den här delen av skatteförvaltningens verksamhet vald för utvärdering? Varför inte t.ex. en studie i vilka fel förvaltningen gjort sig skyldig till och hur den typen av fel ska kunna undvikas i framtiden?
- Finns det bred uppslutning bland personalen om syftet med den här utvärderingen? Den berör till stor del personalens arbetssätt.
- Vad ska utvärderingen leda fram till? Kommer t.ex. nedskrivna standarder eller checklistor tas fram för personalens arbetssätt?

Mervärdet i den här typen av analys, jämfört med att bara deskriptivt redovisa skattskyldigas betyg, är att man får en fingervisning om vad som kan vara lämpligt prioritera i kvalitetsarbetet. Underlaget för prioriteringen ges av skattningar av den grad av relativ betydelse som företagen tillmäter vissa kvalitetsdimensioner. Det är relativ betydelse, så det finns ingen information om betydelse i absolut mening, än mindre om en skillnad i relativ betydelse är intressant. En svaghet i många statistiska förfaranden är att de kan påvisa en skillnad men inte säga något om den skillnaden är intressant. Den här typen av utvärdering behöver kompletteras med andra studier som tillsammans vägleder myndigheten i prioriteringen av insatser.

Den här utvärderingen säger strängt taget inget alls om skatteförvaltningen gör ett bra jobb eller inte. Vad säger det vägda medelbetyget 3.3 för "personalen är kunnig när det gäller mer komplicerade ärenden angående den allmänna självdeklarationen för företag"? Får personalens kunnskap godkänt eller är företagen missnöjda med vad de tycker är medelmåttigt kunnande? Nej, utvärderingen försöker inte uttala sig om hur väl skatteförvaltningen utför sitt uppdrag. Den syftar till att stödja prioriteringen av vissa delar av kvalitetsarbetet.

Sammanfattande bedömning av utvärderingen

Det finns mycket intressant i rapporten som skatteförvaltningen säkert har nytta av. T.ex. ger de företag som svarat i genomsnitt högre betyg för skatteförvaltningens talspråk än skriftspråk. Huvudresultaten visas effektivt i diagram som säkert inspirerar till fortsatt förbättring av kvaliteten.

Utvärderingen försöker inte uttala sig om hur väl skatteförvaltningen utför sitt uppdrag. Den syftar till att stödja prioriteringen av vissa delar av kvalitetsarbetet.

Eftersom beskrivningen av de metoder som utvärderingen använt är knapphändig är inte utvärderingen fullt ut tillgänglig för utomstående. Vissa oklarheter i förfarandet gör att utvärderingens trovärdighet naggas i kanten.

Bedömning 2

Beskrivningens syfte

Syftet med utvärderingen framgår inte i rapporten, förutom i titeln. Av själva analysen kan man gissa sig till att syftet är att utvärdera kvalitetsnivån inom ett urval av skatteförvaltningar genom en enkätundersökning riktad mot företag.

Kommentar: Syftet som framgår från analysen överensstämmer inte fullt ut med titeln, vilken leder tankarna till en analys av effekter av åtgärder för att höja kvaliteten på skatterevision av och information till de berörda företagen. En tydlig syftesbeskrivning är förstås alltid viktig och skulle i detta fall kunnat hindra denna tvetydighet.

Utvärderingens metod: Metoden att mäta nivån på kvaliteten utgår från ett antal (6) kriterier som anses utgöra delkomponenter (kvalitetsdimensioner) av begreppet total kvalitet. För att få ett mått på var och en av dessa dimensioner har ett frågeformulär komponerats med totalt 31 delfrågor summerat över de sex kvalitetsdimensionerna, eller delområdena som de benämns. Dessutom har tre frågor ställts under rub-

riken total kvalitet. Frågor har även ställts beträffande inställningen (hos företagen/företagarna) till skattesystemet i allmänhet (om det fyller en viktig samhällsfunktion, om det fungerar bra, är rättvist etc.). Varje fråga besvaras enligt en femgradig skala från lägsta betyg (1) till högsta betyg (5). Det slutliga indexet multipliceras med 20 så dess skala går från 20 till 100. Med utgångspunkt från svaren på enkätfrågorna konstrueras ett index för varje delområde, detta index betraktas sedan som en förklarande variabel och en regressionsanalys utförs där indexet för total kvalitet är den beroende variabeln. Regressionskoefficienterna för de sex delområdesindexerna visar vilka delområden som bidrar starkast till ett ökat total kvalitetsindex. Detta index skall användas för att göra jämförelser över tiden och mellan skatteförvaltningar.

Bedömning

A Dokumentation: Dokumentationen består av enkätformulären, men inga referenser till andra studier som kan styrka riktigheten och relevansen av den valda metoden. Inga andra tidigare utvärderingar hänvisas till. Det är för en utanförstående bedömare omöjligt att sätta denna rapport in i ett sammanhang som motiverar dess utförande (klagomål från kunder, egna upplevda problem, ineffektiva rutiner etc.). Dessutom finns ett avsnitt som kallas "Analysinstruktioner", men det framgår inte om detta är instruktioner givna av uppdragsgivarna, vilket vore den naturliga tolkningen av denna rubrik. Detta är viktigt eftersom det inte är klart om enkätfrågorna konstruerats av uppdragsgivarna eller av utvärderarna, eller eventuellt i samarbete mellan båda.

B Data: De data som används är enkätsvaren dessa dokumenteras riktigt i form av frekvenstabeller och -diagram. Inga sekundära datakällor används. Bortfallsanalyser genomförs och avvikelser i frågerespons i bortfallet diskuteras och analyseras. Det bör utifrån denna analys gå att acceptera slutsatsen att bortfallet inte systematiskt snedvridet resultaten.

C Metod: Utvärderingens metod beskrevs ovan, här ges en del kommentarer beträffande dess relevans.

En kommentar till metoden är att de index som konstrueras verkar bäst lämpade för jämförelser över tiden för en och samma skatteförvaltning om syftet är att mäta effekterna av vidtagna åtgärder för kvalitetsutveckling. Detta gäller speciellt om man frågar samma företag vid två tillfällen. Ingen sådan jämförelse görs dock (men är kanske ämnad för framtiden). Konstruktionen av variabler och index beskrivs kortfattat i en bilaga. En mer detaljerad beskrivning skulle varit av godo för att fullt ut förstå metoden, t.ex. med ett exempel.

Jämförelser mellan olika skatteförvaltningar är motiverad med utgångspunkt från en mätning av nivån på kvaliteten. Ett problem är dock att eftersom man frågar olika företag i olika län så är svaren inte direkt jämförbara. De indexsiffror som presenteras ligger numeriskt mycket nära varandra och man skulle behöva utföra ett statistiskt test av noll-hypotesen om ingen skillnad i kvalitetsnivå mellan förvaltningarna. Det kan hända att vissa förvaltningar haft "otur" i urvalet och fått en högre andel av företag som är missnöjda än vad som gäller för hela populationen. Populationerna av företag kan även skilja sig åt mellan olika delar av landet, t.ex. på grund av olika näringslivsstruktur, vilket i sin tur kan innebära att vissa förvaltningar har svårare uppgifter att lösa i genomsnitt. Ett sådant förhållande kan leda till mer klagomål och lägre upplevd kvalitet, även om den är samma som i andra delar av landet. Standardmetoder för att utföra test av skillnader i medelvärden bör kunna utföras, liksom även test av skillnader i varians i svaren. Det sistnämnda kan vara en viktig kvalitetsaspekt eftersom ett givet genomsnitt kan gömma stora skillnader i upplevd kvalitet mellan respondenterna inom ett och samma urval (län). En del företag kan ha blivit mycket väl bemötta, medan andra mycket dåligt bemötta när de hänvänt sig till sin skatteförvaltning t.ex. En hög kvalitet är givetvis det man bör sträva efter först och främst, men en jämn kvalitet kan indikera att ett program för kvalitetsutveckling trängt i genom (accepterats) av alla (de flesta) medarbetarna. En hög varians i kvaliteten kan indikera att man inte nått fram till alla medarbetarna.

Ett annat metodproblem är att författarna behandlar sitt index som om man kunde betrakta kvalitet som mätt enligt en kardinal skala. Detta innebär att man kan säga att förvaltning A har en dubbel så hög kvalitet än förvaltning B om kvoten mellan deras respektive index är lika med två. I själva verket är skalan ordinal, d.v.s. man kan bara säga att A uppfattas som bättre än B, men inte hur mycket. Dock är

regressionsanalysmetoden rimlig eftersom den bör identifiera de områden som är relativt sett viktigast för den upplevda total kvaliteten (i ordinal mening), och även identifiera "problemområden". I stort sett kan man betrakta den använda metoden som lämplig och även riktigt tillämpad, trots invändningarna ovan.

D Slutsatserna: Inga slutsatser dras, resultaten presenteras i ett stort antal tabeller och fyrfältsdiagram.

E Kostnadseffektivitet: Varken total kostnaderna, eller målet för undersökningen är givna. Eftersom effektivitet definieras som graden av måluppfyllelse dividerat med förbrukade resurser, kan denna fråga inte besvaras. Om man gissar att undersökningens inriktning överensstämmer med det mål som uppdragsgivarna gett utredarna, och gör en bedömning av vilken arbetsinsats som dylika enkäter brukar kräva, så kan kostnadseffektiviteten förmodligen betraktas som hygglig. Den skulle dock vara betydligt högre om presentationen av rapporten förbättrades avsevärt, speciellt vad gäller slutsatserna, vilka alltså inte finns. Det skulle dock kunna åtgärdas genom att begära ytterligare arbete på rapporten i detta avseende (vilket kanske redan skett).

F Objektivitet: Eftersom uppdragets formulering inte är given kan inte heller denna fråga besvaras. Det framgår inte, som påpekats ovan, om enkätfrågorna konstruerats i samarbete med uppdragsgivaren eller inte, eller vilka instruktioner som utredarna fått. Dock verkar frågorna relevanta och objektiva, givet det huvudsyfte som utvärderingen verkar ha.

Sammanfattande omdöme

Utvärderingen är omfattande och synes relevant. Det finns brister i den statistiska analysen, vilket påpekats ovan. Det största bristen är dock i presentationen av rapporten. Den är rörig, utan inledning, syftesformulering, metodpresentation (förutom i en bilaga där den inte bör vara), samt saknar slutsatser. Presenterad som en uppsats på B- eller C-nivån i vilket högskoleämne som helst så skulle dessa brister leda till att den underkännes. Detta är inte enbart formella påpekande utan även en fråga om resursslöseri eftersom de nedlagda resurserna

på undersökningen har lett till resultat som bör vara till stor nytta för skatteförvaltningarna i deras kvalitetsutvecklingsarbete, men den undermåliga redovisningen av resultaten reducerar den potentiella effekten.

- 16 Aid and Macroeconomics:
An Evaluation of Swedish Import
Support to Guinea-Bissau,
Mozambique Nicaragua and
Vietnam
(SIDA)
- 17 Fungibility and Sustainability
– Import Support to Tanzania
(SIDA)
- 18 Swedish Support to Cape Verde
(SIDA)

Bedömning 1

Inledning

En del av svenskt bistånd är inriktat på att genom importstöd häva den restriktion på ekonomisk utveckling som härrör ur brist på utländsk valuta. Importstöd ges antingen i form av ett tillskott av utländsk valuta eller genom att givarlandet direkt förser biståndsmottagaren med produkter som inte kan erhållas på den inhemska marknaden.

Under senare år har denna typ av bistånd ökat, samtidigt som dess inriktning förändrats. Övergripande kan sägas att från att ha varit projektbunden har stödet antagit en mer generell karaktär. Det skiljer sig från det sk betalningsbalansstödet som har en vidare applikation och också kan användas för t ex avbetalningar på mottagarlandets utlandsskuld. För att utröna effekten och måluppfyllelsen av detta stöd har SIDA låtit utföra ett flertal utvärderingar som genomförts av såväl SIDA:s egen personal som externa konsulter.

Expertgruppen för studier i offentlig ekonomi (ESO) har startat ett projekt som avser att belysa tillförlitligheten och objektiviteten i olika myndigheters utvärderingsverksamhet. Föreliggande rapport är en del i detta projekt och syftet är att bedöma relevansen i de utvärderingar som SIDA låtit genomföra vad beträffar importstödet betydelse för mottagarländernas ekonomiska utveckling. Enligt ESO:s direktiv till bedömningen av dessa utvärderingar ska följande aspekter särskilt beaktas:

1. Dokumentation, dvs kan läsaren utifrån de uppgifter som presenteras i utvärderingen bilda sig en uppfattning om dess förlitlighet.
2. Metodens lämplighet.
3. Slutsatser. Är dessa rimliga och följer de av resultaten i utvärderingen.
4. Kostnadseffektivitet. Är resursinsatsen i paritet med värdet av utvärderingen.
5. Objektivitet. Finns det skäl att ifrågasätta utvärderarens oväld.

Bedömningen av SIDA:s utvärdering av importstödet betydelse i Cap Verde, Guinea-Bissau, Mozambique, Nicaragua, Tanzania och Vietnam, kommer att fokuseras kring dessa fem punkter. Vi vill göra läsaren uppmärksam på att i det följande tas inte ställning till biståndets effektivitet i övergripande mening i respektive land, utan rapporten är strikt koncentrerad till utvärderingarnas kvalitet och relevans.

Rapporten är uppbyggd på följande sätt. Inledningsvis ges några allmänna synpunkter som berör samtliga utvärderingsstudier. Därefter går vi igenom utvärderingen för respektive land, med betoning på ovanstående fem punkter. Avslutningsvis sammanfattas slutsatserna angående utvärderingarnas kvalitet, relevans och oväld.

Några övergripande kommentarer

Utvärderingarnas syfte är att utifrån de kriterier som formulerats av SIDA bedöma importstödets effekter på såväl mikro- som makroekonomisk nivå. Under senare år har en allt större tonvikt lagts vid den senare aspekten och i föreliggande utvärderingar är det särskilt stödets effekter på tillväxten, eller tillväxtförutsättningarna, som ska beaktas.

Utvärderingarna tar som utgångspunkt uppfyllelsen av de mål eller föresatser som ställts upp av SIDA. Med en sådan uppläggning framtonar två övergripande, eller principiella, problem. För det första, när beställaren av utvärderingen sammanfaller med den myndighet som ska utvärderas, ställs naturligtvis extra stora krav på klarsynthet och objektivitet. Av naturliga skäl kan beställaren vara "hemmablind", men det är dessutom lättare för utomstående att hävda att dessa dubbla roller underminerar trovärdigheten för utvärderingen. För det andra förefaller det vara en svaghet att målet i sig inte omfattas av utvärderingen ifråga, dvs utvärderarna är låsta av SIDA:s ursprungliga målformuleringen. Är det uppsatta målet i sig tveksamt eller inadekvat, blir utvärderingen av begränsat intresse. I ett par fall dristar sig dock författarna att kritisera, eller ifrågasätta, möjligheten att på ett meningsfullt sätt utvärdera de mål som SIDA uppställt för importstödet.

I framförallt två av de tre utvärderingarna, läggs stor möda vid att visa om importstödet, eller förändringar i strukturen på detta stöd, påverkar tillväxten och tillväxtförutsättningarna i respektive land. För att kunna dra några slutsatser om tillväxteffekter krävs emellertid ett långsiktigt perspektiv, samtidigt som man kontrollerar för andra faktorer som kan tänkas påverka tillväxten. Sådana kontrollvariabler är t ex konjunktoreffekter, kapitalackumulering, storlek av offentlig sektor, regleringar och handelshinder. Endast i en av utvärderingarna återfinns ett sådant långsiktigt perspektiv (Tanzania), medan ansatsen i de övriga utvärderingarna är mer kortsiktigt. Men även i Tanzania-fallet är det en kortsiktig förändring i importstödets utformning, som endast omfattar ett fåtal år, som i första hand ska belysas. Generellt kan sägas att denna uppgift torde vara den ekonomiska vetenskapen övermäktig, dvs det går inte att utifrån ett tidsperspektiv om fem år säga någonting om tillväxteffekterna. Utvärderaren ställs inför det mindre angenäma valet att antingen ta på sig en uppgift som egentligen är omöjlig, eller att frångå de direktiv som ställts upp för utvär-

deringens utförande av SIDA. I både fallen riskerar utvärderingen att hamna på fel spår.

Den andra sidan av detta mynt är att i samtliga länder, även om ett par av dessa är extremt biståndsberoende, är den svenska biståndsandelen försumbar eller mycket liten i förhållande till ländernas hela ekonomi. Detta försvårar naturligtvis ytterligare uppgiften att bedöma effekterna av de svenska insatserna på tillväxt och andra makroekonomiska variabler. För t ex Mozambique, som erhåller ett förhållandevis stort importstöd från Sverige, så uppgår detta till ca 4 procent av landets total import.

Slutligen kan påpekas att endast i ett fall (Tanzania) framgår klart vad syftet är med utvärdering och de olika kriterier som skall appliceras. Detta beror framförallt på att SIDA:s direktiv för utvärdering bifogats i ett appendix, vilket gör att läsaren lättare kan tillgodogöra sig syftet med studien. Den utvärdering som omfattar fyra länder, eller – mer korrekt – sammanfattar fyra fristående utvärderingar, är det möjligt att upplägget i sig skymmer syftet¹. En mer fullödlig bild av utvärderingen hade sannolikt förmedlats om läsaren istället fått ta del av de ursprungliga utvärderingarna.

I den utvärdering som behandlar stödet till Cap Verde är syftet särskilt diffust. Läsaren får bl a veta att en av de övergripande frågorna som ska besvaras är huruvida "...should import support to Cap Verde follow the guidelines outlined in the *Principles of Import Support* (SIDA, January, 1993) and recently adopted by SIDA, or would Cap Verde's unusual need best be met by a different policy of import support?". Däremot invigs inte läsaren i vad dessa guidelines består i. Slutsatsen blir följaktligen att ett supplement som innefattar direktiven är till stor nytta för läsarens möjlighet att bedöma utvärderingens syfte och riktighet.

¹ Summeringen i sig är mycket välskriven men syftet med den är just att summera de tidigare genomförda utvärderingarna.

Kommentar till respektive utvärdering

I ovanstående avsnitt berördes utredningarna på ett övergripande plan med avseende på metod, samt i någon mån dokumentation och slutsatser. Jag kommer nu att kortfattat diskutera respektive utvärdering.

Fungibility and Sustainability. Import support to Tanzania (SIDA Evaluation Report 1993/5)

Tanzania har sedan länge varit ett av de största mottagarländerna för svenskt bistånd. Av de tre utvärderingar som här behandlats är detta den mest genomarbetade. På ett förtjänstfullt sätt redogör författarna för Tanzanias ekonomiska utveckling och biståndshistoria sedan 1970. Biståndsberoendet, effekten på industristruktur, och tendensen mot en de-industrialisering, kopplas till biståndets struktur och omfattning. Hur skiljer sig effekterna av projektstöd eller investeringsstöd i förhållande till ett mer generellt importstöd? Bör stödet inriktas på kapacitetshöjande åtgärder eller är det brist på komponenter och insatsvaror som hämmat utvecklingen? I detta sammanhang diskuteras också i vilken utsträckning biståndet konserverat en obsolet näringslivsstruktur och effekten av att en stor del av biståndet kanaliseras till statliga företag, så kallade parastatals, vilka är ökända för sin ineffektivitet.

Delar av teorin kring utveckling behandlas också utförligt och likaså diskuteras hur betydelsen av importstöd empiriskt kan underbyggas genom att t ex studera hur pass känslig produktionen är för förändringar i tillgång på utländsk valuta²². Vidare behandlas importstödet effekter på makroekonomiska variabler som inflation, växelkurs och det därmed sammanhängande "Dutch Decease" fenomen, dvs. om tillförsel av utländsk valuta genom bistånd leder till en övervärderad valuta. Slutligen diskuteras tillväxtkonsekvenserna av importstöd och dess fördelning på olika sektorer. Bl a konstateras att trots att tillväxten i första hand härrör ur jordbruksproduktionen, erhöll denna sektor en blygsam andel av biståndet. Vidare visas hur inkomstfördelning antagit en skevare form sedan 1975.

²² Produktionens så kallade utbudselasticitet med avseende på utländsk valuta eller import.

Utvärdering utmynnar i två huvudsakliga slutsatser. För det första, om importstödet ska fortsätta bör det ändra karaktär. Istället för att försöka övervaka och styra stödets användning, bör uppföljningen ske i termer av ett på förhand definierat resultat, t ex utgiftsnivåer för vissa sektorer, utbudsindikatorer, etc. Dessutom bör den liberalisering av fördelning av utländsk valuta som påbörjats, ytterligare förstärkas. Här föreslår författarna att ett auktionsförfarande etableras. För det andra riktar författarna in sig på importstödet lämplighet som medel att uppfylla biståndets mål, vilket definieras som att primärt avhjälpa fattigdom i utvecklingsländerna. Mot den bakgrunden föreslås att importstödet fasas ut, eller minskas, och att större resurser satsas på skola och utbildning i Tanzania.

Den senare rekommendationen är lätt att instämma i och den understöds också av flera studier som visar på positiva tillväxteffekter av en kvalitativ grundutbildning. Åtgärden har också en direkt koppling till modern tillväxtteori som betonar kunskap som den främsta drivkraften bakom tillväxt. Även den förstnämnda rekommendationen synes välmotiverad, dock förefaller problemen med att mäta ett resultat i termer av t ex utbudsindikatorer, givet Sveriges begränsade insatser, vara av minst lika svår art som de som diskuteras ovan. Det övergripande intrycket är dock att detta är en väl genomarbetad utvärdering som uppfyller de mål som stipulerats i direktiven.

Aid and Macroeconomics

(SIDA Evaluation Report 1993/2)

Utvärdering skiljer sig från de övriga så tillvida att dels sammanfattas resultaten av fyra fristående utvärderingarna avseende länderna Guinea-Bissau, Mozambique, Nicaragua och Vietnam, dels redogör författaren generellt för målen med svensk bistånd och biståndets makroekonomiska effekter.

Beträffande den sistnämnda delen är detta en i sig kompetent redogörelse för de teorier ur vilka slutsatsen kan härledas att brist på utländsk valuta utgör en hämsko på ekonomisk utveckling och tillväxt. Likaså redogörs för de makroekonomiska störningar som ett sådant stöd kan ge upphov till. Således kan tillförsel av utländsk valuta i biståndsform leda till att den långsiktiga exportpotentialen minskar, att

valutan apprecieras (vilket ytterligare försvårar export) och att inflationistiska mekanismer byggs in i ekonomin. Det finns ingenting att erinra mot denna framställning.

Resten av utvärderingen ägnas åt en sammanfattning av de fyra länderutvärderingarna. Samtliga länder lider av olika makroekonomiska obalanser, en svag exportutveckling, regleringar som hämmar pris- och löneflexibilitet samt ett högt biståndsberoende. Utvärderingarna förenas av den gemensamma ståndpunkten att brist på utländsk valuta har bromsat den ekonomiska utvecklingen och att ett importstöd därför har en positiv effekt på tillväxten. Trots detta så kommer de olika utvärderingarna fram till olika slutsatser när det gäller bibehållandet av importstödet i framtiden, vilket är något oväntat.

Bortsett från en rekommendation angående importstödet vara eller icke vara, innehåller utvärderingarna också en mängd andra förslag. Dessa förefaller ligga utanför direktiven och kommenteras inte i någon större utsträckning i sammanfattningen. Några axplock ur dessa ger dock vid handen att för t ex Guinea-Bissau anser utvärderaren att marknaden saknar möjlighet att klara fördelningen av utländsk valuta³. Likaså förordas att det svenska biståndet används för att motverka inflytandet av Världsbanken och Internationella Valutafonden. För Mozambique del rekommenderas en administrativ fördelning av utländsk valuta medan en total omvärdering av biståndet bör präglade den framtida relationen gentemot Nicaragua. Uppenbarligen kan flera invändningar framföras mot dylika tankegångar, det är dock svårt att på grundval av sammanfattningen säga mycket mer än så.

Istället bör kritiken riktas mot den grundläggande slutsatsen i de fyra fristående utvärderingarna, dvs att importstödet har haft en positiv inverkan på tillväxten. Utifrån sammanfattningen är möjligheten att bedöma rimligheten i denna slutsats begränsad. Är det det svenska biståndet som avses eller är det en mer generell diskussion? Mot bakgrund av den korta tidsperiod som utvärderingarna omfattar förefaller slutsatsen ogrundad. Frågan rörande samverkan mellan andra för tillväxten positiva faktorer och det svenska biståndet förbigås helt eller ägnas all för litet utrymme. I vilken utsträckning har utvecklingen

³ Detta och följande exempel återfinns i sammanfattningen av de fyra fristående utvärderingarna.

präglats av avreglering och mer marknadsinriktade lösningar i de olika ekonomierna? Detta tas inte upp i sammanfattningen av utvärderingarna. Det är därför svårt, för att inte säga omöjligt, att utifrån sammanfattningen göra sig en rättvisande bild av de respektive utvärderingarna.

Swedish Support to Cap Verde.

Denna utvärdering avviker på flera sätt från de två som ovan behandlats, dessutom verkar den inte ingå i SIDA:s serie av "evaluation reports". Vidare är syftet något oklarare jämfört med de två utvärderingar som behandlats ovan. Frågan om importstödet inverkan på tillväxt och andra makroekonomiska variabler berörs mycket knapphändigt. På ca två sidor konstateras att importstödet varit inflationsdrivande samt bidragit till en övervärderad valuta i Cap Verde. Som bl a påpekades i den förstnämnda utvärderingen (Tanzania), kan dock den inhemska regeringen neutralisera sådana effekter genom adekvat ekonomisk politik. Eftersom författarna hävdar att ekonomin i Cap Verde avreglerats och moderniserats, förefaller det inte otänkbart att förutsättningarna för att genomdriva sådana åtgärder borde ha funnits i Cap Verde.

Tyngdpunkten i utvärderingen ligger i stället kring importstödet nuvarande utformning. Författarna hävdar att en omfattande byråkrati, inte minst från SIDA:s håll, förhindrat att målet med importstödet uppnås. Dessutom har en omfattande detaljreglering av stödet lett till att vissa företag och institutioner gynnats på ett otillbörligt sätt. Slutsatsen som dras i utvärderingen är att stödet bör omformas till ett betalningsbalansstöd och alla ex ante krav på stödets användning bör upphöra. För att samtidigt försäkra sig om att stödet används på ett effektivt sätt bör denna åtgärd kombineras med att en "Bureau for International Market Information" etableras och övervakas av SIDA.

Även om en övergång till ett mer generellt system förefaller rimligt, verkar sättet på vilket detta ska ske inte särskilt genomtänkt. Ett auktionsförfarande är kanske en bättre metod än det som rekommenderas av författarna, i synnerhet som den ekonomiska utvecklingen och förståelsen för ekonomins funktionssätt, enligt författarna, varit entydigt positiv i Cap Verde.

Avslutande kommentar

Genomgående kan konstateras att – med mer eller mindre precision – uppfyller utvärderingarna de syften som direktiven utvisar. Det centrala problemet är att målet i sig, i detta fall effekten av importstöd på ekonomiernas tillväxt, inte underkastas en kritisk granskning. Enligt min mening är det, av skäl som anförts ovan, i samtliga utvärderingar varken meningsfullt eller möjligt att empiriskt säga något om importstödet effekter på ekonomisk tillväxt. Är ökad tillväxt målet med bistånd görs kanske de effektivaste insatserna genom att förbättra utbildningssystemet, från grundskola till kvalificerad managementutbildning.

Aspekter som behandlar strukturella problem i utvecklingsländerna, och vars symptom bl a är en brist på utländsk valuta, förtjänar i detta sammanhang en större uppmärksamhet. Vilka är de tillväxthämmande regleringarna? Hur ser konkurrensen ut i dessa ekonomier? Vilken roll och marknadsmakt har de statliga, ofta mycket stora, företagen? Leder den ekonomiska politiken till bättre tillväxtförutsättningar? Detta talar för en större internationell koordination av biståndet, vilket naturligtvis går stick i stäv med den rekommendation som framkom ovan, nämligen att svenskt bistånd skulle användas för att minska Världsbankens och Internationella Valutafondens inflytande.

Avslutningsvis bör man beakta den oheliga allians som kan förekomma mellan de så kallade "fristående" konsulter som gör utvärderingen och den beställande myndigheten. Det finns inget som hindrar att dessa konsulter också är inblandade i andra projekt som finansieras av myndigheten. Därför vore det önskvärt om uppdragsgivaren av utvärderingen var frikopplad från den myndighet som granskas, och istället sköttes av t ex Riksrevisionsverket. Alternativt kan man på andra sätt försäkra sig om att konsulten verkligen är fristående. Exempelvis kunde ett krav vara att utvärderaren inte har, eller tidigare haft, samarbete med myndigheten ifråga. Dessa problem uppstår naturligtvis endast för myndigheter där utläggning av uppdrag på konsulter är ett naturligt led i verksamheten. SIDA är just en sådan myndighet.

Bedömning 2

Alla tre rapporterna bygger till stor del på samma teoriapparat. Detta gäller framförallt utvärderingarna a) och b). Därför får mina principiella argument i tillämpliga delar gälla alla tre rapporterna. Jag väljer därför att föra den principiella argumentationen utifrån utvärdering a). Jag har dessutom ett antal specifika kommentarer till utvärderingarna b) och c).

BEDÖMNING AV UTVÄRDERING a)

1 Utvärderingens syfte

Syftet med utvärderingen är att kartlägga de makroekonomiska effekterna av svenskt importstöd till Tanzania. Enligt författarna har tidigare utvärderingar i första hand utgått från de mikroekonomiska effekterna av biståndet.

2 Metod

Författarna illustrerar Tanzanias utveckling under den tid då importstödet tillämpats med hjälp av grafer som beskriver import-, export- och BNP-utveckling under perioden. Ingen sofistikerad ekonometrisk metod har tillämpats. De teoretiska resonemangen bygger på traditionell makroteori.

3. A och B. Dokumentation

Jag är själv inte speciellt inläst på detta område. Därför kan min bedömning av studiens tillförlitlighet och relation till tidigare studier inte bli annat än rudimentär.

Studien följer gängse standard att hänvisa ordentligt till tidigare studier samt till använda källor (som nämnts kan jag emellertid inte ta

ställning till om dessa referenser är relevanta/uttömmande). Dokumentationen ger dock i stort sett ett vederhäftigt intryck.

C. Metod

Den empiriska metoden är som nämnts tämligen enkel. Så vitt jag kan bedöma bygger de slutsatser som dras uteslutande på grafiska tidssektorerstudier och enkla korrelationer. (Det är möjligt att det ligger en mer grundlig analys bakom. Det framgår emellertid inte av skriften.) Denna metod tillåter i strikt mening inte att slutsatser dras om samband mellan variabler, i synnerhet inte slutsatser om *causala samband*. Det är fullt möjligt att andra variabler, som inte tas med i analysen kan förklara de samband som förefaller gälla. Det är dessutom möjligt att någon variabel drivs av "inneboende" mekanismer vilket gör gängse ekonometrisk metod otillförlitlig.

De teoretiska resonemangen bygger på traditionell makroteori, anpassad till länder med svagt utvecklade institutioner och starka planekonomiska inslag. I huvudsak verkar författarna vara väl förtrogna med teorin även om tveksamhet föreligger i vissa resonemang. Ett exempel är synen på växelkursens bestämningsfaktorer och implikationer. Det finns visst empiriskt och teoretiskt stöd för slutsatsen att den *nominella* växelkursen, på längre sikt, i första hand bestäms av den relativa inflationsutvecklingen länder emellan, d.v.s. den relativa penningpolitiska inriktningen, medan nettoexportaspekter och "spekulativa" finansiella flöden är av stor vikt på kortare sikt. Författarna verkar, i likhet med traditionen inom biståndsteori i allmänhet, (exemplifierat av utvärderingarna b och c ovan) bortse från det långsiktiga perspektivet och betrakta växelkursen som bestämd av reala flöden.

Detta synsätt har även implikationer för distinktionen nominell-real växelkurs. På sid 49 argumenterar författarna för att ett inflöde av utländsk valuta orsakat av importstödet leder till en appreciering, vilket tränger undan nettoexporten. Det är rimligt att anta att inflödet av forex leder till en *nominell* appreciering på kort sikt. Vad som är avgörande för nettoexporten är emellertid den *reala* växelkursen. Den senare bestäms, t.ex. av skiftande preferenser för länders produkter samt efterfråge- och produktivitetsutvecklingen i tradable- och non-tradable-sektorerna, vilka diskuteras på annat håll i skriften. Den reala

växelkursen torde dock inte, mer än mycket kortsiktigt, bestämmas av in- och utflöden av forex orsakade av importstöd. Det är således inte säkert att dessa flöden får bestående effekter på nettoexporten under en så lång tidsperiod som det är fråga om i detta fall.

Det finns även vissa kopplingar till modern s.k. endogen tillväxtteori. Således påpekas vikten av humankapital- och infrastrukturinvesteringar. En ståndpunkt, som är förenlig med modern tillväxtteori och som hävdats av Robert Lucas, är att investeringar i fysiskt kapital inte kan orsaka "tillväxtmirakel", d.v.s. varaktigt högre tillväxttakt jämfört med omvärlden eller jämfört med tidigare perioder i ett lands historia. Det finns i stället starka krafter, s.k. "upphinnareffekter" som gör att tillfälliga tillväxtövertag orsakade av investeringar i fysiskt kapital äts upp av internationella kapitalflöden. En annan faktor som begränsar tillväxteffekterna av kapitalinvesteringar är substitution mellan olika produktionsfaktorer inom landet. Båda dessa faktorer gör att marginalavkastningen av kapitalinvesteringar tenderar att avta, vilket i sin tur stoppar upp tillväxtökningen.

En implikation av teorin är snarare att investeringarna *beror på* tillväxten, d.v.s. avkastningen på investeringarna. Viktiga förutsättningar för tillväxt *och* investeringar är enligt detta synsätt ordnade äganderättsförhållanden och stabila institutioner i samhället.

Däremot kan, enligt modern teori, vissa typer av *humankapitalinvesteringar* permanent göra ett lands relativa tillväxtbana brantare. Exempel på denna typ av humankapitalinvesteringar är "korsbefruktningar" mellan branscher, där erfarenheter från en bransch förs över till andra branscher. Om denna teori stämmer borde tillväxten vara lägre i länder som investerar relativt mycket i mogna branscher och därmed konserverar industristrukturen och högre i länder där satsningar görs på nya branscher (jämför t.ex. Sverige och Japan/Korea).

En nödvändig förutsättning för att dessa tillväxtmekanismer ska kunna verka är sannolikt både en existerande industristruktur och en grundläggande utbildningsnivå i samhället. Det första har stimulerats genom importstöd riktade till tanzanisk industri och det sistnämnda kan uppnås genom satsningar på utbildning för alla. I enlighet med dessa argument propagerar författarna även för att svenskt importstöd fortsättningsvis bör inriktas på utbildning av de stora massorna.

D Slutsatserna

Författarnas huvudslutsatser är att den ensidiga inriktningen av importstödet under 1970-talet på stora investeringar snarast har bidragit negativt till Tanzanisk import, export, tillväxt och fattigdomsbekämpning. Följden har blivit ett ökat beroende av importstöd. Man propagerar därför för ett framtida importstöd inriktat på infrastruktur- och utbildningssatsningar. Syftet med stödet måste enligt författarna vara att befärja produktivitetens utvecklingen i jordbrukssektorn, eftersom Tanzanias komparativa fördelar även fortsättningsvis kommer att finnas inom denna bransch. På det sättet skulle Tanzania kunna sluta gapen mellan inhemskt sparande/investeringar och mellan exportförmåga/importbehov.

P.g.a. begränsningar i både empirisk och teoretisk metod vilar slutsatserna delvis på lösan grund. Detta är i och för sig inte unikt för de aktuella studierna. Varje försök att klarlägga orsakssamband i så komplicerade sammanhang som gällande ekonomiers funktionssätt är med nödvändighet behäftade med stor osäkerhet. Det förefaller dock inte som om författarna till fullo utnyttjat "batteriet" av tillgängliga tekniker för att minska osäkerheten. En enkel regressionsanalys hade t ex kunnat ge slutsatserna något större tyngd. Användandet av modern tidsserieanalys hade kunnat ytterligare bidra till att öka tillförlitligheten.

Det är svårt för mig som saknar ingående kunskaper om de aktuella länderna att bedöma slutsatsernas rimlighet liksom andra potentiella förklaringar till den dåliga ekonomiska utvecklingen i allmänhet i många utvecklingsländer. En näraliggande alternativ- eller kompletterande förklaring tycks dock vara den fortsatta bristen på stabila institutioner, fasta spelregler och ekonomiska incitament. Detta intryck förstärks av det faktum att även andra ekonomier, *som inte varit utvecklingsländer och inte mottagit utländskt importstöd* men i huvudsak tillämpat planekonomisk resursallokering, brutit samman under de senaste 7-8 åren. Författarna är naturligtvis medvetna om bristerna i ekonomins funktionssätt men anser att importstödet har bidragit ytterligare till den negativa utvecklingen.

Dessutom är naturligtvis länder vars export närmast uteslutande utgörs av jordbruksprodukter, beroende av klimatologiska - och andra faktorer som påverkar jordbruksproduktionen. Många länder i Afrika,

däribland Tanzania, har t.ex. drabbats av svår jordförstöring p.g.a. urskilningslös avverkning av skogar. Andra, mer globala faktorer har ytterligare bidragit till en klimatförändring som kan ha försämrat produktiviteten inom jordbruket. Sådana "utbudschocker" förefaller kunna vara väl så viktiga för utvecklingsländers importförmåga.

E Kostnadseffektivitet

Utvärderingen av importstöds effekter på Tanzania har resulterat i en mycket omfattande rapport. Det är möjligt att rapporten hade kunnat vara något kortare utan att den förlorat nämnvärt i substans. Tillämpande av en striktare empirisk metod hade kunnat öka resultatets tillförlitlighet. Jag är övertygad om att författarna förfogar över de nödvändiga kunskaper som krävs för att genomföra detta utan extra resursåtgång.

F Objektivitet

Det är inte uppenbart att författarna drivs av någon annan ambition än att göra en objektiv utvärdering av importstöds politiken. Däremot är det alltid så att man genom valet av empirisk och teoretisk utgångspunkt i viss mån "tar ställning".

4 Allmänna synpunkter och kommentarer

Under tidigare skeden var svensk importstöd till Tanzania först och främst inriktat på uppbyggnad av den tanzaniska kapitalstocken. Under slutet av 1970-talet blev det alltmer uppenbart att behovet av ytterligare *utbyggnad* av kapitalstocken var underordnat behovet av ett ökat *utnyttjande* av existerande kapitalstock. Syftet med importstödet sedan 1980-talets början har varit att öka importen av intermediära varor vilka i sin tur ska bidra till att öka kapacitetsutnyttjandet i den Tanzaniska ekonomin.

Som framgår av titeln är "fungibility" ett av nyckelbegreppen i utvärderingen. Så vitt jag vet brukas begreppet uteslutande inom litte-

raturen om biståndsekonomi. I princip tycks emellertid "fungibility" vara liktydigt med begreppen "indirekta effekter" eller "crowding-out" (alternativt "crowding-in"). Begreppet omfattar dock även faktorer som inte vanligtvis inbegrips i "crowding-out", t ex korruption. Författarna grundar sitt betonande av "fungibility" på det faktum att den tanzaniska importen *föll* som andel av BNP i en tid (slutet av 1970-till slutet av 1980-talen) då de utländska importstöden var betydande. Det fanns uppenbarligen faktorer som gjorde att nettoeffekterna på importen inte motsvarades av volymen på det ursprungliga importstödet.

Författarnas beskrivning av "fungibility"-effekterna i Tanzania är kortfattat följande: 1970-talets projektstöd var inriktat på stora investeringar. Dessa påverkade den tanzaniska ekonomin på flera sätt. Tanzanias exportvaror, nästan uteslutande jordbruksvaror, prissätts på marknader med givna världsmarknadspriser. Den direkta effekten av investeringsprojekten var att driva upp kostnaderna för inhemska produktionsfaktorer; en form av "cost-push inflation". Även om kostnadsökningarna initialt var lokala tenderade de att sprida sig till hela ekonomin, inklusive den exportinriktade sektorn. Den indirekta, multiplikatoreffekten, betonas mer av författarna. När produktionsfaktorer anställdes inom projekten ökade inkomsterna och därmed efterfrågan på varor, såväl "tradable" som "non-tradable", i första hand jordbruksvaror. Eftersom priserna på "tradables" är givna på världsmarknaden steg relativpriset på "non-tradables", vilket innebar en real tanzanisk appreciering. Relativprisförändringen vred produktionen bort från "tradables" och till fördel för "non-tradables". Båda dessa mekanismer innebar att Tanzanias internationella konkurrenskraft försämrades varvid exporten minskade.

Importstödet tar formen av transfereringar av utländsk hårdvaluta. Denna överförs sedan på olika vägar till inhemska företag. Den grundläggande tanken är att de inhemska företagen skall betala för den utländska valutan med motsvarande belopp i tanzanisk valuta (s k "countervalue funds"). Detta fungerade inte av två skäl. Dels avstod myndigheterna i realiteten, i många fall, från att kräva betalning (korruption m.m.) eller fördröja inkrävandet, dels var den officiella valutakursen konstlat hög (d.v.s. den tanzaniska shillingen konstlat stark). Följden blev att importstödet positiva effekt på valutareserven övervägde indragningar av inhemska krediter, varvid penningutbudet öka-

de. Importstödet bidrog således till en monetarisering av ekonomin, till de inflationistiska tendenserna och konkurrenskraftförsämringen.

Både den direkta- och multiplikatoreffekten tenderade således att minska den tanzaniska exporten och därmed importförmågan. Möjliga förvärrades situationen ytterligare, åtminstone på kort sikt, av ytterligare inflation p.g.a. hanteringen av "countervalue funds". Eftersom det inte skedde någon nämnvärd inhemsk produktion av intermediära industrivaror medan importförmågan försämrades, minskade tillgången till industrivaror och därmed kapacitetsutnyttjandet i de tanzaniska industri- och jordbrukssektorerna. Författarna hävdar att denna process inte bara delvis *urholkade* tillväxteffekterna av 1970-talets investeringsstödspolitik. Man menar att undanträngningseffekterna o.m. övervägde de direkta effekterna av projektstödet, så att netto-effekterna faktiskt var en *fallande* tanzanisk produktion. Urholkandet av den tanzaniska importförmågan förstärktes dessutom av kraftigt stigande oljepriser i slutet av 1970-talet, vilket ledde till närmast en kollaps av den tanzaniska ekonomin.

Det andra nyckelbegreppet är "sustainability". Huvudfrågan är här i vilken utsträckning 1970-talets importstöd ledde till att Tanzania blivit mer eller mindre beroende av importstöd i framtiden. Frågeställningen sammanfattas enligt två kriterier: I vilken utsträckning har importstödet bidragit till att minska gapet mellan inhemskt sparande och investeringar respektive gapet mellan exportförmågan och importen? Grundat på argumenten ovan drar författarna slutsatsen att importstödet snarast har bidragit till att försvåra uppfyllandet av de två kriterierna, vilket i sin tur ökade behovet av importstöd.

Slutomdöme

Som jag hoppas framgår av mina kommentarer anser jag att utvärderingen till stora delar är mycket bra och relevant. Det är dessutom intressant läsning för en oinitierad i dessa frågor. Den pedagogiska uppbyggnaden bidrar till att göra rapporten läsvärd, även om den i vissa stycken är onödigt lång.

BEDÖMNING AV UTVÄRDERING b)

Rapport b) är en sammanfattning av fyra separata bedömningar som gjorts om effekterna av svenskt importstöd till Guinea-Bissau, Mozambique, Nicaragua och Vietnam. Huvudslutsatserna som refereras i rapporten är att svenskt importstöd har varit för litet för att påtagligt påverka tillgången på utländsk valuta, importförmågan, tillväxten och fattigdomen i de aktuella länderna. Delrapporternas författare vågar ändå hävda att effekterna varit marginellt positiva.

Författaren till rapport b) gör inget försök att bedöma de enskilda rapporternas kvalitet. Han konstaterar emellertid att SIDA:s riktlinjer rekommenderar att stöd i första hand bör riktas till länder där importstödet verkningsgrad kan förväntas vara högst. Detta torde vara fallet i länder som genomfört betydande avregleringar och makroekonomiska stabiliseringar. Författaren noterar att bedömningarna resulterar i motsatt ståndpunkt: "Öka importstödet till Guinea-Bissau, där verkningsgraden är lägst och minska stödet till Vietnam med bäst förutsättningar att dra nytta av stödet."

De aktuella rapporterna tillämpar samma teoretiska maskineri och samma empiriska metod som bedömning a).

Slutomdöme

Rapport b) är en välskriven och föredömligt kortfattad sammanfattning av teorin samt av de fyra delrapporterna. I övrigt gäller samma principiella omdömen som för rapport a).

BEDÖMNING AV UTVÄRDERING c)

Den tredje utvärderingen har en delvis annorlunda utgångspunkt. Den är mer inriktad på de institutionella och byråkratiska hinder på såväl svensk som Cape Verdisk sida som hindrat ett effektivt utnyttjande av svenskt importstöd. Syndabocker har varit såväl SIDA och Nord-banken som Cape Verdes centralbank och importmyndigheter. Dess-utom betonas, liksom i de övriga bedömningarna, effektivitetsförluster av att stödet i alltför stor utsträckning gått till icke-

konkurrensutsatta verksamheter/företag. På det sättet hävdar författarna till bedömning c), har importstödet mest bidragit till att bevara ineffektiva strukturer.

Bedömning c) är mindre välskriven och pedagogisk än de övriga. Den fokuserar väl mycket på siffror och ingående resonemang om utvecklade byråkratiska procedurer. Resonemangen är till stor del omöjliga att följa i detalj. Mycket av detta hade kunnat sammanfattas i betydligt kortare form, varvid utrymme kunnat frigöras för mer fundamental analys.

Trots dessa brister får man en relativt klar bild av en mycket byråkratisk apparat som till stora delar omöjliggör positiva effekter av importstödet. Dessutom har uteblivna betalningar av countervalue funds varit betydande. Författarnas slutsats av detta är: "What this means is that the real effect of much of the Swedish support has been to subsidize a selected number of public companies (often loss-making) and to provide free credit to private entrepreneurs. The development effect of such usage of the Swedish support is, at best, doubtful."

Författarna rekommenderar att svenskt stöd i fortsättningen ska utgöras av betalningsbalansstöd, d.v.s. obundna transfereringar av utländsk hårdvaluta. Man bör således lämna systemet där SIDA lämnar importstöd efter speciell prövning. Därmed förlorar SIDA möjligheterna att påverka utvecklingsstrategin i Cape Verde.

Slutomdöme

Rapporten gör inget ambitiöst försök att bedöma de makroekonomiska implikationerna av svenskt importstöd, t.ex. så som görs i de övriga bedömningarna. Man följer i stort sett stödets färd från Sverige till hamnarna i Praia och Mindelo men inte längre. Visserligen kan man dra vissa slutsatser genom att studera vilka företag som fått ta del av stödet och på vilka grunder detta har skett, men det räcker inte för att bedöma effekterna på ekonomin i sin helhet. Några illustrationer av Cape Verdes ekonomiska utveckling under stödperioden hade varit en tillgång i detta sammanhang.

19 Resultat av särskilda skogsvårdsinsatser i norra Sverige (Skogsstyrelsen)

Bedömning 1

Utvärderingens syfte

Utvärderingen syftar till att bedöma effekterna av ett selektivt bidragsprogram (SSI-programmet), avsett att höja aktiviteten i skogsvården hos privata skogsägare i Norr- och Västerbotten, och administrerat av Skogsstyrelsen/skogsvårdsstyrelserna i respektive län. Det treåriga programmet startade sommaren 1988. Kostnaden var totalt 80 miljoner kronor därav en sjundedel för administration. Bidragen kom att delas ut till 3300 privata skogsfastigheter (knappt en sjättedel av alla privata skogsfastigheter i området), och stimulerade till en ökning av sådana skogsvårdsåtgärder som har en mycket låg, eller negativ, privatekonomisk lönsamhet – exempelvis röjning, gallring och stamkvistning – men som bland jägmästare allmänt anses bedrivas i för liten omfattning. Med utvärderarnas egna ord skulle de "kvantifiera merinsatserna i skogsvård och avverkning", "fastställa effekter på kort och lång sikt", "kartlägga eventuella positiva och negativa bieffekter" av nämnda bidragsprogram och "bedöma och förklara programmets måluppfyllelse". I utredningen framgår inte helt klart vilka mål staten hade för bidragsprogrammet.

Utvärderingens metod

Utvärderarna beskriver själva sin metod som en köksstol, där de fyra benen utgörs av:

- 1) En totalundersökning av de 3300 fastigheterna som berörts av SSI-programmet, insamlad av Skogsvårdsstyrelserna för administration av detsamma och för utbetalning av bidragsmedel. Fastighetsägarna rapporterar här faktiskt genomförda åtgärder de senaste tre åren, och planerade åtgärder inom SSI-programmets ram för de följande tre åren. Därtill finns uppgifter från Skogsvårdsstyrelsen om i vilken utsträckning planerade åtgärder faktiskt genomförts.
- 2) En enkätundersökning av drygt 300 skogsägare som inte berördes av programmet ("referensgruppen"), utformad på liknande sätt.
- 3) En enkätundersökning om vad skogsägare (inom och utom SSI-programmet), skogsbolag, skogsägareföreningar, kommuner och konsulenter hos Skogsvårdsstyrelserna tycker om SSI-programmet. Totalt erhöles 648 svar på 885 utskick.
- 4) Insamling av "bakgrundsstatistik", dvs av data om de totala röjningarna och gallringarna i tio (av 30–40) skogsvårdsdistrikt i de två nordligaste länen under åren 79–92, samt data om statliga bidrag till dessa aktiviteter under denna tid i de två länen.

Stolens sits utgörs av "modeller för skattning av effekter". Med detta menas i huvudsak deskription av data och en "modell" som säger att om de 3300 skogsägarna som fick särskilda bidrag ("programgruppen") inte hade fått dessa bidrag, så skulle de ha betett sig som de 330 skogsägarna i referensgruppen. Med andra ord: eftersom referensgruppen och programgruppen historiskt betedde sig ungefär lika (programgruppen var dock även tidigare mer aktiv), så kan man anta att de skillnader som observerades medan SSI-programmet var i kraft förklaras av just detta. Därutöver presenteras nuvärdeskalkyler för den privatekonomiska (?) lönsamheten av att genomföra fyra specifika skogsvårdsåtgärder, vilka berättigade till bidrag i SSI-programmet, vid en specifik bonitet ("jordmån"). Den enda förutsättningen som preciseras, förutom bonitet, är kalkylräntan, vilken sattes till 2 procent.

Slutligen utgörs stolens rygg av analys och slutsatser – se nedan.

Dokumentation

I huvudsak är dokumentationen utförlig – mängden diagram och tabeller är ibland nästan överväldigande. Bitvis saknas dock viktig dokumentation. Det finns inga referenser till tidigare studier (bortsett från en egen pilotstudie) – bara en litteraturlista i slutet av rapporten. Den mest uppenbara bristen på dokumentation gäller dock modellen för nuvärdeskalkyl – se ovan.

Data

Datamaterialet i sig är i huvudsak bra, möjligen alltför omfattande (se nedan under kostnadseffektivitet). Den enkät som gjorts om åsikter är dock av tvivelaktigt värde: samtliga tillfrågade parter gynnas bidragen, oavsett om de är samhällsekonomiskt riktiga eller ej. Det är därför föga förvånande att nästan alla är positiva. Desto mer anmärkningsvärt är det istället att personalen på skogsvårdsstyrelsen, som har mest insyn i projektet och som bör vara den grupp som lägger mest vikt vid samhällsekonomisk effektivitet, är negativa till projektet.

Ett särskilt dataproblem är att programgruppen inte valdes ut slumpmässigt (däremot valdes referensgruppen slumpmässigt). Detta diskuteras under slutsatser.

Metod

De data som samlas in är i stor utsträckning relevanta, men det kriterium som används för att mäta graden av måluppfyllelse i bidragsprogrammet är alltför snävt. Man har bara sökt belysa om och i vilken utsträckning mängden skogsvårdsåtgärder ökat; inte huruvida det är bra om dessa ökar, och inte heller om de ökat i en sådan utsträckning att programmets kostnader motiveras. Vad gäller beräkningen av ökningen av antalet dagsverken förefaller det finnas en hel del metodmässiga problem. Detsamma gäller beräkningarna av nuvärdeskalkyler – som nämnts ovan finns dock ingen dokumentation som möjliggör en ordentlig bedömning av detta. Slutligen kan man ifrågasätta värdet

av att fråga de som gynnas av ett bidrag om vad de tycker om detta bidrag – se vidare under slutsatser.

Slutsatser

I korta drag finner utredarna att programmets målsättning uppnåtts, enär aktiviteten (räknat i antal hektar) ökat med cirka 85 procent. Ökningen var dramatisk (flera hundra procent) för de skogsvårdsåtgärder som gavs särskilt stora bidrag – röjning, förstagångsgallring och stamkvistning. Vidare visade enkätundersökningen om åsikter att alla tillfrågade grupper var positiva till SSI-programmet – utom de skogsvårdskonsulenter som administrerade det. Det faktum att det totala virkesutbudet minskade, som följd av att den slutavverkade arealen minskade med 25 procent, fördunklar, enligt utredarna, inte programmets lysande framgång. Slutsatserna förefaller mig mycket svagt underbyggda.

För det första: det diskuteras aldrig om det är bra att öka skogsvårdsåtgärderna – det förutsätts att så är fallet. Rimligen bör samhället subventionera näringsverksamhet endast i de fall marknaden inte själv åstadkommer ett tillfredsställande resultat. Ett tänkbart skäl för att subventionera skogsvård, som i någon mån berörs, är att det kan skapa jobb i Norrlands inland. Man finner att 40 000 dagsverken skapats av SSI-programmet, men man räknar aldrig ut att detta skett till en kostnad av 2000 kronor per dagsverke, och diskuterar följaktligen aldrig om kostnaden är låg eller hög.

För det andra: man diskuterar aldrig huruvida SSI-programmet lett till en tidsmässig omfördelning av åtgärder. Detta trots att ett uttalat syfte med utredningen vara att "fastställa effekter på kort och lång sikt". En mycket rimlig tolkning är att de skogsägare som valts ut att delta i programmet bedömer bidragen som temporära. Det är därför lönsamt för dem att genomföra den typ av åtgärder som ges särskilt förmånliga subventioner under den tid projektet löper, och följaktligen att förlägga åtgärder som ej subventioneras – t ex slutavverkning – till efter projektets sluttid. Eftersom träden växer, och eftersom tillväxten avtar mycket långsamt, så är kostnaden för att uppskjuta avverkning till tre år efter den optimala tidpunkten låg. Den genomsnittliga fastighetsägaren erhöll extra bidrag om totalt 24 000 kr, vil-

ket i många fall säkert räcker för att göra en sådan omDisposition lönsam. Slutsatsen bör alltså vara att de långsiktiga effekterna av ett permanent bidrag skulle bli betydligt mindre. Detta resonemang är konsistent med det faktum att såväl utredarna som andra bedömare anser att effekten av SSI-programmet varit starkare än effekten av tidigare, mer permanenta, bidragssystem.

Ett problem som endast i någon mån berörs hänger ihop med hur de 3300 skogsägarna valdes ut. Det har tydligen varit upp till varje skogsvårdsdistrikt att utforma urvalsprocessen. Man kan ifrågasätta varför ett godtyckligt urval av skogsägarna ska gynnas av subventioner på i genomsnitt 24 000. Det tycks också som om man medvetet gynnade stora och aktiva skogsägare – kanske dessutom just sådana som hade behov av att genomföra skogsvårdsåtgärder under den tid projektet löpte. Uppenbarligen kan detta, åtminstone delvis, förklara att aktivitetsgraden varit högre i programgruppen än i referensgruppen.

Slutligen kan nämnas att den långsiktiga sysselsättningseffekten endast helt kort berörs. Kanske beror detta på att avverkningsvolymen faller, och på att man antar att varje ytterligare kubikmeter genererar ett sysselsättningstillskott som motsvarar den genomsnittliga arbetskraftsåtgången per förbrukad kubikmeter i skogsindustrin. Detta är givetvis ett mycket förenklat synsätt, som bl a bygger på att industrins efterfrågan på virke är fullständigt elastiskt, och på att marginella nytillskott av kapacitet är baserade på samma teknik som en genomsnittlig befintlig anläggning. I själva verket är t ex MoDos nya tidningspappersmaskin tre gånger mindre personalintensiv än den genomsnittliga tidningspappersmaskinen.

Kostnadseffektivitet

Det förefaller som om alldeles för mycket resurser lagts ned på datainsamling. Var det verkligen nödvändigt med dels en totalundersökning av de 3300 berörda skogsägarna, dels en totalundersökning av röjning och förstagallring i 10 skogsvårdsdistrikt, och dels en enkät om åsikter om programmet till nästan 900 personer? Uppenbarligen inte, vilket framgår av kritiken ovan! Rimligen hade enkäten kunnat strykas, eller åtminstone inskränkas till ett betydligt mindre antal, och

kanske framför allt till skogsvårdsstyrelsernas personal. Vidare hade det räckt med ett urval av de 3300 skogsägarna. Istället skulle betydligt mer möda lagts ned på analys av datamaterialet. Eventuellt borde man också ha försökt matcha referensgruppen bättre med programgruppen, dvs man borde ha försökt tillämpa samma urvalsprinciper för båda grupper.

Objektivitet

Objektivitet är ett värdeladdat ord; att säga att utredarna brustit i objektivitet kan uppfattas som en allvarlig anklagelse. Det förefaller dock helt klart för mig att utredarna haft en synnerligen okritisk inställning till det bidragssystem man fått i uppgift att utreda. Inom de snäva kriterier de satt upp för sig själva har de i huvudsak varit objektiva, utom kanske i valet av diskonteringsränta (vid mer rimliga diskonteringsräntor blir nästan alla skogsvårdsåtgärder olönsamma, till de flesta jägmästares stora förtvivlan). En mer fullständig objektivitet förutsätter dock en mer kritisk inställning än vad utredarna visat upp.

Sammanfattande bedömning

Inom det granskade utvärderingsprojektet har man faktiskt utvärderat sig själva. Detta kunde dock inte förhindra ett tämligen undermåligt resultat. Att döma av omfattningen av datainsamlingen, och av antalet presenterade diagram, har kostnaden för utvärderingen varit relativt hög. Redan hösten 1988 började Forskningsstiftelsen Skogsarbeten arbeta med utvärderingen; slutresultatet presenterades 31 oktober 1991 och är skriven av tre författare. Insamlingen av det slutliga datamaterialet påbörjades i mars 1991.

Förutom den uppenbara kostnadsineffektiviteten, så kan man främst kritisera utvärderingen för en alltför okritisk och snäv syn på utvärderingsuppdraget. Man har inriktat sig på att dokumentera *att* skogsvårdsinsatserna har ökat, men man har egentligen aldrig analyserat huruvida det är önskvärt att de ökar, eller om i så fall ökningen har varit tillräckligt omfattande för att motivera SSI-programmets kostnader. Vidare diskuteras knappast heller *varför* skogsägarna ökar sin aktivi-

tet. En rimlig hypotes är att dessa utnyttjade temporära bidrag, och att därför en stor del av den observerade ökningen berodde på att man utförde åtgärder som man annars skulle ha utfört senare, och istället sköt upp andra, ej bidragsberättigade åtgärder.

Bedömning 2

3 Utvärderingens syfte;

Ett treårigt försöksprogram – "särskilda skogsvårdsinsatser i norra Sverige" – inleddes sommaren 1988 med avsikt att öka den totala aktiviteten inom Västerbottens och Norrbottens privatskogsbruk med 10 % samt ge underlag för långsiktigt ökad avverkning och sysselsättning. Programmet finansierades med statsbidrag omfattande 80 miljoner.

Utvärderingen i fråga syftar till att:

- kvantifiera programmets effekter i form av merinsatser i skogsvård och avverkning,
- fastställa kort- och långsiktiga effekter,
- kartlägga bieffekter, samt
- bedöma måluppfyllelsen.

4 Metoden

Utvärderingen redovisar mätbara förändringar av programmet samt söker understödja slutsatser genom indirekt analys av attityder till projektet. Den huvudsakligen metoden går ut på att kartlägga omfattningen av skilda skogsbruksåtgärder före och efter introduktionen av programmet, dels på de fastigheter som omfattas av programmet (programfastigheter) och dels på en population av övriga fastigheter (referensfastigheter). De resultat som erhålles belyses vidare med hjälp av allmän statistik samt en enkätundersökning.

A) Dokumentationen av utvärderingen får betraktas i huvudsak väl genomförd. Tillförlitligheten av det sifferunderlag som framtagits

framstår som tillfredsställande. *Resultatens* tillförlitlighet är dock förknippade med större frågetecken. Av mindre betydelse är att jämförelse av utvecklingsförlopp genomförs utan att hänsyn tas till skillnader mellan fastighetsgrupperna. Andra allvarigare problem, vilka diskuteras närmare nedan, gör emellertid att själva slutsatsernas tillförlitlighet blir svag.

B) Presenterade data framstår som valda på ett rimligt sätt och som konsekvent analyserade. Däremot är det svårt att bedöma tillförlitligheten av den analys som baseras på dessa data. Orsaken är dels att programfastigheternas beteende sannolikt påverkats av förväntningar om programmet samt, dels att de fastigheter som ej omfattats, av samma skäl, inte går att betrakta som goda referensobservationer. En möjlig väg att undgå de svårigheter som uppstått kunde ha varit att välja en alternativ referenspopulation, bestående av t.ex. fastigheter i en liknande region belägen utanför de omfattade länen, vilka över huvud inte varit aktuella för något program av denna typ. Vissa jämförande uppgifter från andra områden hade under alla förhållanden behövts för att data fullt ut skulle kunna betraktas som "lämpliga".

C) Utvärderingsmetoden är förknippad med det grundläggande problemet att programfastigheternas beteende sannolikt påverkats av förväntningar om programmet och referensfastigheternas av att de inte omfattas därav. Detta påtalas också ett flertal gånger i rapporten, utan att någon egentlig hänsyn tas till saken. Av dessa skäl blir utvärderingsmetodens lämplighet tveksam, även om den i sig förefaller konsekvent genomförd.

D) Slutsatserna, vilka kan sägas vara sammanfattade under rubriken "syntes" s. 46 i rapporten, ger ett blandat intryck. Den huvudsakliga slutsatsen, att en sammanvägning av resultaten ger en mycket positiv bild av programmets utformning och effekter, framstår inte som väl underbyggd. Den från början avsedda ökningen av avverkningen har inte kommit till stånd under den observerade metoden, vilket dock i sig inte behöver vara negativt (vilket också betonas i utvärderingen). De effekter i skogsvård som däremot redovisas (vilka främst består i en ökad aktivitet på programfastigheterna jämfört med referensfastigheterna) kan mycket väl vara en följd av att programfastigheterna, vil-

ka i många fall uppges ha erhållit subventioner längre tillbaka i tiden, haft en låg aktivitet före programmets genomförande i avvaktan på det. Referensfastigheterna å sin sida kan ha dämpat sin aktivitet på grund av programmet, t.ex. i förhoppningen av att även de skulle bli föremål för subventioner i framtiden.

Det skall också konstateras att ingen egentlig analys genomförts av programmets samhällsekonomiska avkastning. De i detta sammanhang mycket viktiga effekterna på industrin kan inte diskuteras så lättvindligt som görs på sid 37. Däremot ska noteras att den diskussion som i övrigt följer under "syntesen" förhåller sig betydligt mindre kategorisk än det först avgivna huvudbudskapet. Man betonar där framför allt de stora svårigheterna att beräkna vilka de slutliga långsiktiga effekterna egentligen blir.

E) Utvärderingen är i långa stycken ambitiöst upplagd och genomförd. Rapporten är på det hela taget välskriven och genomtänkt. Sålunda leder den fram till en rad intressanta observationer. Genom den grundläggande bristen i hanteringen av programmets beteende-effekter framstår utvärderingens kostnadseffektivitet likväl som låg. Man har i alltför stor utsträckning "skjutit mygg med kanoner".

F) Utredarens objektivitet skulle kunna ifrågasättas eftersom huvudslutsatsen förefaller bristfälligt underbyggd. Man kan ju inte heller bortse från möjligheten att Skogsstyrelsen, för vilken utvärderingen genomförts, har ett intresse av att läsa om positiva effekter av subventioner till skogsbruk. Å andra sidan ska poängteras att utredaren lagt sig vinn om att uppmärksamma svagheter i utvärderingen, även om dessa senare inte vägts in i den heltäckande bedömningen. Det är inte heller helt lätt att komma runt de svårigheter som här påtalats. En allmän omedvetenhet föreligger på många håll om hur subventioner på oväntat sätt kan förändra människors förväntningar, incitament och agerande.

Sammanfattande bedömning

Utvärderingen är ambitiöst upplagd, konsekvent genomförd och välskriven. Den ger ett intressant underlag men förbiser att ifrågasätta

grundläggande incitamentseffekter. Av de fyra angivna syftena kan det första anses ha uppfyllts, i vart fall i en snäv mening. Övriga har däremot inte uppnåtts på ett tillfredsställande sätt. Den övergripande slutsatsen saknar tillförlitlighet.

20 Utvärdering av den nya fältorganisationen – en första lägesrapport

Bedömning 1

Utvärderingens syfte. Bakgrunden är införandet av en ny fältorganisation inom Skolverket och rapportens syfte är att ge Skolverkets ledning underlag för att bedöma i vilken utsträckning de beslutade förändringarna genomförts och likaså hur väl den nya organisationen uppfyller de mål som verksamheten ställt upp för dess arbete.

Utvärderingens metod. Att verbalt ge en lägesbeskrivning av den första fasen i implementeringen av Skolverkets nya fältorganisation. Beskrivningen bygger enligt uppgift på ett omfattande bakgrundsmaterial i form av intervjuer, tidredovisningar, attitydundersökningar samt "beskrivningar och bedömningar av olika infallsvinklar".

Bedömning

Dokumentation. Verksamheten som utvärderas, de data som används och resultaten är ej beskrivna på ett sådant sätt att läsaren kan skapa sig en tillfredsställande bild av utvärderingens tillförlitlighet.

Data. Då de data som ligger till grund för utvärderingen endast i ringa utsträckning finns redovisade i utvärderingen, är det mycket svårt att bedöma huruvida de är tillräckliga eller tillförlitliga.

Metod. Utvärderingsmetoden är knappast lämplig, eftersom den är alltför journalistisk och impressionistisk. Utvärderarna har "plockat" aspekter ur bakgrundsmaterialet på ett så pass osystematiskt sätt att en utomstående näppeligen kan bedöma om en rättvisande bild ges.

Slutsatserna. En genomläsning av materialet förmedlar för en utomstående intrycket av att omorganisationen i hög grad varit ett misslyckande: få tecken på ökad kvalitet eller effektivitet, stor splittring i arbetsuppgifterna, i hög grad negativ bild av trivsel och arbetsklimat etc. Exempelvis skriver man att "fältgrupperna är inte bra på att lyfta fram och hantera konflikter samt att utnyttja den befintliga kompetensen. Det är också få som tycker att det finns en gemensam policy gentemot skolhuvudmän och allmänhet". Detta citat förmedlar rimligen en bild av att projektet varit föga framgångsrikt; på ren svenska står det ju att man drar i olika riktningar, har svårt att få användning av sina kunskaper och inte riktigt vet vad man ska göra. Ändå skriver utredarna i rapporten till verksledningen (1995-11-08) att "förändringen av organisationen i viktiga delar varit ganska framgångsrikt". Jag finner inte stöd för denna slutsats i utredningen.

Kostnadseffektivitet. Min bedömning är att man kunnat få ut mer för givna resurser om man valt att på ett mer objektiva sätt redovisa dataunderlaget i tabeller och med hjälp av gängse deskriptiva statistikor, och sedan kommenterat dessa uppgifter.

Objektivitet. Trots att det lyser igenom att den negativa informationen överväger i bakgrundsunderlaget får ej detta fullt genomslag i texten genom att man i flera fall refererar enskilda positiva bedömningar och inslag oavsett om dessa är representativa.

Sammanfattande bedömning. Rapporten är skriven på ett sådant sätt att den är mycket svår att tillgodogöra sig för en utomstående. Vad exakt sysslar fältgrupperna med? Vad är måluppfyllelse? Då det dessutom finns mycket få mätbara kriterier för output i den organisation som utvärderas blir det utifrån det underlag jag erhållit omöjligt att bilda sig en välgrundad uppfattning om hur väl verksamheten fungerat.

Bedömning 2

3 Syfte:

Syftet är att utvärdera den nya fältorganisationen, ej själva utfallet utan processen att få en fungerande fältorganisation.

4 Metod:

Man har valt dessa verktyg: intervju i fältgrupperna, analyser av tidsredovisningar och attitydundersökningar. Den senaste verkar vara något som återkommer med viss regelbundenhet inom verket.

5A Dokumentation:

Rapporten är grundlig. Man delar upp i olika delområden som redovisas, en i taget, med en sammanfattning i slutet. Bra man behöver ej leta efter texten som ska överensstämma med slutsatsen.

5B Data:

Man vet ej huruvida detta är ett urval eller en totalundersökning. Om det är ett urval, hur har man då gått till väga vid urvalsförandet? Står ej redovisat, därför är det svårt att veta om data är tillförlitligt.

5C Metod

Arbetet består som tidigare nämnts av olika delområden, vilket medför att man valt olika metoder för att få ett underlag: intervjuer, tidsredovisningar och attitydundersökningar. Detta gör att man får basfakta som verifieras med ytterligare information.

5D Slutsatser:

Slutsatserna är korrekt dragna av det redovisade materialet och väl underbyggda.

5E Kostnadseffektivitet:

Vem/vilka utförde intervjuerna? Tillkom resor? Svårt att bedöma om man kunde ha hanterat detta på ett annat, kanske billigare sätt. De övriga metoderna förutsätter jag att de sköts m h a datorer och sedvanlig enkätutskick som ej genererar stora kostnader.

5F Objektivitet:

Utvärderaren/utvärderarna har godtagbar objektivitet.

6 Trevlig rapport! Läsvänlig, snygg layout. Saknar trots allt en tabell där tidsredovisningen är sammanställd.

21 Den nationella utvärderingen av grundskolan 1987–1994 – Synpunkter och empiri om intern debatt och lokal användning (Skolverket)

Bedömning 1

3. Utvärderingens syfte

Jag har tolkat min uppgift som att jag skulle bedöma en myndighets (Skolverkets) egen utvärdering av en egen verksamhet.

Jag har fått i uppgift att bedöma en skrift som heter "Den nationella utvärderingen av grundskolan 1987–1994" bestående av tre delar. Av dessa tre har jag läst den första och den andra delen. Den första delen med titeln "Debatt om NU – Design, validitet, användbarhet" är skriven av Lander och Torper vilka är pedagoger verksamma vid Göteborgs respektive Lunds universitet. Den kommer närmast att vara en utvärdering av Skolverkets egen verksamhet. Den andra delen heter "Nytta med NU" och är skriven av Hill, Lander och Thång. Den är i första hand en studie av hur skolor använt sig av den återföring som gjorts av resultaten av den nationella utvärderingen. Den kommenteras därför inte mer här.

Syftet för studien sägs vara att utvärdera Skolverkets styrning och genomförande av det nationella utvärderingsprojektets undersökning 1992, samt verkets roll som övergripande planerings- och samordningsinstans. Skolverket sägs ha betonat följande frågor: Har resurserna för planering och organisation varit tillräckliga? Vilka har varit genomförandestrategins och informationsinsatsernas positiva och negativa sidor? (sid 3) Det saknas dokumentation som beskriver Skolverkets uppdrag till författarna.

Stämmer texten och sammanfattningen med beskrivningen av projektets syfte? När man kommer till kapitel 2 om den nationella utvärderingens (NUs) historia dyker ett nytt syfte upp. Nämligen att återkopp-a resultaten till den lokala skolan för att stimulera utvecklingen (sid 11). Detta syfte har när vi kommer till kapitel 5 blivit ett huvudsyfte.

Min tolkning är att studiens rubrik stämmer bättre med studiens verkliga syfte. Syftet verkar nämligen vara att ge ett referat av den vetenskapliga debatt som den nationella utvärderingen varit föremål för och att ge en bild av de för- och nackdelar som den valda ansatsen har.

Syftet har därför inte i första hand varit att bedöma Skolverkets sätt att genomföra den nationella utvärderingen. Det har därför varit oerhört frustrerande att försöka läsa denna text som en utvärdering av genomförandet. Det har gett mig en helt ny upplevelseförvirring. Kan Skolverket verkligen mena att detta är en utvärdering av Skolverkets egen verksamhet? Varför begriper jag inte denna text har jag frågat mig? Skäl är nog att mina förväntningar på en text som är en utvärdering av en myndighets egen verksamhet inte alls stämde med vad jag läst.

Till att börja har texten formella brister. Det saknas en sammanfattning genom vilken man kan orientera sig om studiens problemställning och huvudresultat. Ett annat problem är att studien inte avgränsat sina bedömningar till ett fåtal väl preciserade aspekter vilka skulle kunna vara påverkbara för Skolverket. Syftet stämmer inte med innehållet. Några av frågeställningarna i det påstådda syftet kommenteras knappt alls.

Texten är rörig. Mål, åtgärder och resultat kommenteras huller om buller. Jag har inte kunnat finna en samlad presentation av författarnas slutsatser beträffande syftet. Hade NU-projektet kunnat styras och genomföras på ett mer ändamålsenligt sätt? Har författarna kunnat dra några lärdomar om planering och samordning? Var resurserna för planering och organisation tillräckliga? Vissa slutsatser finner jag i kapitel 5 Lokal användning av den nationella utvärderingen under rubriken En adekvat strategi? på sid 59–60.

4. Utvärderingens metod

I kapitel 1 beskrivs studiens metod. Den består av läsning och intervjuer. Där definieras även utvärdering. Författarna definierar utvärdering som "en granskning av ett fenomen med syfte att fördjupa kunskaper man har eller att skapa nya kunskaper" (sid 5). De framhåller att uppföljning är en löpande insamling av data som syftar till att kontrollera om resultaten av en verksamhet håller sig inom förväntade gränser. De hävdar att kontroll inte är det primära syftet för utvärderingen.

Riksrevisionsverkets definition av utvärdering för budgetprocessen är att "bedöma och förklara prestationer och händelseförlopp i en myndighetsverksamhet". Utvärderingen "klargör och skärskådar myndighetens prestationer och/eller prestationernas effekter i samhället" (Resultat, verksamhet, ekonomi en handledning för myndigheterna, Riksrevisionsverket 1994). En utvärdering kan sägas syfta till att bedöma om en vald åtgärd är ändamålsenlig och kostnadseffektiv.

För att bedöma om en åtgärd är ändamålsenlig behövs en beskrivning av statsmakternas mål för aktiviteten. Därför bör en tolkning av målen presenteras. Är målen realistiska? Kan man precisera dem ytterligare? Hur skulle en alternativ målbeskrivning ha kunnat se ut?

Vidare behövs en beskrivning av tänkbara alternativ till den valda åtgärden. Ofta kan det räcka att beskriva små variationer för att visa hur förbättringar skulle kunna åstadkommas.

I kapitel 2 beskrivs den nationella utvärderingens (NUs) tillkomst-historia från 60-talet och framåt. Ett syfte är att tolka målen för utvärderingen. Ett ursprungligt mål är att skapa goda betingelser för in-lärning. Som ett medel för att bedöma vad sådana betingelser är formuleras det underordnade målet att mäta elevers kunskaper, färdighe-ter och attityder. Detta underordnade mål kom senare att ges en större tyngd.

Författarna visar att glidningen mellan dessa mål skapar ambivalens i utformningen av den nationella utvärderingen. Det skapar också en oklarhet med vilken norm det faktiskt genomförda utvärderingsarbetet skall bedömas.

Utvärderaren måste dock göra ett val. Ett tänkbart val vore att välja mellan två normer. Låt mig kalla dem Norm 1 och Norm 2.

Norm 1. Om syftet är att följa upp eleverna i enskilda skolor när läroplanens mål så bör data från ett representativt urval av eleverna i alla skolor samlas in.

Norm 2. Om syftet är att bedriva långsiktig forskning om kausalsambanden i inläring, bör data från ett representativt urval av elever, om både elevers resultat, sociala bakgrund samt skolans egenskaper samlas in för samma individer.

När Skolverket genomförde sin första undersökning inom NU-projektet 1992, vilken är föremålet för utvärderingen, var målet enligt författarna (sid 14) att:

- Bedöma kvaliteten i grundskolan i förhållande till läroplanens mål.
- Bidra till förståelse för samband mellan sociala bakgrundsfaktorer, skolfaktorer och elevernas utveckling.
- Utveckla modeller för utvärdering
- Stimulera ämnesdidaktisk forskning
- Bygga upp ett nätverk mellan pedagogiska forskare

För att bedöma dessa mål skulle man kunna formulera alternativa åtgärder för att nå målen var för sig. Vidare skulle man kunna diskutera måluppfyllelse och kostnader för den valda åtgärden. Om möjligt bör man dessutom bedöma om motsvarande eller bättre resultat hade kunnat nås till lägre kostnader. I vilken mån görs det?

Låt mig kortfattat presentera de slutsatser jag kunnat finna till de respektive rubrikerna:

Bedöma kvaliteten i grundskolan i förhållande till läroplanens mål

Skolverkets mål bör rimligen vara att genomföra denna målsättning med tillfredsställande kvalitet och till så låga kostnader som möjligt. Om Skolverket har gjort detta eller ej kräver en beskrivning av alternativa tillvägagångssätt för varje mål och motsvarande åtgärd. Det saknas en genomgång av tänkbara alternativa uppläggningar för att bedöma den hårda kärnan av mätningar av elevernas kunskaper, färdig-

heter och attityder. Alternativen skulle beträffande Skolverkets genomförande ha kunnat bestå av den nationella utvärderingen, kunnat bestå av olika urval av ämnen som granskas, antal skolor som granskas m.m.

Författarna drar slutsatsen att "Såvitt vi begriper har NU-projektet alls inga möjligheter att mäta skolors relativa effektivitet i elevprestationer" sid 65. Skälet verkar vara att undersökningens resultat inte kan användas på detta sätt. Referenser eller hänvisningar till var dessa bedömningar görs saknas i anslutning till denna slutsats.

Inga slutsatser dras om kostnaderna för att mäta kvaliteten i grundskolan mätts på ett kostnadseffektivt sätt.

Bidra till förståelse för samband mellan sociala bakgrundsfaktorer, skolfaktorer och elevernas utveckling

På sid 19 och 35 refereras en invändning mot NU-ansatsen som kan tolkas om att teoretiska utgångspunkter för arbetet saknats. Jag har dock inte kunnat finna författarnas egna bedömningar eller slutsatser.

Utveckla modeller för utvärdering

För vissa ämnen har nya testmetoder utvecklats inom ramen för arbetet med den nationella utvärderingen (sid 15 och 25).

Stimulera ämnesdidaktisk forskning

En del av forskarna ökade sin kompetens. De lärde sig ett statistikprogram, SPSS. Andra lärde sig avancerade metoder, t.ex. lisrel (sid 49).

Bygga upp ett nätverk mellan pedagogiska forskare

Projektet anses ha bidragit till att det avsedda nätverket bildats för forskare (sid 49). I övrigt kan man knappast tala om att det uppstått nå-

got nätverk (sid 60). Konferenserna har varit stimulerande och intressanta (sid 49).

Låt mig nu kommentera frågorna i syftet.

Skolverkets styrning och genomförande av det nationella utvärderingsprojektet samt verkets roll som övergripande planerings- och samordningsinstans

Två huvudinvändningar mot skolverkets planering och samordning förs fram. Den första huvudinvändningen är att man arbetat under tidspress. Författarna hävdar att rapporteringscyklerna var för korta sid 65. Argument för denna tes saknas dock. Kan inte kortare tider delvis kompenseras med större resurstilldelning eller lägre ambitioner?

Den andra huvudinvändningen är att man konstaterar att samordningen brustit (sid 56). Det avser främst återkopplingsarbetet gentemot skolorna.

En tredje invändning går ut på att man introducerar testmetoder i full skala innan de var tillräckligt utvecklade som prototyper. Författarna konstaterar utan något egentligt argument att "NU-projektet synes vara ett sådant fall" sid 41.

Jag saknar en diskussion av en ändamålsenlig avgränsning av forskning och administration. Är det lämpligt att låta ett statligt verk administrera forskning och resultatmätning i ett och samma projekt?

Författarna konstaterar i kapitel 6 att "det är svårt att säga hur styrningen av ett projekt av NU:s karaktär skall se ut för att vara ändamålsenlig" sid 63. Det är också svårt att "definiera hur riksdag och regering skall styra nationell utvärdering" sid 67. På sidan 68 antyder författarna att Skolverket inte styrt projektet "verket formulerar projektets allmänna syften och sedan är dess uppgift att på olika sätt bistå forskarna utan att lägga sig i hur de löser uppgiften. Genom en sådan hållning definieras NU-projektet som ett traditionellt forskningsprojekt."

Det saknas argument för varför man bedriver utveckling av mätinstrumenten i samma projekt som man utför uppgiften att faktiskt följa upp vilka resultat som nås i skolorna. Ett tänkbart alternativ vore att genomföra utvecklingsarbetet separat.

Med undantag för Skolverkets betoning av återkopplingens betydelse för den lokala utvecklingen har Skolverket upplevt sin egen styrning som luddig (sid 68). Skolverkets antyds ha använt en olämplig upphandlingsform och låtit delprojekten arbeta för oberoende av varandra. På (sid 70) föreslås att en dubblering av forskargrupper bör användas för att undvika för stark auktoritetstro(?). Vidare borde Skolverket(?) utlyst en tävlan om övergripande principer för utformningen av projektet.

Skolverket verkar ha skickat ut "jämförelser av rådata" som inte ger "möjligheter att dra slutsatser om en enskild skolas prestationer" (sid 65). Författarna bedömer att återrapporteringen till skolorna i former som inbjöd till jämförelser mellan skolor borde undvikits (sid 65). Genom att Skolverket inte ville avstå från den uppmärksamhet som möjligheten att jämföra sig med andra skolor skulle ge resultat (sid 53).

Har resurserna för planering och organisation varit tillräckliga?

Strategin kan sägas ha innehållit en rad komponenter, t.ex. en satsning på pilotskolor, uppmuntran till kommuner och skolor att själva informera om utvärderingen m.m. Författarna verkar dock tycka att en samordnad strategi för att stimulera återkoppling saknats (sid 56).

Det påstås att en rapportering vart tredje år leder till stark tidspress (sid 64), att kravet på snabb återföring tog ganska mycket tid vilket ledde till att möjligheterna att analysera materialet begränsades.

Vilka har varit genomförandestrategins och informationsinsatsernas positiva och negativa sidor?

Skolverket har negligerat kommunernas potentiella roll i strategin för återkoppling (sid 66). Kommunerna har haft stora svårigheter med bl.a. besparingar under den studerade perioden.

I kapitel 3 diskuteras debatten kring uppläggnings av NU-projektet. Normen i detta avsnitt är främst i vilken mån projektledningen exponerat sig för synpunkter och debatt. Där diskuteras främst om den val-

da strategin med avseende på forskning med den andra strecksatsen ovan som syfte.

Slutsatsavsnittet konstaterar att det förekommit intern och extern diskussion av projektet. Det verkar som om författarna anser att den återföring som skett till merparten av de berörda lärarna och skolorna inte på ett enkelt sätt kan användas för att bedöma den enskilde lärarens eller skolans prestationer.

6. Sammanfattande bedömning

Jag vill inte bedöma skriftens kvaliteter som en vetenskaplig översiktsuppsats i pedagogik. Studien är dock inte en utvärdering av Skolverkets verksamhet i Riksrevisionsverkets avseende av följande skäl:

- 1) Målen för utvärderingen och Skolverkets genomförande preciseras och tolkas inte separat.
- 2) Alternativa tillvägagångssätt diskuteras inte systematiskt.
- 3) Bedömningar av relativ måluppfyllelse för olika tillvägagångssätt görs ej av författarna.

Jag har därför inte heller kommenterat den andra delen som utvärderar i vilken utsträckning den nationella utvärderingens återkoppling till skolorna givit upphov till ett utvecklingsarbete på skolorna, eftersom studien inte relaterar detta till Skolverkets sätt att genomföra arbetet.

Bedömning 2

3. Utvärderingen syftar till att utvärdera det nationella utvärderingsprojektets (NU-projektets) styrning och genomförande samt Skolverkets roll som övergripande planerings- och samordningsinsats.

Inom det ovannämnda har två speciella frågor prioriterats:

- Har resurserna för planering och organisation varit tillräckliga?
- De negativa och positiva sidorna med genomförandestrategin respektive informationsinsatserna.

4. För det första är det frågan om tre olika separata utvärderingar. Den första delutvärderingen handlar om den nationella utvärderingens design, validitet och användbarhet (Debatt om NU). Den andra delutvärderingen handlar om den lokala användningen av det nationella utvärderingsprojektet för grundskolans högstadium 1992–1994, och slutligen den sista delrapporten är en utvärdering om hur de deltagande skolorna fick sina första resultat tillbaka från det Nationella utvärderingsprojektet 1993. Var och en av dessa delutvärderingar tillämpar en egen metod.

Del I:

Utvärderingens metod består av läsning av projektets skrifter och viss debatlitteratur om projektet, samt deltagande i två av projektets konferenser. Intervjuer med ledarna för de forskargrupper, som konstruerat instrument samt informella samtal med ledningen för forsknings- och utvärderingsprogrammen på Skolverket.

Del II:

Telefonintervjuer, besök med intervjuer av rektor, den lokala samordnaren för NU-projektet, en ämnesföreträdare för no-blocket samt en företrädare för annat ämne som rektorn valt ut.

Lärarenkäter i samband med en personalkonferens, intervjuer med lärargrupper utvalda av skolledningen. Intervjuer av skolledningen och den lokale samordnaren.

Del III:

Enkäter, varians- och regressionsanalyser i samband med analyserna.

5. Bedömning av utvärderingen

Allmänna kommentarer (efterfrågas egentligen under punkt 6):

Målgruppen för utvärderingens resultat? Är den endast avsedd för internt bruk inom Skolverket? Rapporten är inte särskilt lättläst. Beskrivningen av tillvägagångssättet vid utvärdering och återkopplingen av resultaten m.m. hade kunnat förkortas ned betydligt. En kronologisk framställning av de olika rapporterna hade också underlättat inläsningen. Som framställningen är nu hoppar man fram och tillbaka (tidmässigt) i texten, vilket försvårar inläsningen.

A) Dokumentation

Dokumentationen är omfattande. Rapporten består av de tre delarna; 1. Debatt om NU, 2. Nyttan med NU och 3. Återkopplingens pedagogik. Bl.a. i den första delrapporten beskrivs verksamheten som utvärderas, de tidigare utvärderingarna, datainsamlingen och -metoder ingående. Som konstateras i de allmänna kommentarer hade framställningen t.o.m kunnat förkortas en del för att läsaren skulle kunna bibehålla intresset för texten. Resultaten för del III (Återkopplingens pedagogik redovisas delvis i bilagor. Förklaringarna till tabellerna saknas dock, och det kan vara svårt för den som inte är väl insatt i statistiska metoder att förstå vad man egentligen vill förklara i de olika tabellerna. Analyserna av de olika resultaten ges ett alltför litet utrymme.

B) Läsning av projektets skrifter och viss debatlitteratur om det samt deltagande i två av projektets konferenser, har antagligen lett till beskrivningar i avsnitten i den nationella utvärderingen – en kort historia (kap. 2. Del I) och till avsnittet om validitetsdiskussion (kap. 3. del I).

Utkast till en skrift, som beskriver tillkomsten av NU-programmet har gett information om NU-projektets tillkomst. Data av det slag som erhållits genom dessa metoder (litteraturstudier) är tillräckligt lämpliga och tillförlitliga för de ändamål som de är avsedda för.

Data har erhållits framför allt genom intervjuer. 19 intervjuer har genomförts med ledarna för de forskargrupper som konstruerat instrument. Tidpunkten för dessa intervjuer framgår inte av texten, inte heller innehållet i intervjuerna, vilket gör det svårt att bedöma kvaliteten på dessa data. Senare på sidan 12 (del I) talas det om omsättningen av personal, som arbetat med den nationella utvärderingen, vilket innebär att i vissa fall har de personer som i början bistått utvärderingen med data i form av intervjuer, bytts ut. Hur detta kan påverka utvärderingen kan diskuteras. Intervjutiden anges till 1–2 timmar och -formen till fria samtal. Det är alltid mycket svårt att uttala sig om kvaliteten på intervjuer som har hållits som fria samtal. Intervjuarens egenskaper och erfarenhet spelar en stor roll i dessa sammanhang.

Även informella samtal med ledningen för forsknings- och utvärderingsprogrammen på Skolverket, har genomförts.

På sidan 12 (del I) kan man läsa följande:

"Sammanfattningsvis kan man säga att den nationella utvärderingen inte utformats under gynnsamma förhållanden. Osäkerheten i verkets roll i det framtida styrsystemet gör att det är svårt att utforma ett detaljerat program för den nationella utvärderingen Samtidigt sker det successivt ett utbyte av tongivande personer, vilket påverkar möjligheterna att hålla samman och koordinera projektgruppens arbete. " Citatet belyser problematiken runt igångsättandet av ett projekt av denna omfattning. Man kan också fråga sig om det är just Skolverket som central myndighet som bör koordinera ett så gigantiskt arbete, utan skulle det kunnat vara bättre att t.ex. geografiskt dela landet in i olika regioner som arbetar mera lokalt med utvärderingen av skolan.

När det gäller de övriga delutvärderingarna så är ju både telefonintervjuer och enkätsintervjuer bland de vanligaste datainsamlingsmetoderna. För att kunna bedöma kvaliteten på dessa data hade det varit nödvändigt att se vilka frågor som har ställts vid de olika intervjutillfällena. (Intervjumallarna hade kunnat bifogas till rapporterna i form av bilagor).

C. Litteraturstudier, intervjuer, enkätsintervjuer, vid analyser har bl.a. variansanalyser och multipla regressionsanalyser använts.

På sidan 5 (delrapport I) kan man läsa följande;

"Vi har inte gjort intervjuer bland Skolverkets undervisningsråd, som deltog i NU genom att utföra s.k. lägesbedömningar av varje deltagande skola samt genom att delta i återkopplingen av resultaten till dessa skolor."

Det finns inga förklaringar till varför man inte har genomfört dessa intervjuer. Mycket viktig information har säkerligen gått förlorad i och med att undervisningsråden inte intervjuats. Speciellt då återkopplingen av resultaten till lokal nivå verkar ha varit ett problem.

D. I del II (Nytta med NU) har analyser gjorts bl.a. efter nivåer. Det material som har samlats in har analyserats på individnivå (rektor), gruppnivå (arbetsenheterna, ämnesgrupperna), organisationsnivå (kollegiet eller skolan som helhet) samt samhällsnivå (kommunernas och Skolverkets engagemang). De slutsatser som finns i dessa analyser är intressant läsning. Det märkligaste är att kommunerna har haft en sådan liten roll i hela NU-projektet. Det konstateras på sidan 105 (del II) att "problemet ligger, enligt vår mening i att Skolverkets strategi tog för lite hänsyn till den kommunala medverkan i användningen av NU överhuvudtaget." och vidare på samma sida "oavsett vad man anser om Skolverkets strategi för att involvera kommunerna kan man tycka att dessa haft dålig ork att själva ta initiativ för att lära sig något substantiellt om utvärderings- och uppföljningsinstrument och design runt dem."

Just kommunerna borde ha intresse av att ta del av NU-utvärderingarnas resultat. Vilka ansträngningar har man över huvud taget gjort för att engagera kommunerna (de berörda förvaltningarna, politikerna m.m.) i arbetet med NU-projektet.

I den sista delrapporten har man velat förklara återkopplingens effekt i sig på skolorna med hjälp av multipla regressionsanalyser. I texten konstateras följande. "Vi har lyckats konstruera enkäten så, att vi med multipla regressionsanalyser lyckats förklara mellan 26–28 procent av variationen av de mått vi använt på utfall." Vidare skriver man följande "Om vi hade packat våra modeller med många oberoende variabler, eller nöjt med mycket låga svarsfrekvenser, kunde vi komma upp i 35 procent, men flera variabler har ett så obetydligt förklaringsvärde, att vi måste nöja oss med att lyckas förklara strax under 30 procent." Vidare fortsätter man genom att konstatera att inom den samhällsvetenskapliga analysen är man mycket nöjd om man

lyckas komma upp till 50–60 förklarad varians och om man kommer högre bör man snarast bli misstänksam mot sin modell (sid 188, del III). Varför har man inte försökt fånga in flera oberoende variabler i de olika modellerna? Förklaringsvärdet på 26–28 procent är väldigt lågt även inom de samhällsvetenskapliga analyserna. Att man vid högre förklaringsvärden borde bli misstänksam mot sin modell stämmer inte. Det är snarare så att konstruktionen av de olika enkäterna vid datainsamlingen varit sådan att det har varit omöjligt att konstruera flera oberoende variabler.

Det har visat sig att återkopplingens kvalitet och ledningens roll har varit de faktorer som har mest förklarat den beroende variabeln. (Y = konstruktiv debatt, Y = all påverkan) finns det resultat för i framställningen. Detta är intressanta resultat. Skolledningen kan med sitt agerande se till att en konstruktiv debatt överhuvudtaget uppstår. För all påverkan var ledningens roll blygsammare, däremot var det i första hand återkopplingskvalitet som var avgörande. Hur man har mätt återkopplingens kvalitet framgår inte av texten.

E. Vad menas med kostnadseffektivitet i detta sammanhang? kan man fråga sig. Det finns inga uppgifter om kostnaderna för de olika delutvärderingarna. Inte heller några uppgifter om den exakta tidsåtgången för sammanställningen av dessa rapporter. Uppgifter om personal saknas o.s.v. Detta gör det svårt att uttala sig om projektens kostnadseffektivitet. Det som nämns är att studerande vid olika högskolor har medverkat t.ex. i den sista delundersökningens datainsamling, -bearbetning och rapportering. Även i samband med den andra delrapporten har studerande medverkat i insamlingen av data, vilket säkerligen har inneburit stora kostnadsbesparingar för projektet.

F. När det gäller urvalsförfarandet vid den andra delstudien, blir man lite fundersam på urvalsförfarandet. På sidan 147 (Del II) kan man läsa följande; "Vårt planerade urval omfattade alla skolorna i sydvästra, sydöstra och södra Sverige." Hade det inte varit mera objektivt att genomföra ett slumpmässigt urval av alla de 101 högstadieskolor som deltog i den Nationella utvärderingen. Som det är nu verkar det att man har valt de skolor som geografiskt inte ligger alltför långt ifrån just Göteborg.

För övrigt har de tre utvärderarna varit objektiva i sin beskrivning av tillvägagångssätt, val av frågeställning och tolkning av resultaten. Ibland saknar man t.o.m författarnas egna reflektioner på t.ex. resultaten.

6. Se också punkten 5!

Sammanfattningsvis kan konstateras att tiden för inläsningen för de olika rapporterna har varit alltför kort (knappt en månad).

22 Pilotstudie av Lex Maria-anmälningar och beslut (Socialstyrelsen)

Bedömning 1

3 Syfte:

Syftet med undersökningen var att utvärdera Socialstyrelsens agerande vid Lex Maria-ärenden och hur anmälaren upplever detsamma.

4 Metod:

Man har valt att intervjua ett antal personer under ett så kallat "fritt samtal" som avslutades med att den som intervjuar och den som blev intervjuad tillsammans fyllde i ett frågeformulär.

5A Dokumentation:

Det finns inga tidigare utvärderingar att jämföra med, eftersom detta är en pilotstudie. Dokumentet i sig ger läsaren en god uppfattning om tillförlitligheten och innehållets substans.

5B Data:

Materialet är för litet för att vara statistiskt tillförlitligt. 20 utvalda ärenden av ca 1500 ärenden/år är ej representativt. Vikten bör läggas på att detta är en validering av frågeformuläret för fortsatta undersökningar men med bättre urval.

5C Metod:

Intervju är alltid ett svårt utvärderingsverktyg. Ställer stora krav på personen i fråga som ställer frågorna att vara neutral, att ej ställa frågor på ett ledande sätt. Allt för att det slutgiltiga materialet ska vara jämförbart över lag. Med andra ord det kräver utbildning av intervjuaren.

5D Slutsatser:

Av det som återges i redovisningen av frågeformuläret så är det korrekt dragna slutsatser. Man har även varit måna om att framhålla avvikande synpunkter i sammanfattningen.

5E Kostnadseffektivitet:

Intervju kostar alltid mer pengar än om man sänder ut en enkät/frågeformulär. Kostnader som uppstår är: resekostnad, intervjuarens tid och den intervjuades tid. Lösningen i detta fall hade varit att skicka ut enkät vilket skulle ha medfört att man kunde ha fått in fler ärenden att bedöma. Man når fler på kortare tid och kostnaden blir lägre.

5F Objektivitet:

Utvärderaren har godtagbar objektivitet.

6. Rapporten är lite torftig. Man radar upp siffror i texten vilket medför att det blir svårläst. Dessa siffror skulle man kunna redovisa i tabellform. Detta medför att texten blir mer flytande och lättläst. Det är bra att redovisa citat men i denna rapport är det för många citat. Detta kan upplevas som en brist hos författaren, att denne ej kommer på en text själv.

Bedömning 2

3. Utvärderingens syfte

Utvärderingen är en pilotstudie i syfte att ge underlag "för konstruktion av en enkät avsedd för en mer omfattande undersökning och värdering av Lex Maria-verksamheten".

Studien handlar om "en utvärdering av effekten av utredningar som Socialstyrelsen, SoS, utfört och beslut som fattas efter anmälningar enligt Lex Maria." Genom personliga intervjuer skulle studien svara "på frågor om varför och när i förhållande till händelsen Lex Maria anmälan gjordes, vilka åtgärder som vidtogs mellan händelsen och anmälan, efter denna men före beslutet samt efter detta. Hur stor betydelse hade den av SoS utförda utredningen? Vad spelade SoS besök på platsen för roll och hur uppfattades det av de berörda befattningshavarna?"

4. Utvärderingens metod

Ett urval av 20 Lex Maria-ärenden gjordes genom att SoS regionala enheter tillskrevs med begäran om kopior av vardera ca 10 ärenden som avgjorts under 1993 (totalt avgjordes ca 1 500 ärenden). Av totalt 55 inkomna kopior valdes 21 ut (varav en senare föll bort därför att ingen av de inblandade längre fanns kvar på den aktuella kliniken) "som varande intressanta med hänsyn till att utredningen påvisat brister i fråga om organisation, föreskrifter och/eller rutiner som viktigaste orsak till den anmälda händelsen." "..... fall där

händelsen främst kunde tillskrivas felhandling utförd av enskild befattningshavare" undveks.

Intervjuerna utfördes som "fritt samtal där vissa huvudpunkter berördes". Samtalet avslutades med att undersökaren tog fram ett intervjuformulär med 27 frågor för ifyllande tillsammans med den eller de intervjuade (en till tre personer intervjuades samtidigt).

"De utvalda intervjuobjekten tillskrevs av Lars Steen (SoS, min anm.) med förfrågan om de ville och hade möjlighet att medverka till den planerade intervjun som kunde beräknas ta en till två timmar i anspråk. I brevet framhölls att jag ej var anställd hos SoS utan tillfälligt engagerad i det aktuella projektet."

Utredaren skriver själv att: "Planeringen och genomförandet av pilotstudien kunde på grund av snäva tidsmarginaler och därav förorsakad skyndsamhet förutses ha vissa brister. Dessa farhågor har besannats särskilt vad gäller urvalet av ärenden. Resultaten måste därför granskas kritiskt och värderas med stor försiktighet."

Samtidigt uttrycks dock slutsatsen: "Resultaten av intervjuerna är påfallande homogent och bör därför tillerkännas viss värde. En fortsatt, större och mer representativ undersökning i anslutning till pilotstudien skulle sannolikt ge ett liknande utslag. Utredaren rekommenderar SoS att avvakta något eller några år med en ny undersökning av effekterna av Lex Maria-utredningar och annat kvalitetssäkringsarbete."

A) Dokumentation

Det är svårt att skapa sig en rimlig bild av utvärderingens tillförlitlighet med hjälp av den redovisade dokumentationen. En enda hänvisning finns till tidigare utvärderingar och den uttrycks enligt följande:

"Ju längre tiden har gått, dess svagare synes kritiken mot SoS verksamhet ha blivit. *Den nyligen publicerade rapporten från riksdagens revisorer utgör i detta hänseende ett anmärkningsvärt undantag* (min kursivering). Enligt min (utredarens, min anm.) uppfattning bygger mycket av rapportens innehåll på felaktigt faktaunderlag eller missförstånd. Den visar emellertid på behovet för SoS del av välgrundade argument för en fortsättning och utökning av tillsynsverksamheten och

kvalitetssäkringsarbetet. Handläggningen av Lex Maria-ärenden är en viktig del av denna verksamhet."

Vad som kritiserats i rapporten från riksdagens revisorer framgår inte och någon ytterligare diskussion i frågan finns inte.

Eftersom inga uppgifter lämnas om hur de årligen ca 1500 Lex Maria-ärendena fördelar sig på olika typer av anmälningar och beslut går det inte att bedöma representativiteten i urvalet av 20 ärenden. Dock framgår enligt ovan (se "Utvärderingens metod") att endast en viss sorts ärenden valts ut, nämligen de som resulterat i åtgärder i organisation och rutiner m.m.

5B) Data

De data som ligger till grund för utvärderingen känns inte tillräckligt tillförlitliga för de slutsatser som dras och de rekommendationer som ges (se också "Utvärderingens metod" ovan och "Slutsatser" nedan).

Dock ligger det i sakens natur att en pilotstudie som syftar till att ta fram underlag för en fördjupad studie i sig inte kan omfatta alla önskvärda data för att besvara den ev. kommande huvudstudiens frågor.

5C) Metod

Metoden med personliga intervjuer bedömer jag som lämplig. Där emot kan såväl urvalet av ärenden (se ovan under "Dokumentation"), som urvalet av respondenter ifrågasättas. Endast huvudansvariga chefsöverläkare (16), överläkare (10), chefsläkare (2), sjukhemschef (1), medicinskt ansvariga sjuksköterskor (3), vårdavdelningsföreståndare (4) och klinikföreståndare/vårdenhetschefer (2) intervjuades.

Utredaren "avstod också medvetet från samtal med enskilda befattningshavare som ställts till ansvar för felhandling eller i annat avseende blivit kritiserade i SoS utredning." Berörda patienter och anhöriga tillfrågades inte alls.

5D) Slutsatser

De slutsatser som dras beträffande det sannolika utfallet av en större studie samt beträffande rekommendationen att vänta några år med en sådan studie känns inte tillräckligt väl underbyggda:

"En fortsatt, större och mer representativ undersökning i anslutning till pilotstudien skulle sannolikt ge ett liknande utslag. Utredaren rekommenderar SoS att avvakta något eller några år med en ny undersökning av effekterna av Lex Maria-utredningar och annat kvalitets-säkringsarbete."

Det kan inte uteslutas att ett annat urval ärenden och andra responder hade givit en åtminstone delvis annan bild och andra slutsatser.

5E) Kostnadseffektivitet

Såvitt jag kan bedöma hade det knappast varit möjligt att göra en väsentligt mycket bättre utvärdering med samma förutsättningar beträffande resurser och tid.

Utredaren skriver att: "Tidsramen för denna pilotstudie medgav inte ens försök till ett från statistik synpunkt representativt urval av ärenden som avgjorts under t.ex. det sista eller de två senaste åren". "Intervjuerna kunde endast genomföras fr o m . sista veckan i maj t o m första veckan i juli. Semestrar försvårade planläggningen.

Varför utvärderingen måste bedrivas med sådan brådska och under semestertid förklaras inte. Det förefaller märkligt mot bakgrund av att Lex Maria tillkom 1936 och att rekommendationen är att vänta med huvudstudien några år framåt i tiden.

En mer tillförlitlig utvärdering hade alldeles uppenbart krävt en längre tid för urval och informationsinsamling.

5F) Objektivitet

Tillvägagångssättet (bl a SoS's tidplan) och utredarens tolkning av resultaten känns inte övertygande för utvärderingens objektivitet. Utvärderaren understryker gång på gång de positiva omdömen som ges om SoS arbete och nya inriktning. Den lilla kritik som framskymtar mot

bristande erfarenhet och kompetens hos några av SoS inspektörer dränks bland lovorden. Jag kan inte undgå misstanken att det kan finnas lojalitets- och taktikfällor i den valda metoden. Kan det finnas sådana förklaringar till att svarsutfallet blev så homogent trots det lilla urvalet? Det gäller i synnerhet som endast ärenden som lett till organisatoriska åtgärder etc undersöktes.

Sammanfattande bedömning

Helhetsomdömet är att jag som utomstående läsare av Lex Maria-utvärderingen inte känner mig övertygad om att felbehandlingar inom sjukvården hanteras på ett tillfredsställande sätt. Utvärderingen tolkar t ex inte rimligheten i de angivna tiderna mellan fel och anmälan respektive mellan anmälan och beslut. Detta trots att sådan kartläggning och analys var en huvudfråga för utredningen.

Utvärderingen utgår ensidigt från det professionella perspektivet och bortser helt från de krav, önskemål och perspektiv som kan finnas hos de 1500 felbehandlade patienterna.

Som pilotstudie för att tjäna som underlag för en huvudstudie kan utvärderingen möjligen räcka till. Men när slutsatsen blir att det inte behövs någon huvudstudie förrän om några år känns studien inte tillfredsställande.

23 Rapport till Socialstyrelsen (Socialstyrelsen)

Bedömning 1¹

1 Syfte med undersökningen

Syftet med undersökningen var att undersöka om

- chefsöverläkare
- medicinskt ansvariga sjuksköterskor
- vårdhemsläkare
- tandläkare med ledningsansvar inom folktandvården
- legitimerade privatpraktiker
- yrkesverksamma legitimerade läkare
- tandläkare
- sjukgymnaster
- psykologer som inte är privatpraktiserande

hade:

1. Sett eller hört talas om socialstyrelsens föreskrifter om kvalitetssäkring inom hälso- och sjukvården inklusive tandvård.
2. Om man sett den folder som socialstyrelsen givit ut som heter "Att säkra kvalitet i hälso- och sjukvården".

¹Anm. Den ursprungliga bedömningen innehåller också en innehållsförteckning.

3. Om man fått muntlig information om föreskrifterna.
4. Om man har informerat sin personal om kvalitetssäkring.
5. Om man har kvalitetssäkring på arbetsplatsen.
6. Orsaken till att man startat med kvalitetssäkring.
7. Vad man gör för att kvalitetssäkra.
8. Betygsättning av föreskrifterna.

2. Metod

Utvärderingen är en kombination av totalundersökning och urvalsundersökning. Ett totalurval har genomförts bland följande grupper:

- medicinskt ansvariga sjuksköterskor
- vårdhemsläkare
- tandläkare med ledningsansvar inom folktandvården

Bland övriga yrkeskategorier har ett slumpmässigt urval gjorts. Både totalurvalet och det slumpmässiga urvalet har genomförts med hjälp av socialstyrelsens register. Data till undersökningen har samlats in via telefonintervjuer från SKOPS intervjucentral i Stockholm. Telefonintervjuerna genomfördes under perioden 12–29 juni och 1–24 augusti 1995. Urvalet uppgick till att börja med till 3.895 personer. 422 personer uteslöts på grund av att de förekom flera gånger eller inte tillhörde undersökningspopulationen. Det slutliga urvalet blev 3.473 personer. Antalet intervjuer blev 2.908 stycken, vilket ger en svarsfrekvens på 84 procent. (*Räknat på 3 895 personer blir svarsfrekvensen 75 procent*).

3. Kommentarer till undersökningen

A) Dokumentation

Mitt första intryck av rapporten är att den påminner om en omkullvält bokhylla. Genom att i den löpande texten hänvisa till tabellerna längst bak i rapporten, tvingas jag att bläddra fram och tillbaka för att få en vettig överblick. Rapporten saknar en innehålls-

förteckning, vilket yt-terligare förstärker intrycket av rörighet. Författaren blandar "generella" med "exakta" redovisningar. I en sammanfattning kan det vara på sin plats att skriva "*En av sex är chefsöverläkare.....*", men att sedan blanda det med procentsiffror gör det svårlöst "*En av sex (17 procent) är chefsöverläkare.....*". Tabellerna kunde ha delats upp och lagts in i den löpande texten, vilket skulle ha gett en mer lättläst rapport. I en bilaga finns en teknisk beskrivning av hur undersökningen har genomförts, samt en kommentar om statistisk osäkerhet i allmänhet. Men i de tabeller som författaren redovisar, saknar jag kommentarer om eventuella osäkerheter i skattningarna i den här undersökningen. Om författaren ändå publicerar så omfattrika tabeller, hade det varit på sin plats att redovisa konfidensintervall och känslighetsanalyser (av de grupper som varit föremål för en totalundersökning).

En stor brist i dokumentationen är att det inte förs en diskussion vilka eventuella vinster som metoden att kombinera en total- och urvalsundersökning kan ge. Genom att göra en totalundersökning av delgrupper, som är små i förhållande till större grupper, kan en ökad precision i skattningarna erhållas i det fall man gör det på ett förnuftigt sätt. I dokumentationen saknas all kommentar om varför man gjort på det sättet. Mitt intryck är att författaren inte tagit hänsyn till finessen med denna typ av design.

Den grafiska utformningen av rapporten gör den också svårläst. Valet av typsnitt och storlek på typerna gör att jag har svårt att läsa den.

Det är svårt för mig som läsare, att få en uppfattning om utvärderingens tillförlitlighet.

B) Data

Det finns en bilaga till rapporten, som kort redovisar population, urval, datainsamlingsmetod, period för datainsamling samt svarsfrekvens och bortfall. I den redovisningen saknar jag uppgifter om storleken på den population som urvalet är draget ifrån, och hur stickprovsstorleken har bestämts. Ingen sammanställning ges om hur många individer som ingår i de tre yrkeskategorier som skall totalundersökas. Det saknas uppgifter om hur urvalet fördelat sig på yrkeskategorier och andra redovisningsgrupper. I tabell 8 ges en procentuell redovisning av de

svarandes fördelning på yrkesgrupper, men där redovisas inte heller något bortfall. Jag saknar en diskussion om hur de 422 personer, vilka bortföll från det första urvalet, (antingen för att de var dubletter eller inte skulle ingå i undersökningspopulationen), får för konsekvenser på tillförlitligheten i resultaten.

Datansamligen har till stor del utförts under semesterperioden, hur kan det ha påverkat svarsfrekvensen. I redovisningen har inte rapporterats om det kan finnas eventuella "semestereffekter" vid valet av intervjuperioder. Det hade varit intressant att se om svarsfrekvensen skilde sig mellan de två undersökningsperioderna. Ingen kommentar lämnas om strategin varit att försöka få tag i personer som varit lediga första perioden under den andra för att på så sätt minska bortfallet.

Bristerna medför att jag har svårt, som utomstående att få en klar bild av tillförlitligheten i datamaterialet.

C) *Metod*

Metoden att använda både en total- och en stickprovsundersökning är lämplig. Framför allt i en situation där vissa redovisningsgrupper är små, och där man har en klar bild över vilka som ingår i totalpopulationen. Författaren har tycks inte dragit nytta av det i analysen.

I avsnittet där intervjupersonerna skall betygsätta föreskrifterna i kvalitetsförbättrande effekt på sjukvården, har författaren använt sig av ett medelbetyg (medelvärde). Jag skulle ha redovisat resultaten i form av medianvärde (och eventuellt kvartiler). Orsaken är att betygsättningen inte kan betraktas som en kvotskala, eventuellt välvilligt som en intervallskala, men framförallt en rangordning. Detta eftersom intervjupersonerna ombeds att rangordna föreskrifternas betydelse för kvalitetssäkring på en skala från *ingen betydelse* – *stor betydelse*. Därför är medelvärdet ett olämpligt mått i detta sammanhang. Att använda medelvärde, t.ex. om motivet är att skatta antalet barn i en befolkning, är helt i sin ordning, men att redovisa ett medelvärde när mätningen ger upphov till en rangordning är olämpligt. Det kan ge läsaren av rapporten en bild av falsk noggrannhet.

D) Slutsats

Genom att inte ta hänsyn till (i alla fall som jag som läsare uppfattar det), kombinationen av total- och stickprovsundersökning, samt utelämna en diskussion av bortfallets betydelse för resultaten är det svårt att värdera resultaten av undersökningen. I redovisningen av olika undergrupper kunde författaren ha kommenterat eventuella vinster i skattningarnas precision där hänsyn tagits till de två olika urvalsmetoderna. Om det till exempel var 40 vårdhemsläkare i totalpopulationen och endast 20 svarade, kan den verkliga andelen som sett föreskrifterna (i redovisningens tabell 1), variera mellan 50 och 100 procent om en känslighetsanalys redovisats.

På frågan om man som *arbetsledare eller chef* har informerat sin personal om innehållet i föreskrifterna, har man redovisat att 38 procent har gjort det. Längre fram i den redovisningen talas det om att 65 procent av arbetsledarna/cheferna har informerat sin personal. I detta sammanhang skulle de personer som inte var arbetsledare utslutits från redovisningen. Därmed blir hela redovisningen i tabell 4 och motsvarande stycke i rapporten missvisande. Samma sak gäller slutsatsen om Göteborg, där omräknat 37 procent har informerat om föreskrifterna, medan det i Skaraborg var 77 procent av arbetsledarna som informerat sin personal.

E) Kostnadseffektivitet

Valet av telefonintervju som datainsamlingsmetod, har troligen bidragit till den höga svarsfrekvensen (84 eller 75 procent beroende på hur man räknar). Min erfarenhet är att brevenkäter inte givit samma svarsfrekvens, framförallt i den population som undersökts här.

Då jag vet vad Socialstyrelsen har betalat SKOP för den här undersökningen, är det för mig svårt att kommentera kostnaden. Men jag har en känsla av att det borde vara möjligt att producera en bättre redovisning av undersökningen till samma pris, som Socialstyrelsen betalat. Framför allt gäller det rapportens grafiska utformning, bearbetningen och redovisningen av datamaterialet.

F) Objektivitet

Frågeställningarna tycker jag är klara och tydliga. Tillvägagångssättet anser jag vara lämpligt, med att kombinera en total- och stickprovsundersökning. Tolkningen och redovisningen av undersökningens resultat kunde göras bättre.

4. Sammanfattande bedömning och ytterligare kommentarer

Vem vänder sig undersökningsrapporten till? Man kan tänka sig att arbeta "quick and dirty", i fall att det enbart riktar sig till de närmast sörjande och dessa är medvetna om de antaganden som ligger till grund för undersökningen. Men är avsikten att vända sig till en större publik, exempelvis personer, vilka arbetar med kvalitetssäkringsfrågor inom landsting, är innehållet i rapporten inte så fullständig som jag skulle vilja ha det. Om den varit gjord av en av mina elever på momentet undersökningsmetodik, på en A-kurs i statistik, hade den fått gå en vända till innan jag som lärare godkänt rapporten.

Jag anser att de största svagheterna med utvärderingen är följande:

- Svårt att följa med i resultatredovisningen.
- Ingen redovisning av eventuella vinster med att kombinera total- och stickprovsundersökningen.
- Bortfallsanalysen saknas till stor del.
- Den grafiska utformningen.
- Felaktig bearbetning av frågeställningen om man informerat sin personal (tabell 4).
- Användning av ett olämpligt mått vid redovisningen av tabell 7, medelvärde i stället för median.

Bedömning 2

3. Utvärderingens syfte

Utvärderingens syfte framgår inte av rapportens huvudtext. Socialstyrelsens (SoS) uppdrag till SKOP har inte heller bilagts utredningen. Sammantaget innebär detta att det bara går att göra antaganden om syftet med undersökningen (författarens egen benämning).

Det är uppenbart att undersökningen syftar till att ge svar på om SoS:s föreskrifter om kvalitetssäkring i hälso- och sjukvården, inkl tandvården, har nått ut till sin tänkta målgrupp. Sannolikt syftar den också till att svara på om föreskrifterna har haft betydelse för om kvalitetssäkring kommit till stånd eller ej.

Det är mer osäkert om den också syftar till att utvärdera om föreskrifterna har givit någon vägledning vid kvalitetssäkringen samt om föreskrifterna har haft betydelse för utformningen av kvalitetssäkringen. Det är således osäkert om undersökningen överhuvudtaget har varit tänkt att vara en utvärdering. Dessa oklarheter påverkar möjligheten att bedöma undersökningen.

4. Metod

Undersökningen är en intervjuundersökning med 2 908 personer. Ett totalurval gjordes i grupperna medicinskt ansvariga sjuksköterskor, vårdhemsläkare samt tandläkare med ledningsansvar inom folktandvården. I övriga grupper – chefsöverläkare, legitimerade privatpraktiker samt yrkesverksamma legitimerade läkare, tandläkare, sjukgymnaster och psykologer som inte är privatpraktiserande – gjordes ett urval som ombesörjdes av SoS. Bortfallet var 16 procent.

Intervjuerna genomfördes av SKOP:s intervjucentral i Stockholm. Av det och av rapporten kan man dra slutsatsen att intervjuerna var mycket standardiserade. Bortsett från möjligheten att redovisa exempel på kvalitetssäkring förekom vad jag förstår inga öppna frågor.

5. Bedömning

A) Dokumentation

Redan av det ovanstående framgår att rapporten lämnar mycket övrigt att önska vad gäller dokumentation.

- Uppdrag och syfte framgår inte, inte heller vilket krav på dokumentation som ställts av SoS. Om arbetet skett i nära samarbete med SoS, och rapporten inte är tänkt att läsas av andra än beställaren, utgör möjligen detta inget problem. Om rapporten däremot är tänkt att läsas av andra, t ex bilagd en fördjupad anslagsframställning, så hade det varit önskvärt att syftet med undersökningen hade framgått.
- Metoden redovisas kortfattat i bilaga. Av den framgår population, urval, datainsamlingsmetod, datainsamlingsperiod, svarsfrekvens och bortfall. Redovisningen är dock rudimentär. Vad gäller urvalet framgår t ex inte urvalens storlek i förhållande till hela populationen för varje yrkesgrupp. Det framgår inte heller vilken typ av intervjuundersökning som genomförts. Har de t ex varit mycket bunden eller öppen. Det framgår inte heller hur stort bortfallet varit inom de olika yrkeskategorierna.
- Resultaten av undersökningen presenteras huvudsakligen i tabeller. Dessa sammanfattas mycket kortfattat i text. Varken tabeller eller text ger, med något undantag, någon vägledning till hur resultaten kan tolkas. T ex anges i tabellerna antalet svar för varje yrkesgrupp, arbetsområde eller län. Detta utan att det framgår hur många i varje yrkesgrupp, arbetsområde eller län som ingått i urvalet. Det gör uppgiften om antalet svar svår att värdera. Författaren drar inga egna slutsatser av undersökningen.
- I en bilaga till rapporten redovisas en tabell som visar hur stor en observerad skillnad mellan två grupper minst bör vara för att man med 95 procents säkerhet skall betrakta den som en verklig skillnad. I rapportens huvudtext kommenteras dock inte resultatens säkerhet. Läsaren får själv söka värdera denna med stöd av tabellen.
- På sid 5 i rapporten hänvisas till en bilaga 2 med exempel på kvalitetssäkring. Bilagan finns inte med i den bedöma rapporten. Det

är därför svårt att säga vilket värde den skulle kunna haft. Innehållet i bilagan kommenteras eller värderas inte alls i rapporten.

Förteckning över ESO:s tidigare publicerade rapporter

Expertgruppen för studier i offentlig ekonomi (ESO) är en kommitté (B 1981:03) under Finansdepartementet. Dess huvuduppgift är att bredda och fördjupa underlaget för budgetpolitiska och samhällsekonomiska avgöranden. En särskild uppgift är att belysa frågor om den offentliga sektorns effektivitet och produktivitet. Expertgruppen arbetar främst genom att lämna uppdrag till myndigheter, institutioner och forskare. Rapporterna publiceras i departementsserien eller i annan lämplig form.

I Ds-serien har ESO tidigare utgivit följande rapporter

1982

- Perspektiv på besparingspolitiken (Ds B 1982:3)
- Inkomstfördelningseffekter av livsmedelssubventioner (Ds B 1982:7)
- Perspektiv på budgetunderskottet, del 1. Budgetunderskottens teori och politik. Statens budgetfinansiering och penningpolitiken (Ds B 1982:9)
- Offentliga tjänster på fritids-, idrotts- och kulturområdena (Ds B 1982:10)
- Ökad produktivitet i offentlig sektor – en studie av de allmänna domstolarna (Ds B 1982:11)

1983

- Staten och kommunernas expansion – några olika styrmedel (Ds Fi 1983:3)
- Enhetligt barnstöd? (Ds Fi 1983:6)
- Perspektiv på budgetunderskottet, del 2. Fördelningseffekter av budgetunderskott. Hushållsekonomi och budgetunderskott (Ds Fi 1983:7)
- Minskad produktivitet i offentlig sektor – en studie av patent- och registreringsverket (Ds Fi 1983:18)
- Driver subventioner upp kostnader? – prisbildningseffekter av statligt stöd (Ds Fi 1983:19)
- Administrationskostnader för några transfereringar (Ds Fi 1983:22)
- Generellt statsbidrag till kommuner – modellskisser (Ds Fi 1983:26)
- Produktivitet i privat och offentlig tandvård (Ds Fi 1983:27)
- Perspektiv på budgetunderskottet, del 3. Budgetunderskott, portföljval och tillgångsmarknader. Modellsimuleringar av offentliga besparingar m.m. (Ds Fi 1983:29)
- Fördelningseffekter av kommunal barnomsorg (Ds Fi 1983:30)
- Administrationskostnader för våra skatter (Ds Fi 1983:32)

1984

- Vem utnyttjar den offentliga sektorns tjänster? (Ds Fi 1984:2)
- Perspektiv på budgetunderskottet, del 4. Budgetunderskott, utlandsupplåning och framtida konsumtionsmöjligheter. Budgetunderskott, efterfrågan och inflation (Ds Fi 1984:3)
- Konstitutionella begränsningar i riksdagens finansmakt – behov och tänkbara utformningar (Ds Fi 1984:7)
- Är subventioner effektiva? (Ds Fi 1984:8)
- Marginella expansionsstöd – ekonomiska och administrativa effekter (Ds Fi 1984:12)
- Transfereringar och inkomstskatt samt hushållens materiella standard (Ds Fi 1984:17)
- Parlamentet och statsutgifterna – hur finansmakten utövas i nio länder (Ds Fi 1984:18)

- Återkommande kostnads- och prestationsjämförelser – en metod att främja effektiviteten i offentlig tjänsteproduktion (Ds Fi 1984:19)

1985

- Statsskuldräntorna och ekonomin – effekter på inkomst- och förmögenhetsfördelningen samt på den samlade efterfrågan i samhället (Ds Fi 1985:2)
- Produktions-, kostnads- och produktivitetens utveckling inom offentligt bedrivna hälso- och sjukvård 1960–1980 (Ds Fi 1985:3)
- Produktions-, kostnads- och produktivitetens utveckling inom den sociala sektorn 1970–1980 (Ds Fi 1985:4)
- Transfereringar mellan den förvärsarbetande och den äldre generationen (Ds Fi 1985:5)
- Frivilligorganisationer – alternativ till den offentliga sektorn? (Ds Fi 1985:6)
- Organisationer på gränsen mellan privat och offentlig sektor – förstudie (Ds Fi 1985:7)
- Produktions-, kostnads- och produktivitetens utveckling inom vägsektorn (Ds Fi 1985:9)
- Skatter och arbetsutbud (Ds Fi 1985:10)
- Sociala avgifter – problem och möjligheter inom färdtjänst och hemtjänst (Ds Fi 1985:11)
- Egen regi eller entreprenad i kommunal verksamhet – möjligheter, problem och erfarenheter (Ds Fi 1985:12)

1986

- Produktions-, kostnads- och produktivitetens utveckling inom armén och flygvapnet (Ds Fi 1986:1)
- Samhällsekonomiskt beslutsunderlag – en hjälp att fatta beslut (Ds Fi 1986:2)
- Effektivare sjukvård genom bättre ekonomistyrning (Ds Fi 1986:3)
- Effekter av statsbidrag till kommuner (Ds Fi 1986:7)
- Byråkratiseringstendenser i Sverige (Ds Fi 1986:8)
- Svensk inkomstfördelning i internationell jämförelse (Ds Fi 1986:12)
- Offentliga tjänster – sökarljus mot produktivitet och användare (Ds Fi 1986:13)
- Kostnader och resultat i grundskolan – en jämförelse av kommuner (Ds Fi 1986:14)
- Regleringar och teknisk utveckling (Ds Fi 1986:15)
- Socialbidrag. Bidragstagarna: antal och inkomster. Socialbidragen i bidragssystemet (Ds Fi 1986:16)
- Produktions-, kostnads- och produktivitetens utveckling inom den offentligt finansierade utbildningssektorn 1960–1980 (Ds Fi 1986:17)
- Offentliga utgifter och sysselsättning (Ds Fi 1986:29)

1987

- Att leva på avgifter – vad innebär en övergång till avgiftsfinansiering? (Ds Fi 1987:2)
- Vågar ut ur jordbruksprisregleringen – några idéskisser (Ds Fi 1987:4)
- Kvalitetsutvecklingen inom den kommunala äldreomsorgen 1970–1980 (Ds Fi 1987:6)
- Produktkostnader för offentliga tjänster – med tillämpningar på kulturområdet (Ds Fi 1987:10)
- Integrering av sjukvård och sjukförsäkring (Ds Fi 1987:11)

1988

- Kvalitetsutvecklingen inom den kommunala barnomsorgen (Ds 1988:1)
- Från patriark till part – spelregler och lönepolitik för staten som arbetsgivare (Ds 1988:4)
- Produktivitetens utvecklingen i kommunal barnomsorg 1981–1985 (Ds 1988:5)
- Prestationer och belöningar i offentlig förvaltning (Ds 1988:18)
- Subventioner i kritisk belysning (Ds 1988:28)

- Hur stor blev tvåprocentaren? Erfarenheter från en besparingsteknik (Ds 1988:34)
- Effektiv realkapitalanvändning i kommuner och landsting (Ds 1988:51)
- Alternativ i jordbrukspolitiken (Ds 1988:54)
- Kvalitet och kostnader i offentlig tjänsteproduktion (Ds 1988:60)
- Vad kan vi lära av grannen? Det svenska pensionssystemet i nordisk belysning (Ds 1988:68)

1989

- Hur man mäter sjukvård – exempel på kvalitets- och effektivitetsmätning (Ds 1989:4)
- Lönestrukturen och den "dubbla obalansen" – en empirisk studie av löneskillnader mellan privat och offentlig sektor (Ds 1989:8)
- Beställare-utförare – ett alternativ till entreprenad i kommuner (Ds 1989:10)
- Vad ska staten äga? De statliga företagen inför 90-talet (Ds 1989:23)
- Statsbidrag till kommuner: allt på en check eller lite av varje? En jämförelse mellan Norge och Sverige (Ds 1989:26)
- Produktivitetsmätning av folkbibliotekens utlåningsverksamhet (Ds 1989:42)
- Bostadsstödet – alternativ och konsekvenser (Ds 1989:47)
- Kommunal förmögenhetsförvaltning i förändring: citykommunerna Stockholm, Göteborg och Malmö (Ds 1989:56)
- Hur ska vi få råd att bli gamla? (Ds 1989:59)
- Arbetsmarknadsförsäkringar (Ds 1989:68)

1990

- Bostadskarriären som förmögenhetsmaskin (Ds 1990:29)
- Skola? Förskola? Barnskola! (Ds 1990:31)
- Statens dölda kapital. Aktivt ägande: exemplet Vattenfall (Ds 1990:36)
- Sjukvårdskostnader i framtiden – vad betyder åldersfaktor? (Ds 1990:39)
- Läkemedelsförmånen (Ds 1990:81)

1991

- Målstyrning och resultatuppföljning i offentlig förvaltning (Ds 1991:19)
- Metoder i forskning om produktivitet och effektivitet med tillämpningar på offentlig sektor (Ds 1991:20)
- Vad kostar det? Prislista för statliga tjänster (Ds 1991:26)
- Det framtida pensionssystemet – två alternativ (Ds 1991:27)
- Skogspolitik för ett nytt sekel (Ds 1993:31)
- Prestationsbaserad ersättning i hälso- och sjukvården – vad blir effekterna? (Ds 1991:49)
- Ostyriga projekt – att styra och avstyra stora kommunala satsningar (Ds 1991:50)
- Marginaleffekter och tröskleffekter – barnfamiljerna och barnomsorgen (Ds 1991:66)
- S.J, Televerket och Posten – bättre som bolag? (Ds 1991:77)

1992

- Skatteförmåner och särregler i inkomst- och mervärdskatten (Ds 1992:6)
- Frihandeln ett hot mot miljöpolitiken – eller tvärtom? (Ds 1992:12)
- Växthuseffekten – slutsatser för jordbruks-, energi- och skattepolitiken (Ds 1992:15)
- Fattigdomsfällor (Ds 1992:25)
- Vad vill vi med socialförsäkringarna? (Ds 1992:26)
- Statliga bidrag – motiv, kostnader, effekter? (Ds 1992:46)
- Hur bra är vi? Den svenska arbetskraftens kompetens i internationell belysning (Ds 1992:83)

- Slutbudsmetoden – ett sätt att lösa tvister på arbetsmarknaden utan konflikt (Ds 1992:88)
- Kommunerna som företagsägare – aktiv koncernledning i kommunal regi (Ds 1992:111)
- Press och ekonomisk politik – tre fallstudier (Ds 1992:124)
- Statsskulden och budgetprocessen (Ds 1992:126)

1993

- Presstödet effekter – en utvärdering (Ds 1993:20)
- Hur välja rätt investeringar i transportinfrastrukturen? (Ds 1993:22)
- Lönar sig förebyggande åtgärder? Exempel från hälso- och sjukvården och trafiken (Ds 1993:37)
- Social Security in Sweden and Other European Countries – Three Essays (Ds 1993:51)
- Idrott åt alla? – Kartläggning och analys av idrottsstödet (Ds 1993:58)

1994

- Att rädda liv – Kostnader och effekter (Ds 1994:14)
- Varför kulturstöd? – Ekonomisk teori och svensk verklighet (Ds 1994:16)
- Kvalitets- och produktivitetens utvecklingen i sjukvården 1960–1992 (Ds 1994:22)
- Kvalitet och produktivitet – Teori och metod för kvalitetsjusterade produktivitetmått (Ds 1994:23)
- Den offentliga sektorns produktivitetens utveckling 1980–1992 (Ds 1994:24)
- Det offentliga stödet till partierna – Inriktning och omfattning (Ds 1994:31)
- Den svenska insolvensrätten – Några förslag till förbättringar inom konkurshanteringen m.m. (Ds 1994:37)
- Budgetunderskott och statsskuld – Hur farliga är de? (Ds 1994:38)
- Bensinskatteförändringars effekter (Ds 1994:55)
- Skolans kostnader, effektivitet och resultat – En branschstudie (Ds 1994:56)
- Den offentliga sektorns produktivitetens utveckling 1980–1992. Bilagor (Ds 1994:71)
- Valfrihet inom skolan – Konsekvenser för kostnader, resultat och segregation (Ds 1994:72)
- En Social Försäkring (Ds 1994:81)
- Fördelningseffekter av offentliga tjänster (Ds 1994:86)
- Nettokostnader för transfereringar i Sverige och några andra länder (Ds 1994:133)
- Skatter och socialförsäkringar över livsrytmen – En simuleringsmodell (Ds 1994:135)
- En effektiv försvarspolitik? – Fredsvinst, beredskap och återtagning (Ds 1994:138)

1995

- Försvarets kostnader och produktivitet (Ds 1995:10)
- Företagsstödet – Vad kostar det egentligen? (Ds 1995:14)
- Hushållning med knappa naturresurser – Exemplet allemansrätten, fjällen och skotertrafik i naturen (Ds 1995:15)
- Vad blev det av de enskilda alternativen? En kartläggning av verksamheten inom skolan, vården och omsorgen (Ds 1995:25)
- Kostnader, produktivitet och måluppfyllelse för Sveriges Television AB (Ds 1995:31)
- Hushållning med knappa naturresurser – Exemplet sportfiske (Ds 1995:47)
- Invandring, sysselsättning och ekonomiska effekter (Ds 1995:68)
- Generationsräkenskaper (Ds 1995:70)
- Kapitalets rörlighet – Den svenska skatte- och utgiftsstrukturen i ett integrerat Europa (Ds 1995:74)

1996

- Hur effektivt är EU:s stöd till forskning och utveckling? – En principdiskussion (Ds 1996:8)
- Reglering som spel – Universiteten som förebild för offentliga sektorn? (Ds 1996:18)
- Nästa steg i telepolitiken (Ds 1996:29)

I andra serier utgivna rapporter m.m.

I följande förteckning ingår dock inte de rapporter som publicerats som bilagor till långtidsutredningarna.

Tidigare publicerat

- Besparingar genom avregleringar (RRV, 1982, Dnr 1982:999)
- Vem skall betala jordbrukets rationalisering? (Statskontorets småskrifter 9, 1983)
- Inför omprövningen. Alternativ till dagens socialförsäkringar (Liber Förlag 1983)
- Statsförvaltningen behöver nya organisationsformer – förstudie (RRV, 1984, Dnr 1983:18)
- Kostnader för offentliga tjänster i Norden. KRON-projektet (Statskontoret 1983:18)
- Hur stor är den offentliga sektorn? Johan A. Lybeck (Liber Läromedel, 1984)
- Varför blev de dyrare? Kostnadsutvecklingen för statliga reformer (RRV, 1984. Dnr 1984:334)
- Erfarenheter av stora omorganisationer, styrning – genomförande (Statskontoret 1985:4)
- Erfarenheter av stora omorganisationer, tre fallstudier (Statskontoret 1985:5)
- Statliga tjänsteproduktion – produktivitetens utveckling 1960–80 (Statskontoret 1985:15)
- Långa handläggningstider i offentlig verksamhet (RRV 1985, Dnr 1984:695)
- Kvalitetssäkring – att mäta, värdera och utveckla sjukvårdens kvalitet (Spri-rapport 230, 1987)
- FoU – en resurs för utveckling av offentliga tjänster? En studie av lokalt utvecklingsarbete inom kriminalvård och barnomsorg (Statskontoret 1989:39)
- En "skattereform" för socialförsäkringar? Red. Klas Eklund (Publica, 1993)
- I fädrens spår... "Den svenska infrastrukturens historiska utveckling och framtida utmaningar" Arne Kaijser (Carlssons 1994)
- Social Security in Sweden – How to Reform the System
Engelsk sammanfattning av En Social Försäkring (Ds 1994:81)
- Riv skolan, börja lära! Reformrörelser i USA. Sven Wiberg (SNS Förlag 1995)
- The Mobility of Capital – The Swedish Tax and Expenditure Structure in an Integrated Europe
Engelsk översättning av Kapitalets rörlighet – Den svenska skatte- och utgiftsstrukturen i ett integrerat Europa (Ds 1995:74)

Departementsserien 1996

Kronologisk förteckning

1. Effektivare ekobrottsbekämpning. Ju.
 2. Underhållsstöd till barn till särlevande föräldrar. S.
 3. The Swedish Marketing Act. C.
 4. Regeringens och myndigheternas befogenheter vid svåra påfrestningar på samhället. Fö.
 5. Distributionsstödet till dagspressen. Ku.
 6. Ersättning vid dammbrott. Ju.
 7. Romkonventionen – införlivande med svensk rätt av EG-konventionen om tillämplig lag för avtalsförpliktelser. Ju.
 8. Hur effektivt är EU:s stöd till forskning och utveckling? – En principdiskussion. Fi.
 9. Svenska författningar i översättning till främmande språk. En förteckning. SB.
 10. Beskattning av vissa EU-bidrag. Fi.
 11. Rekrytering och grundutbildning av poliser. Ju.
 12. Förutsättningar för avveckling av Bromma flygplats. K.
 13. Genetisk integritet – vem har rätt att använda information från genetiska undersökningar? S.
 14. Värdering av kvinnors respektive mäns meriter vid tjänsteställning inom universitet och högskolor. U.
 15. Säkerhet och utveckling i Afrika. UD.
 16. Lärarutbildning i förändring. U.
 17. Ändrad grund för förvaltningen av vissa statligt reglerade bostadslån. N.
 18. Reglering som spel – Universiteten som förebild för offentliga sektorn? Fi.
 19. Lag om kriminalvårdsregister. JU.
 20. Ändringar i lotterilagen. C.
 21. Forskning om offentlig sektor. Bättre kontakter mellan forskare och praktiker. C.
 22. Ökat inflytande för allmänheten vid prövning av tillstånd enligt kärntekniklagen. M.
 23. Remissvar på Kulturutredningen. Ku.
 24. Vissa företagsskattefrågor. Fi.
 25. Jämkning av skadestånd enligt trafikskadelagen, trafikförsäkringsavgifter m.m. Ju.
 26. Genusperspektiv i forskningen. U.
 27. Beskattningen av utländska kapitalförsäkringar. Fi.
 28. Remissammanställning. Kvinnofrid (SOU 1995:60). S.
 29. Nästa steg i telepolitiken. Fi.
 30. Kommunal redovisning. Fi.
 31. Ungdomsfrågor i EU. Förslag till en svensk strategi. In.
 32. Förslag till lag om tillsyn över hundar och katter. Jo.
 33. Förbud mot rasistiska symboler m.m. Ju.
 34. Beskattning av bilförmån. En ny modell. Fi.
 35. Studiefinansiering och examina i forskarutbildningen. U.
 36. Kan myndigheter utvärdera sig själva? Fi.
-

Departementsserien 1996

Systematisk förteckning

Statsrådsberedningen

Svenska författningar i översättning till främmande språk. [9]

Justitiedepartementet

Effektivare ekobrottsbekämpning. [1]
Ersättning vid dammbrott. [6]
Romkonventionen – införlivande med svensk rätt av EG-konventionen om tillämplig lag för avtalsförpliktelser. [7]
Rekrytering och grundutbildning av poliser. [11]
Lag om kriminalvårdsregister. [19]
Jämkning av skadestånd enligt trafikskadelagen, trafikförsäkringsavgifter m.m. [25]
Förbud mot rasistiska symboler m.m. [33]

Utrikesdepartementet

Säkerhet och utveckling i Afrika. [15]

Försvarsdepartementet

Regeringens och myndigheternas befogenheter vid svåra påfrestningar på samhället. [4]

Socialdepartementet

Underhållsstöd till barn till särlevande föräldrar. [2]
Genetisk integritet – vem har rätt att använda information från genetiska undersökningar? [13]
Remissammanställning.
Kvinnofrid (SOU 1995:60). [28]

Kommunikationsdepartementet

Förutsättningar för avveckling av Bromma flygplats. [12]

Finansdepartementet

Hur effektivt är EU:s stöd till forskning och utveckling? – En principdiskussion. [8]
Beskattnings av vissa EU-bidrag. [10]
Reglering som spel – Universiteten som förebild för offentliga sektorn? [18]
Vissa företagsskattefrågor. [24]
Beskattnings av utländska kapitalförsäkringar. [27]
Nästa steg i telepolitiken. [29]
Kommunal redovisning. [30]
Beskattnings av bilförmån. En ny modell. [34]
Kan myndigheter utvärdera sig själva? [36]

Utbildningsdepartementet

Värdering av kvinnors respektive mäns meriter vid tjänsteställning inom universitet och högskolor. [14]
Lärarytelse i förändring. [16]
Genusperspektiv i forskningen. [26]
Studiefinansiering och examina i forskarutbildningen. [35]

Jordbruksdepartementet

Förslag till lag om tillsyn över hundar och katter. [32]

Kulturdepartementet

Distributionsstödet till dagspressen. [5]
Remissvar på Kulturutredningen. [23]

Näringsdepartementet

Ändrad grund för förvaltningen av vissa statligt reglerade bostadslån. [17]

Civildepartementet

The Swedish Marketing Act. [3]
Ändringar i lotterilagen. [20]
Forskning om offentlig sektor. Bättre kontakter mellan forskare och praktiker. [21]

Inrikesdepartementet

Ungdomsfrågor i EU.
Förslag till en svensk strategi. [31]

Miljödepartementet

Ökat inflytande för allmänheten vid prövning av tillstånd enligt kärntekniklagen. [22]
